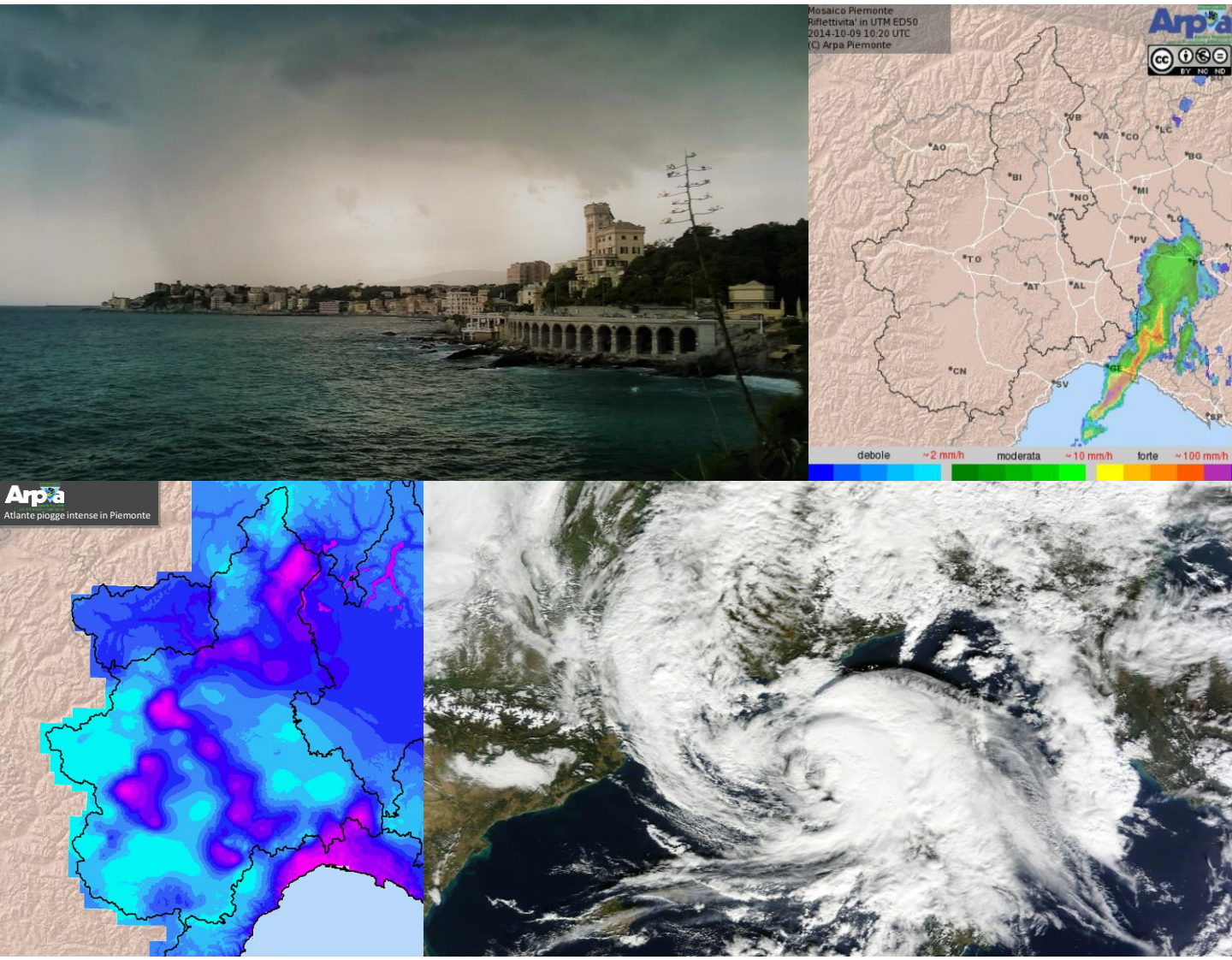
Advances in the space-time analysis of rainfall extremes

ScuDo
Scuola di Dottorato ~ Doctoral School
WHAT YOU ARE, TAKES YOU FAR

29th cycle – Doctoral Program in Environmental Engineering

Doctoral Dissertation



Advances in the space-time analysis of rainfall extremes

Andrea Libertino

ScuDo
Scuola di Dottorato ~ Doctoral School
WHAT YOU ARE, TAKES YOU FAR

Doctoral Dissertation
Doctoral Program in Environmental Engineering (29^{th}cycle)

Advances in the space-time analysis of rainfall extremes

By

Andrea Libertino

******

Supervisor(s):
Prof. Pierluigi Claps, Supervisor
Ing. Paola Allamano, Co-Supervisor

Doctoral Examination Committee:
Prof. Marco Marani , Referee, Università degli Studi di Padova, Padova, IT
Prof. Ramesh Teegavarapu, Referee, Florida Atlantic University, Boca Raton, US
Prof. Roberto Deidda, Università degli Studi di Cagliari, Cagliari, IT
Ing. Salvatore Gabriele, CNR-IRPI Cosenza, Rende (CS), IT
Prof. Davide Poggi, Politecnico di Torino, Torino, IT
Prof. Roberto Rudari, CIMA Research Foundation, Savona, IT

Politecnico di Torino
May, 2017

Declaration

I hereby declare that, the contents and organization of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

Andrea Libertino
May, 2017

Cover images (clockwise from top left): The city of Genova, Andrea Libertino, 2015; Radar reflectivity mosaic for the Piemonte region 2014-10-09 10:20 UTC (c) ARPA Piemonte; MODIS Terra image 8-11-2011 10:30 UTC (c) MODIS-NASA-GSFC; Precipitation Atlas of the Piemonte region (c) ARPA Piemonte.

* This dissertation is presented in partial fulfillment of the requirements for **Ph.D.** degree in the Graduate School of Politecnico di Torino (ScuDo).

Ma verso la fine d'ottobre piovve in montagna e piovve in pianura,
il fiume Tanaro parve rizzarsi in piedi tanto crebbe.
La gente ci vide il dito di Dio [...]
Il fiume esagerò al punto che si smise d'aver paura della repubblica
per cominciare ad averne di lui.
Beppe Fenoglio, *I ventitrè giorni della città di Alba*, 1952.

Nu l'è l'aaegua de 'na rammâ 'n calabà 'n calabà
Fabrizio de Andrè, *Dolcenera*, 1996.

To my family.

Acknowledgements

First, I want to thank my advisor Pierluigi Claps, for offering me countless opportunities for growing from both a scientific and a human perspective.

Thanks to Francesco Laio, Paola Allamano, Daniele Ganora and Luca Ridolfi for supporting me throughout the course of my Ph.D. and to all the other professors, researchers and students that offered me interesting insights, encouraging me to develop my investigation in various directions. Particularly, thanks to Marco Marani and Ramesh Teegavarapu for their insightful comments that allowed to improve the quality of this manuscript, and to Secondo Barbero, Roberto Cremonini and Milena Zaccagnino for providing technical support in many fields of this work.

Thanks to Ashish Sharma and Venkat Lakshmi for advising me during my research period in Sydney and to all the staff of the Water Research Center of the Sydney UNSW for giving me the chance to live a unique enriching experience.

I owe a lot to all of my colleagues, for their friendship and their patience in helping me dealing with my "innate optimism" kriging the daily difficulties. Thanks for contributing immensely to my personal and professional growth.

Last but not the least, I would like to thank my family and my friends for supporting me every day of my life.

Data providers are aknowledged: CF Regione Valle d'Aosta, ARPA Piemonte, ARPAL-CFMI-PC, ARPA Lombardia, Meteotrentino, UI Provincia Autonoma di Bolzano, ARPAV, UI Regione Friuli Venezia Giulia, ARPA Emilia Romagna, SIR Toscana, DICA - Univesità degli Studi di Firenze, Regione Umbria, DPC Regione Marche,CF Regione Lazio, UIM Regione Abruzzo, CF Regione Molise, CF Regione Campania, DPC Regione Puglia, CFM-ARPACAL, DPC Regione Basilicata, Osservatorio delle Acque Regione Siciliana, ARPAS.

Abstract

Statistical estimation of design rainfall is considered a consolidated topic in hydrology. However, extreme rainfalls and their consequences still constitute one of the most critical natural risks worldwide, particularly in the urban environments. Additional efforts for improving the spatio-temporal analysis of extreme rainfall are then required, particularly at the regional scale. In this work, a new set of data and techniques for improving the spatial statistical analysis of extreme rainfall is proposed. Italy is considered a challenging case study, due to its specific geographic and orographic settings, associated with recurring storm-induced disasters.

At first, the rain-gauge data patchiness resulting from the evolution of the monitoring agencies and networks, is tackled with the "patched kriging" methodology. The technique, involving a sequential annual interpolation, provides complete annual maxima series consistent with the available data. This allows to extract all the information avaialble from the gauge records, considering also the information "hidden" in the shortest series, increasing the robustness of the results.

Interpolation techniques, however, can only reflect the estimation variance determined by the spatial and temporal data resolution. Additional improvements can be obtained integrating the rain gauge information with remote sensing products, able to provide more details on the spatial structure of rainstorms. In this direction, a methodology aimed at maximizing the efficiency of weather radar when dealing with large rainfall intensities is developed. It consists in a quasi-real-time calibration procedure, adopting confined spatial and temporal domains for an adaptive estimation of the relation between radar reflectivity and rainfall rate. This allows one to follow the well-known spatio-temporal variability of the reflectivity-rainfall relation, making the technique suitable for a systematic operational use, regardless of the local conditions. The methodology, applied in a comprehensive case study reduces

the bias and increases the accuracy of the radar-based estimations of severe rainfall intensities.

The field of the satellite estimation of preciptation is then explored, by analyzing the ability of both the Tropical Rainfall Measurement Mission (*TRMM*) and the recently launched Global Precipitation Measurement (*GPM*) mission to help identifying the timing of severe rainfall events on wide spatial domains. For each considered product, the date of occurrence of the most intense annual daily records are identified and compared with the ones extracted from a global rain-gauge database. The timing information can help in tracking the pattern of deep convective systems and support the identification of localized rainfall system in poorly gauged areas.

The last part of the work deals with the analysis of rainfall extremes at the country scale, with a particular focus on the most severe rainfall events occurred in Italy in the last century. Many of these events have been studied as individual case studies, due to the large recorded intensities and/or to their severe consequences, but they have been seldom expressly addressed as a definite population. To try to provide new insights in a data-drived approach, a comprehensive set of annual rainfall maxima has been compiled, collecting data from the different regional authorities in charge. The database represents the reference knowledge for extremes from 1 to 24 hours durations in Italy, and includes more than 4500 measuring points nationwide, with observation spanning the period 1916-2014. Exploratory statistical analyses for providing information on the climatology of extreme rainfall at the national scale are carried out and the stationarity in time of the highest quantiles is analysed by pooling up all the data for each duration together. The cumulative empirical distributions are explored looking for clues of the existence of a class of "super-extremes" with a peculiar statistical behavior. The analysis of the spatial the distribution of the records exceeding the 1/1000 overall empirical probability shows an interesting spatial clustering. However, once removed the influence of the uneven density of the rain gauge network in time and space, the spatial susceptibility to extraordinary events seems quite uniformly distributed at the country scale.

The analyses carried out provide quantitative basis for improving the rainstorm estimation in gauged and ungauged locations, underlining the need of further research efforts for providing maps for hydrological design with uniform reliability at the various scales of technical interest.

Table of Contents

List of Figures

List of Tables

List of Symbols

Roman Symbols

$\hat{R}$ Radar-estimated rainfall

c_1 Sill of a variogram

c_2 Range of a variogram

c_3 Nugget of a variogram

CV Coefficient of variation

d Timescale - duration

h Rainfall depth

h_d^* Annual maximum rainfall depth for duration d normalized on the median

h_d Annual maximum rainfall depth for duration d

$h_{d,0}$ Detrended annual maximum rainfall depth for duration d

i Rainfall intensity

K_T Growth factor with T return period

L Lag-distance

$L-CA$ Coefficient of L-skewness

$L-CV$ Coefficient of L-variation

$L-KUR$ Coefficient of L-kurtosis

l_r r order L-moment

P Cumulative probability

R Rainfall rate measured at-gauge

R^2 Coefficient of determination

Std Standard deviation

T Return period

$V(L)$ Sample variogram

Var Sample variance

Y Year

Z Radar reflectivity factor in mm^6/m^3

Z^* Radar reflectivity factor in decibels

Z^*_{th} Reflectivity threshold in ATS technique

z Elevation

Greek Symbols

$\gamma(L)$ Theoretical variogram

σ_i^2 Kriging variance associated to the i-th element of a cored series

θ_1 Location parameter

θ_2 Scale parameter

θ_3 Shape parameter

Subscripts

cor Cored series

$orig$ Original series

Acronyms / Abbreviations

$3IMERGHH$ Integrated Multi-satellitE Retrievals Half-Hourly - Version 3

asl Above Sea Level

ATS Adaptive in Time and Space technique

BM Block Maxima method

$CUBIST$ Characterization of Ungauged Basins by Integrated uSe of hydrological Techniques

DDF Depth-Duration-Frequency

DSD Drop Size Distribution

DTM Digital Terrain Model

$EV1$ Extreme Value Type 1 distribution, i.e., Gumbel distribution

GEV Generalized Extreme Value

GPM Global Precipitation Measurement

IDW Inverse Distance Weight

IR InfraRed

ISPRA Italian National Institute for Environmental Protection and Research

JAXA Japan Aerospace Exploration Agency

$KOSTRA$ KOmmune-STat-RApportering

$LOESS$ LOcally wEighted Scatter-plot Smoother

MCS Mesoscale Convective System

$MIUR$ Italian Ministry of Education and Research

NASA National Aeronautics and Space Administration

PMW Passive MicroWave

POD_{N_g} Probability Of Detection for a certain N_g

POT Peak Over Threshold method

PP Primary Peak

PR Precipitation Radar

QPE Probabilistic Quantitative Precipitation Estimates

RED (Italian) Rainfall Extremes Database

SIMN National Hydrographic and Mareographic Service

SIN National Hydrographic Service

SP Secondary Peak

TCEV Two Component Extreme Value Distribution

TMPA *TRMM* Multi-satellite Precipitation Analysis

TRMM Tropical Rainfall Measurement Mission

VAPI VAlutazione Piene in Italia

VPR Vertical Profile of Reflectivity

Chapter 1

Introduction

1.1 Foreword

Understanding and monitoring the development of severe rainstorms has always been a priority of the hydrological sciences, due to the significant consequences of extreme rainfalls on the safety of the people, the protection of the natural resources and the organization of the human activities. The Munich Re's NatCatSERVICE [139], one of the world's most comprehensive databases for analysing and evaluating natural catastrophes, shows that between 1980 and 2014 floods and storms have been the main natural disasters in the world, with overall losses of about 2800bn of US$. Adverse hydro-meteorological events (including floods, mass movements and storms) have led to 663000 fatalities, consisting in about 25000 fatalities per year across the globe, without considering the losses indirectly caused (due to famine, collapse of health systems, etc.). This entails additional research efforts, invoked from both the media and the scientific community. The society poses clear questions to the authorities responsible for the safeguard of the territory from floods, concerning the possibility of predicting these events and implementing measures for warning the population and preventing or limitating the damages.

The meteorological mechanism related to the genesis and development of extreme rainfall phenomena is quite well described in the literature for various climatic zones (e.g., [68, 86, 61, 170]). However, more research efforts are needed in order to transform this knowledge in the direction of building concrete and robust tools

for assessing and managing the risk associated to this kind of events on a wide heterogeneous spatial scale.

The aims of this dissertation stem from this considerations. In this work the meteorological aspects are only briefly mentioned, while the main focus is on the development of operational tools and instruments for improving the prediction of extreme rainfalls over large areas.

The spatial analysis of precipitation is a very discussed and evolving topic (e.g., [189, 129]). Rainfall is discontinuous in space and time. The spatialization of the information to ungauged areas, therefore, involves the need to assess both the occurrence and the amount of precipitation. Various problems arise if different time scales are considered and the accuracy of the spatialization methods also depends on other factors, like the type of precipitation, the triggering mechanism, etc. [66].

The framework is further complicated by the inadequacy of the rainfall measures [205]. Traditional rain gauge networks are characterized by a long temporal consistency and a high degree of accuracy but they are often not enough densely or evenly distributed for describing the spatial variability of convective rainfall systems. Coupling remote precipitation measures (e.g., satellite and radar) with gauge precipitation observations is one of the way to bridge the gap between the discrete information provided by the rain gauge network and the real spatial variability of the precipitation. However, these remote sensing tools do not provide a direct measure of the rainfall intensities, and are affected by significant errors and uncertainties (e.g., [2, 3]). Even with improved estimation accuracy, these instruments can not provide series long enough for comprehensive statistical analyses.

The occasional development of rainfall events in which very high intensities (e.g., intensities in the same order of magnitude of the average annual rainfall in the considered site) persist for several hours can add further complexity to the system. When dealing with the estimation of the probability of occurrence of these extraordinary storms, two more significant problem arise. The first is related to the rarity of these kind of events, that prevents a robust statistical approach for the frequency analysis. The second concerns the uneven spatial occurrence of this kind of phenomena, due to their localized scale, that further hampers the assessment of the probability of occurrence of similar amounts in space.

Stemming from the open problems presented above, the main goal of this dissertation is to improve rainstorm prediction on large areas, even in the presence of extraordinary rainfall amounts. The approaches and the methods adopted are data-driven and aim at defining robust statistical tools for analysing both ground and remotely sensed data. The detailed objectives of the work and the structure of thesis are described in section 1.7, to which the experts in this field are addressed. Less expert readers may proceed with the next sections, which provide some introductory concepts on the nature of the rainfall phenomena and on the probabilistic approach commonly adopted for their analysis, that further motivate the research efforts presented in this work.

1.2 Rainfall (extremes) phenomenology

Precipitation is the product of the condensation of atmospheric water vapor, which falls through the lower troposphere and hits the earth's surface, unless it evaporates completely before reaching the ground. According to the characteristics of temperature, pressure and humidity of the air masses, precipitation can occur in different forms: liquid, solid or mixed.

For precipitation to occur, three requirements are to be fulfilled:

- Condition of saturation, mostly due to the lifting of the air masses, is necessary.
- A phase change, from the gas state to the liquid or solid one, must occur.
- Raindrops and ice crystals must grow enough to overcome the updraft force and fall downward.

Three main mechanism (represented in figure 1.1) can induce the lifting of the air masses and cause the production of precipitation:

- Frontal lifting (a), resulting from the collision of air masses at different temperature, usually associated with cyclonic circulation.
- Orographic lifting (b), occurring when air masses are forced to rise over orographic barriers, inducing cooling and precipitation production.

- Convective lifting (c), consisting in strong vertical air motions, induced by the temperature gradient between the ground surface and the air masses.

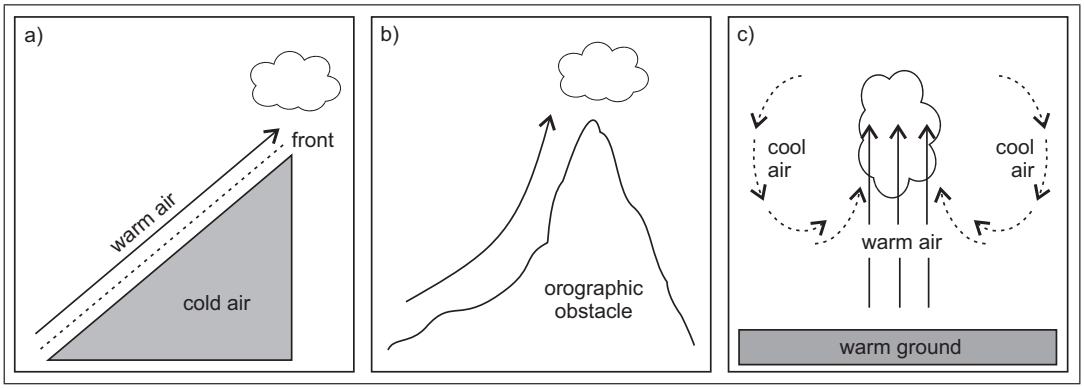


Fig. 1.1 (a) Frontal lifting, (b) orographic lifting and (c) convective lifting.

In the literature, rainfall phenomena are usually classified in stratiform and convective [93, 92]. The first generally present wider spatial scales and lower intensities, that persist for long durations. The second are generated from clouds with significant vertical development. Typical spatial scales vary from several km to some hundreds of kilometers for the *MCS* (Mesoscale Convective System [125]). Rainfall intensities are generally larger and durations in the order of one or few hours.

In the Mediterranean area two main kind of events dominate: frontal events (e.g., [152, 199]), that constitute the predominant meteorological structure, producing ordinary rainfall events on quite wide areas (10^3-10^4 km^2); and isolated convective events (e.g., [209, 111]), short-lived systems, generally characterized by large intensities on a 10-100 km^2 spatial scale. Furthermore, the orographic lift can play an important role in the genesis of rainfall events in the region, especially in complex orography areas (Alps, Appennines, Pyrenees, etc.). The interaction between these structures, under particular morpho-climatic settings, can lead to the genesis of hybrid and complex rainfall systems. Mesoscale Convective Systems and multicell storms [111] are examples of these kind of phenomena: in these cases the interaction between the different spatio-temporale scales lead to persistent severe rainfall.

In addition to the above meteorological structures, recent studies (e.g., [160]) have introduced the Mediterranean hurricanes, or *medicane* [197, 198] category. These events are occasional, but are able to produce extraordinary rainfall amounts

over 10^2-10^3 km^2 areas and present characteristics similar to the tropical cyclones. Motivations for a deep and wide analysis of rainfall extremes are provided by the lack of a statistical description of those exceptional rainstorms.

1.3 Probabilistic modeling of extreme rainfall

Different sources are commonly adopted for collecting rainfall data: from the traditional rain gauge networks, able to provide precise but local rainfall measurements, to the emerging remote sensing instruments, able to span wider spatial domains with a good resolution. Independently from the sources, once data are collected, a number of theories are commonly adopted for merging and modeling rainfall extremes, for different fields and purposes.

Probabilistic modeling of precipitation extremes consists of methods for analyzing the characteristics of the extreme rainfall regime, by applying frequency analysis techniques. Probabilistic modeling of extreme rainfall has a crucial role in flood risk estimation and consequently in the design and management of flood mitigation projects [113].

The first attempts to establish a mathematical relation between the intensity and the frequency of rainfall goes back to as early as 1932 [25]. Since then, many studies have been carried out, aimed at relating rainfall depths with a certain probability of occurrence. Among the most recent efforts the Flood Estimation Handbook in Great Britain [162], the *VAPI* Project in Italy [72], the German *KOSTRA* project (e.g.,[182]) and the Australian Guide to Rainfall and Runoff [154] can be mentioned. A more complete overview on the different approaches adopted from several countries around the globe can be found, e.g, in [185, 186, 114].

Due to the significant developments of the theory of extreme value in the last two decades [55, 180, 163] the methodologies for rainfall frequency analysis are nowadays quite established and robust.

Intensity-Duration-Frequency (IDF) and Depth-Duration-Frequency (DDF) curves are commonly adopted in water resources engineering for both planning, designing and operating of water resource projects and for land and people protection purposes. They are mathematical relationships estimating respectively the average rainfall

intensity i and the average rainfall depth h over a given timescale d for a given return period T (i.e., the average recurrence period of an event over an extended period of time) [114]. These curves are usually evaluated considering the historical records for different durations and a growth curve, aimed at relating a statistical rainfall index to a design rainfall estimate (flood-index method) [59]. This involves the fitting of a proper probability distribution to the series of the maxima. Two approaches are commonly adopted: (i) the "Block Maxima" method (BM), that consists in selecting the maximum rainfall occurring over a fixed period (usually 1 year), and then use the sequence of maxima for the fitting and (ii) the "Peak-Over-Threshold" (POT) method, in which all the rainfall data exceeding some pre-specified threshold are considered [55].

Further information on the distributions commonly adopted for extreme rainfall frequency analysis and on the methodologies available for the parameters estimation and the model verification can be found in the hydrological litterature (e.g., [55, 81]).

1.4 Extreme rainfalls in Italy

Despite its generally mild climate the Mediterranean region is prone to the development of significant rainfall events. It is not uncommon to observe for a certain location the rain accumulated in one hour accounting for a whole montly average and the daily rainfall accounting for a high percentage of the annual average [8].

The work of [127] reports that Italy is the second country in Europe considering the number of flood in the recent period (17% of all Europe) and the first considering the number of causalities (30%). According to the Italian Ministry of Environment [137] 6633 municipalities are classified with a "high hydrogeological risk level" (1492 for landslides, 2023 for floods and 3118 for both of them), corresponding to 81.9% of Italian municipalities. In addition, the risk of natural disasters in Italy is rising due to the increased population density, urbanization, abandonment of mountainous areas, unauthorized buildings, deforestation, and lack of maintenance of slopes and waterways [136, 94]. Moreover, from the morphological point of view, Italy includes a large number of small and steep river catchments that can turn intense precipitations into severe devastating flash floods [161].

A large amount of meteo-climatic data is potentially available for the Italian territory. However the historical evolution of the national monitoring system prevents a straightforward use of all this information, as explained in the following.

Italy can boast of a role at the highest level in the development of meteorological observations [37], with 6 meteorological stations operating since the eighteenth century (Bologna, Milano, Roma, Padova, Palermo and Torino), and 15 stations with observation starting in the first half of the nineteenth century. First attempts of performing a systematic collection of monthly rainfall data go back to as early as 1880 when the National Office for Meteorology and Climate was funded. However, in spite of the huge heritage of data, only a small fraction of the Italian rainfall data is available in a computer-readable format, mainly due to changes in the political-administrative-legislative configuration of the country in the last century that prevented a systematic management of the national database.

The Italian National Hydrographic Service ("Servizio Idrografico Nazionale", *SIN*) was established in 1917 and collected the rainfall data from the local Compartmental Offices in the Hydrological Yearbooks until 1989. In 1989 it has been merged with the Mareographic Service into the National Bureau for Hydrographic and Mareographic Service ("Servizio Idrografico e Mareografico Nazionale", *SIMN*) that inherited the measurement network, and continued the publication of the Hydrological Yearbooks, supported in some departments by local Environmental Agencies. The Hydrological Yearbooks from 1950 onwards and for most of the previous years are divided into two parts: Part I: thermometry and pluviometry; Part II: rainfall, hydrometry, hydrological capacities and budgets, groundwater levels, sediment load, surveys, hydrological studies and exceptional events, tide measurement, mareography.

The Legislative Decree 112/1998 dismantled the *SIMN*, transferring its tasks to the 19 administrative regions and the 2 autonomous provinces of Trento and Bolzano. These authorities designated local Operational Centres and Regional Environmental Agencies to deal with hydro-meteorological monitoring and civil protection issues. Some of the local agencies continued on publishing the Hydrological Yearbooks (some of them in a digitized version), others chose other strategies for the collection and dissemination of the hydro-climatic data. In 2002 the national law forced the unification of the meteorological networks owned by the *SIMN* with those of the

local agencies. In 2008 the Italian National Institute for Environmental Protection and Research (*ISPRA*) was founded to coordinate the work of the agencies.

All these steps lead to a framework of data availability very fragmented in time and space, considered that each shift of the owner of the network implied changes in the network management and configuration, not always accompanied by adequate explanatory notes. The changes, moreover, were not contemporary in the different compartments and regions. Furthermore, in the last decades the *SIMN* and the local agencies, substituted and relocated most of the obsolete manual rain gauges with automatic ones, adding another discontinuity factor in an already complex and fragmented framework.

Various attempts for the configuration of a national database of long daily rainfall series has been made [38, 36], homogenized in a robust daily database in [37], but almost nothing has been made for sub-daily durations. Most of the information collected in the Hydrological Yearbooks is still hidden in the printed publications of the *SIMN* and in the local archives of the 21 Italian local Operational Centres.

1.5 Remote sensing of precipitation

Rainfall data from remote sensing are commonly adopted to improve the spatial characterization of rainfall systems and, considering the peculiar vulnerability of Italy to very localized rainfall extremes, they could play a key role in improving the technical knowledge in the field. However, their potentialities are only partially exploited for the probabilistic assessment of the hazard related to severe rainfall, specially when the local scale is considered. Two main remote sensing instruments are considered in the present dissertation:

- Weather radar
- Meteorological satellites

The radar ability to provide estimates of areal precipitation with a high spatio-temporal resolution (i.e., about 5-15 minutes in time and down to spatial scales finer than 1 km) and over wide areas (i.e., with an effective range of about 200-400 km) [111] makes it a good candidate to supplement the point rain gauge information in the analysis of localized rainfall systems. However, the quality of the rainfall data

obtained with this indirect monitoring technique, still needs considerable improvements, specially when dealing with large rainfall amounts. Moreover, at least for the Italian case, radar maps are not systematically accessible. Despite these limits, radar has found wide application in atmospheric research, weather observation and forecasting, proving the importance of the instrument for the analysis of the spatial rainfall systems (e.g., [19, 67, 201, 33, 116, 192]).

Spaceborne estimates of precipitation provide a significant support in the analysis of complex synoptic-scale rainfall systems and in the monitoring of the development and evolution of weather-related phenomena over the 71% of the Earth's surface covered by the sea [111]. Two type of sensing by satellites have been developed: (a) passive sensing, based on measuring the radiative intensity emitted or reflected by particles in the atmosphere and (b) active sensing, conducted using radar equipment carried by satellites [111]. For more details on the principles and techniques of remote sensing, refer to [21, 69, 183, 106]. The precipitation radar operating on the *TRMM* satellite has been used to study storms in the tropics and subtropics [159]. A wide range of studies has evaluated *TRMM* performances on a global scale (e.g., [40, 143, 128, 167, 74]). Alongside numerous positive results, some drawbacks can be pointed out. For instance, some authors have identified a poor *TRMM* quantification of large precipitation amounts in and near tropical mountainous regions and, in general, in regions characterized by intense deep convection over land (e.g., [215, 208, 65, 97]). Moreover, [194] shows an increase in *TRMM* estimation uncertainty in complex terrains, coastlines and inland water-bodies, cold surfaces, high latitudes and light precipitation. Based on *TRMM* outcomes, *NASA* (USA National Aeronautics and Space Administration)and *JAXA* (Japan Aerospace Exploration Agency) deployed the *GPM* Core Observatory [91].

1.6 The “Black Swan”

In [188] the author observes that the worst disasters in history have been the unexpected ones, because of the inability to brace against them. He names such unexpected large-impact events “Black Swans”. In hydrology, unexpected large events are produced by the nonlinearities of the systems [29, 115]. Climate variability

has the potential to produce "Black Swan" events that are unexpected but have high impact from a societal point of view [111].

In Italy these events are commonly referred to as "water bombs", perhaps derived from the definition of [172], but over-simplified by the media, and nowadays over-used and abused. An area of particular interest for intense rainfall studies in Italy is the Liguria region, with the city of Genova having experienced stunning precipitations in the last years. A record-breaking event was recorded in 2011 (4 November, [73]) with 6 victims and the all-time records for 1 hour in Italy broken with a rainfall depth of 181 mm. Another rain gauge of the city holds the Italian record for both 12 and 24 hours durations from 1970 (respectively 717.8 mm and 948.4 mm [175]). Another very localized intense storm occurred on 10 October 2010 affected only the western part of the city [42] can provide a quite representing example of the problems commonly occurring when dealing with these "water bombs". The rainfall amounts recorded during the events constitute the annual maximum for 12 hours for year 2010 for all the rain gauges of the city. Figure 1.2a show the location of 3 rain gauges of the Liguria Regional Agency for Environmental Protection (*ARPAL*). Considering the annual maxima for 12 hours duration reported in panel (b), during the year 2010 "GE Pegli" rain gauge recorded more than 350 mm in 12 hours, the "GE Castellaccio" gauge about 200 mm, and "GE Centro Funzionale" 100 mm.

Table 1.1 Annual maximum rainfall depth h_{12} for the 12 hour duration for year 2010 and related growth factor K_T.

d (h)	h_{12} (mm)	K_T
GE Pegli	376.8	4.34
GE Castellaccio	203.6	2.16
GE Centro Funzionale	100.6	1.04

The aforementioned classical statistical approaches for the frequency analysis usually involve a local estimate of the expected rainfall for fixed average return period, and then the use interpolation techniques between the available stations. The growth factor of a rainfall amount K_T is evaluated as the rate between the considered amount and the mean of the related series. It can be used as a proxy for assessing the rarity of an extreme event (i.e., the larger is K_T the more the event is rare). Analyzing the growth factors of the rainfall recorded the different gauges, as reported in table 1.1, one would deduce that a extraordinary event is more probable

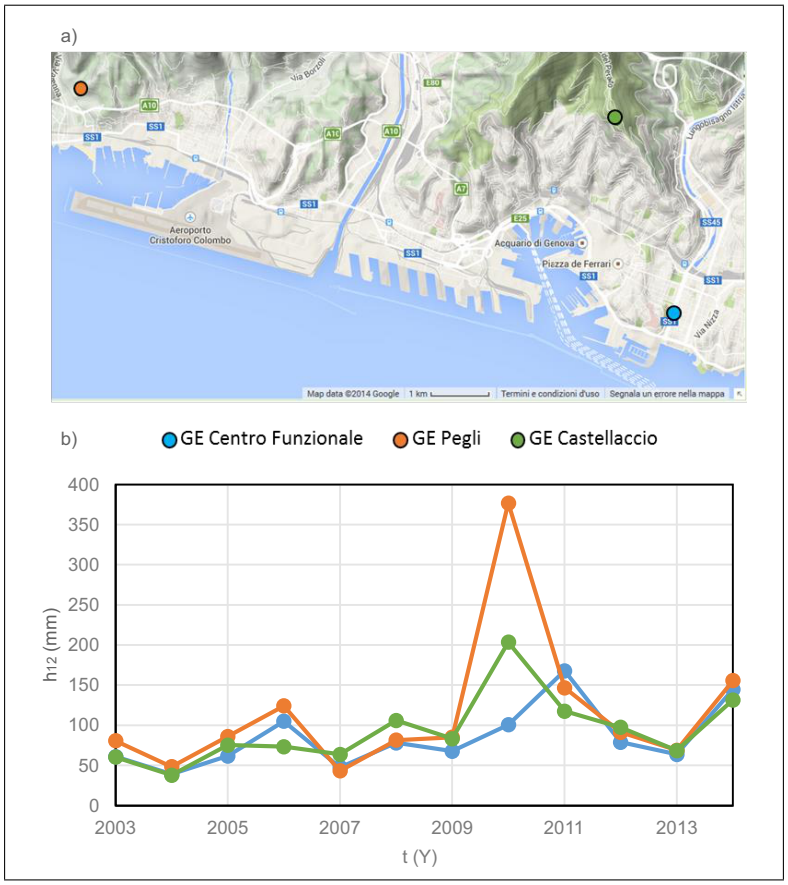


Fig. 1.2 (a) Location of the considered rain gauges and (b) annual maxima series for 12 hours duration.

in correspondence of the rain gauge "GE Pegli" than in correspondence of rain gauge "GE Castellaccio" and of the near rain gauge "GE Centro Funzionale" (which have similar characteristics to the first one in terms of elevation, distance from the sea etc.) as shown in panel (c). Is this an affordable result? Is it a reliable prediction or just a limit of the classical methodologies and of the limitedness in time of available time series? Even the regional techniques, that help overcoming the problems inherent with small data samples, show their limits because of the need to define effectively homogenous regions with sufficient amount of data: the rarity of these events leads to the definition of big regions, with the risk to loose information at the local scale.

It is therefore evident that a comprehensive assessment of the rainstorm hazard can not escape from considering the problem related to this limited population of rainfall events commonly referred as "anomalies" or "outliers" [82], showing extraordinary intensities even when compared with the population of annual maxima. Does the briefly described in section 1.3 classical probabilistic approach for extreme rainfall risk assessment allow people to prepare for unexpected large-impact events? As previously described, predictions are usually based on a large dataset of what happened in the past, and therefore they can not really predict what has not been observed [48]. Moreover, [133] consider some other examples of "hydrological monsters" for demonstrating the limits of the assumptions on which most of the commonly trusted hydrological models are based when dealing with non-stationarity and non-linearity. Different authors underline the importance of not rejecting the "bizzarre" or "monstrous" in hydrology, considering them a great opportunity for improving the predictive capacity of the hydrological models (e.g., [133, 11]).

One of the first explicit statistical treatment of this class of "extraordinary observations" has been made with the Two Component Extreme Event ($TCEV$) distribution [169]. With their work the authors draw the usefulness of gathering these observations in a separate statistical component. A mixture of just two probability distribution is the simplest type of non-stationarity hypothesis and in the context of separation might be worthy of consideration, as it could be useful in explaining the presence of very large outliers [58]. However, despite the development of the climatological and meteorological studies and of the instruments for the analysis and characterization of extreme rainfall events, the opportunity of a systematic separate statistical analysis of the "extraordinary" and "ordinary" components, is still

discussed and far from being reached [62]. Many operational problem have yet to be solved: from the definition of an objective threshold for a systematic identification of the extraordinary component to the extraction of a spatialized information from a small set of uneven distributed large values.

1.7 Aims and outline

This work aims at providing analysis, tools and methodologies for the definition of integrated approaches to the problem of the frequency analysis of extreme rainfall in wide areas, exploring the existence of an "extraordinary" component that can not be interpreted with the classic rainfall analysis methodology.

As [188] notes: "it is much easier to deal with the Black Swan problem if we focus on robustness to errors than improving predictions". The dissertation, therefore, deals with the analysis of the limits of the commonly adopted instruments and procedures for measuring, spatializing and modeling the precipitation-related risk. This is aimed at providing tools and methodologies for improving their robustness and usability in an integrated framework. The very last part of the work consists in an exploratory analysis of the "Black Swan" component, exploring the need to introduce a "super-extreme" population of events, statistically different from the rest of the dataset.

In particular, chapter 2 focuses on the treatment of fragmented and unevenly distributed rain gauge information in view of a regional frequency analysis. A complete environment for a robust frequency analysis is proposed with the adoption of the ordinary kriging equations.

Subsequently, for overcoming the limitations of the point rain gauge measurements, the potentialities of remote sensing products in analyzing the characteristics of precipitation extreme at different scales are explored. In chapter 3, on a mesoscale spatial domain, the ability of weather radar data to provide reliable estimates of rainfall depths during extreme events ia assessed. A quasi-real-time calibration procedure is then proposed. It allows to adapt the $Z-R$ relations to the well-known spatio-temporal variability of the reflectivity-rainfall relation, making the technique suitable for systematic operational use, regardless of local conditions. Radar data

could be highly valuable to patch local spots in the variability of the extremes resulting from the spatial statistical analysis.

In chapter 4 the ability of *TRMM* and *GPM* products to estimate the timing of severe storms on very large spatial domains is analyzed. These tools, due to their spatio-temporal resolution, and to their quasi-global nature, can provide useful information on the spatial structure of rainstorms and the timing of the extremes.

The last part of the work focuses on the investigation of empirical evidences aimed at preliminary assessing the need to identify a "super-extreme" population of events with a different statistical behavior, compared to the rest of the annual maxima database. In chapter 5 the *RED* database, developed in the framework of this study, is introduced. It represents the reference knowledge for extremes from 1 to 24 hours duration in Italy. Its massive consistency allows one to explore the right tail of the distribution of extreme rainfall at the country scale, analyzing that the occurrence of the largest events is too low to be statistically significant at the local scale, and requires a large database to gain the necessary insights. After assessing and correcting the effects of the non homogeneity and fragmentation of the rainfall database, some exploratory analysis aimed at identifying the presence of a threshold separating the "ordinary" and "extraordinary" component of the extreme population are carried out. The spatial distribution of the over-threshold events is also explored in order to identify the presence of areas of the Italian peninsula prone to the development of extraordinary rainfall amounts.

All the branches of the analysis are kept in the different pathways to be useful for the more general objective of understanding the spatio-temporal features of the rainfall extremes in a framework complex from both the topographic and the climatic perspective. On one hand, this study provides a set of operational tools and methodologies useful for better exploring the spatio-temporal characteristics of rainfall extremes at different spatial scales and considering different types of information, that could be potentially merged together in an integrated framework, required to overcome the limits of each instrument. On the other, the developed rainfall database and the exploratory analysis carried out help to assess the importance of further analysis on that family of extraordinary events apparently not significant when considering the single gauging site, but potentially revealing higher frequencies when a wide-area is considered. The need of further research efforts clearly emerges

from the analysis of the spatial distribution of the extraordinary events at the Italian national scale, showing how the statistical characterization of this kind of events is undoubtedly a significant knot to dissolve for avoiding dangerous underestimations of the hydrological risks.

Chapter 2

Spatial rainstorm hazard assessment: dealing with uneven and fragmented records

Rain gauge is the oldest and most accurate instrument for rainfall measurement [111], able to provide long series of reliable data. However, rain gauge records are often plagued by gaps, spatio-temporal discontinuities and inhomogeneities that could affect their suitability for a statistical assessment of the characteristics of rainfall. The non-uniform density and fragmentation of the annual maxima series can prevent a spatially-distributed frequency analysis and lead to an underestimation of the risk connected to severe storms. Furthermore, the need to discard the shorter series for obtaining robust estimates leads to ignore a relevant amount of information which can be essential, especially when large return periods are explored. Dealing with such localized and rare events, any kind of information can be useful.

As an example, figure 2.1a shows, the available series for the 24 hours duration for the "Caselle" rain gauge (45.19°N, 7.65°E, WGS84) in the Piemonte Region. In 2008 a severe localized thunderstorm occurred in the area, with rain rate approaching 300 mm in 24 hours. Only the "Caselle" rain gauge recorded a large rainfall amount, while most of the surrounding stations recorded lower rain rates, as shown in figure 2.1b. All the information related to the severe rainfall event occurred in the area is contained in a 7-year long time series, useless from a statistical point of view, and

ignored from most of the traditional frequency analysis techniques. In this chapter an alternative methodology for a robust rainfall frequency analysis in the presence of short and intermittent records is proposed. The technique allows to preserve this kind of information preserving a robust statistical environment.

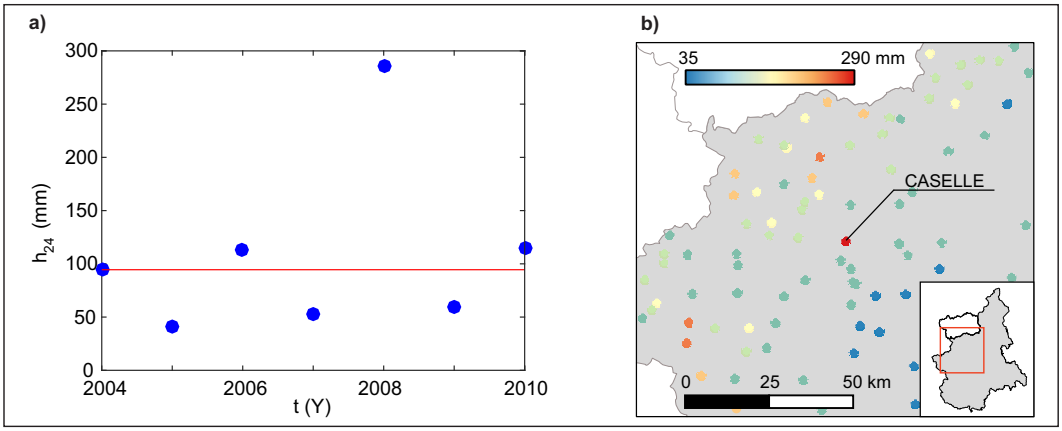


Fig. 2.1 (a) Annual maxima for the 24 hours duration series of the "Caselle" rain gauge (45.19°N, 7.65°E, WGS84). The red line shows the median of the series. (b) Annual maxima for the 24 hours duration recorded during year 2008 at a sub-sample of the database.

2.1 Dealing with spatio-temporal intermittency of rain gauge records

Rain gauge time series are often plagued with a percentage of missing values creating sporadic and/or continuous gaps, but many practical application and statistical methodologies have no tolerance to missing values [148]. The intermittent behavior traces back to the activation and dismissal of rain gauges, attributable to station relocation, service interruptions, replacement/renewal of the sensor, changes in the ownership of the station, etc. The characteristics of the stations (location and elevation, type of sensor, etc.) may also change before and after the interruptions, with consequent problems in attributing the data to a unique sample. The treatment of the historical consistence and migration of precipitation measuring points, especially when dealing with large databases, requires the set-up of specifically-conceived methods to bypass possible inconsistencies [1].

This often leads to the availability of records with too few elements for a robust statistical inference, while long series (i.e. longer than 100 years) are required to overcome the uncertainty related to the estimation of the distribution parameters [112]. When time series are short two approaches can be adopted: (i) a precautionary approach, that consist in excluding all the series shorter than a reference length value and (ii) a preservative approach, focused on the identification of methodologies aimed at extracting all the available information even from the shorter records. While on one hand the approach (i) can lead to ignore important information hidden in the shorter records, negatively affecting the results of the frequency analysis, especially in scarcely-gauged areas, the approach (ii) turns out to be more complex, computationally demanding, and can lead to severe errors when based on not robust assumptions.

Furthermore, understanding complex hydro-meteorological processes requires not only long, but also serially complete and reliable observations [148] from a dense monitoring network. The combined effect of a not-dense or a not-continuous dataset can prevent a thorough application of the aforementioned methodologies and lead to inconsistencies. It is therefore evident that, despite the existence of established rainfall frequency analysis techniques, operational and methodological problems concerning their applications still arise.

A number of procedures for recovery of information from short records can be found in the literature. Different authors propose the adoption of temporal interpolation techniques, to estimate the missing data of environmental series (e.g., linear or logistic regression, polynomial or spline interpolation, inverse distance weighting, ordinary kriging, etc.). More details can be found in the hydrologic literature, e.g. [111]. Recently, additional statistical techniques have been developed, including artificial neural networks and nearest neighbor techniques [70, 71], approaches based on Kalman filters [4] and nonlinear mathematical programming [190]. In [148] it is argued that the complexity and the computational requirements of those techniques often make them unsuitable for an application in the daily routine, that usually relies on the adoption of more traditional patching tools (e.g., filling the gaps with fixed values, often corresponding to the sample mean of the series). A viable alternative, proposed in [148], is based on the analysis of the autocorrelation structure of the series, amenable for a quick filling of sporadic gaps. However, the technique is

useful if the percentage of missing values in the time series is limited. When the gaps are wide and systematic (e.g., in developing countries [51]) and when data show low autocorrelation in time, this approach is not useful.

Even when long rainfall records are available, an *IDF* relation is basically valid only at the point where it is estimated. Gauged station are generally not even distributed in space, and they allow only for a discontinued estimation of the parameters of the rainfall distribution. To face this issue, rainfall data are commonly spatialized either by interpolating the distribution parameters estimated at the station points (e.g., [166, 140]), or by estimating the *IDF*s after pooling data into homogeneous areas, defined by geographical boundaries, or centered around a location of interest (see, e.g., [90]). Regional procedures for rainfall frequency analysis minimize the sampling error due to larger sampling size resulting from pooling. Nonetheless, regionalization is suitable only when the local spatial dependence is weak enough to enable transferring information to a site of interest from surrounding observing sites [39]. Spatial dependence may be important, though, at a larger scale, comparable with the domain size; so, in the presence of high discontinuity in rainfall distribution, due to different climatic or orographic conditions, different approaches should be preferred. In this framework, [200] propose a pure-statistical approach that involves the adoption of a bootstrap algorithm. This kind of approach allows one to overcome the problem of the data filling, but it can deviate significantly from the spatial distribution of the sample values, even when long and reliable time series are available.

In this chapter a simple approach able to provide a dense set of complete series of rainfall data in each location of the domain under analysis is proposed. The methodology, summarized in figure 2.2 is described in section 2.3.1. It is based on the sequential application of the ordinary kriging equation to the values recorded annually in the region of interest. This "patched kriging" procedure preserves the spatio-temporal information of the values recorded by the monitoring network, allowing for a more robust data-based approach in the extreme rainfall prediction.

The methodology is particularly interesting from an operational point of view. It has a low computational cost and can be potentially adopted for filling intermittent time series from different kind of monitoring networks, as it does not require the data to be significantly stationary or auto-correlated in time.

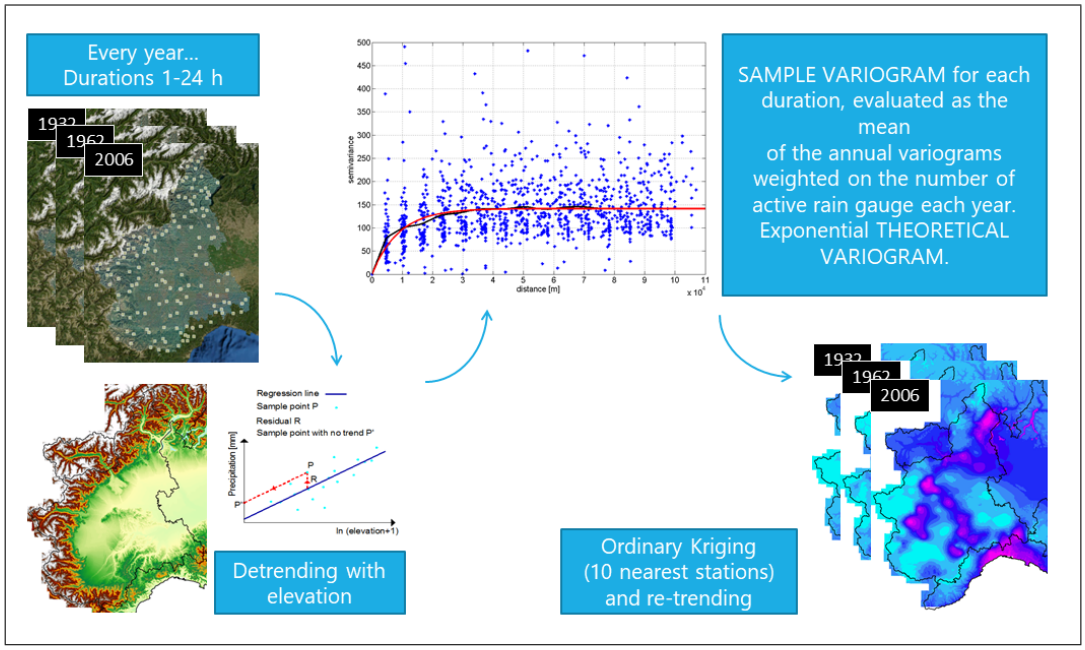


Fig. 2.2 Flow chart of the proposed "patched kriging" methodology.

2.2 The Piemonte region case study

In order to assess the validity of the methodology, a significant case study is considered: the Piemonte region, an area of about 30000 km^2 in the North-Western part of Italy. Piemonte region is considered a significant case study, due to the wide spatial extension of the regional territory, the complex orography and the abundance of rain gauge data (in Italy, as mentioned in the previous chapter, the majority of the rainfall dataset are on a regional basis). The area is characterized by a very heterogeneous orography, flat and hilly in the center, surrounded by the Alps in the North-West and by the Ligurian Apennines in the South, with minimum elevations of the order of a few meters *asl* and maximum over 4000 m *asl* (see figure 2.3a).

A dataset of annual maximum depths of cumulative rainfall over duration intervals of 1, 3, 6, 12 and 24 hours from 1928 to 2010 is considered. The data were collected from the National Bureau for Hydro-Meteorological Monitoring (*SIMN*) until the beginning of the '90s; then the network was taken over by the Regional Environmental Agency (*ARPA* Piemonte) that had removed, substituted or relocated some of the stations. Nearly 500 gauging stations have worked for at least one year

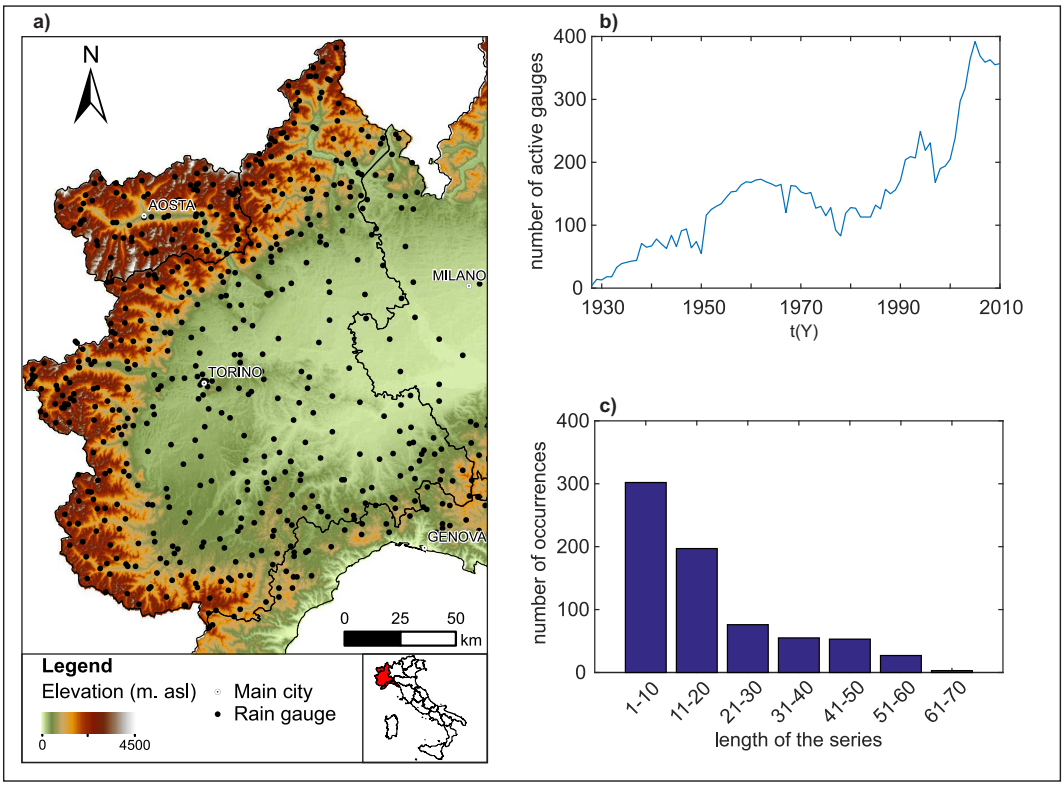


Fig. 2.3 (a) The study area and the location of the available stations. (b) Number of active station per year and (c) number of available series per length class in the study area.

in the considered period. Gauge data from neighboring regions are considered to avoid edge effect.

Figure 2.3b represents the data availability per year. It clearly shows the intermittent nature of the data. This represents a rather typical situation in Italy: only very few of the stations have a complete uninterrupted record, while the large majority has experienced disposal, relocation, interruptions and replacement/renewal of the instrument (e.g., lack of measurements during the period of the disposal of the National Bureau can clearly be distinguished). This explains why more than 50% of the considered rain gauges have series shorter than 20 years, as shown in figure 2.3c, where the stations are classified according to the length of the available series.

In this context it is clear that, despite the large and dense available rain gauge network a regional frequency analysis in the study area would require a preliminary work aimed at tracking the modification to the network, and to re-organize the whole database. This operation, if feasible, would be laborious and time-consuming, and would lead to exclude from the dataset a large amount of information, negatively affecting the robustness of the results. The large amount of information hidden in the shorter series turns out to be significant, especially dealing with intense and localized rainfall events.

2.3 Methods

This section starts with the description of the core of the "patched kriging" technique. The following sub-sections aim at providing instruments for the analysis and the correction of the results in a comprehensive robust statistical framework.

2.3.1 The "patched kriging" technique

The proposed approach, called "patched kriging", allows one to fill-in large spatial datasets by analyzing the available rainfall data year-by-year, assuming that spatial gradients can act as a proxy for temporal gradient [150, 176]. Each measurement is considered a point in the three dimensional $x-y-t$ space, where x and y are planimetric coordinates and t is time. The "patched" procedure is amenable for

application with any spatial interpolation method (e.g., Inverse Distance Weight *IDW*, etc.). In this work, the ordinary kriging is adopted (e.g., [103, 171]). Ordinary kriging is a well known stochastic interpolation method for surface interpolation, based on scalar measurements at different locations and is widely adopted in the spatialization of rainfall data [189]. Detailed description of the kriging algorithm is available in the geostatistical literature (e.g., [98]).

Kriging is based on the regionalized variable theory, which assumes that the spatial variation of the data being modeled is stationary across the surface [145]. Referring to the rainfall field, the assumption of stationarity of data leads to the need to analyze and remove the link between rainfall and elevation, especially in areas characterized by a complex orography [153, 7]. Moreover, the use of topographic information allows one to compensate the lack of information at the small scale, improving the global performance of the method [157]. In this study, the relation between h_d (mm), i.e. the annual maximum precipitation of duration d (h), and elevation z (m) is assumed to linear follow the equation:

$$h_d = m \cdot log(z+1) + m_0 + \varepsilon_d \tag{2.1}$$

where m is the slope of the regression line, m_0 (mm) the intercept and ε_d (mm) the regression residual. The logarithm of elevation is adopted as an independent variable, in order to limit the weight that linear interpolation would attribute to the stations placed at a lower altitude. The regression procedure takes into account the values recorded at all the stations in all the years simultaneously. This stems from the assumption that the relationships between precipitation and altitude is invariant over time. The de-trended at-station precipitation ($h_{d,0}$ (mm)) is obtained for various durations by subtracting the elevation effect from the observed value .

The degree of spatial dependence in the kriging approach is expressed using a sample variogram given by:

$$V(L) = \frac{1}{n(L)} \sum_{L_{ij}} (\alpha_i - \alpha_j)^2 \tag{2.2}$$

where $V(L)$ is the variance, which is defined over observations α_i and α_j lagged successively by lag-distance L and $n(L)$ representing the number of pairs of the sample separated by lag L [189].

De-trended rates are therefore used to define the annual sample variograms. For each year Y in the observation interval, the annual sample variogram $V_Y(L)$ is evaluated according to equation 2.2. A global sample variogram is obtained averaging the annual sample variograms, each weighted by the number of active stations in the considered year. The sample variogram is then converted to an analytical function, the theoretical variogram, $\gamma(L)$. Generally, several variogram models are tested before selecting a particular one. In this study, after some preliminary analysis, the exponential form is adopted:

$$\gamma(L) = c_3 + c_1(1 - e^{-L/c_2}) \tag{2.3}$$

with L (m) lag-distance, c_1 (mm^2), c_2 (m), c_3 (m) the sill, the range parameters and the nugget of the variogram [78]. The nugget effect is neglected by setting c_3=0, considering the rain-gauge records not affected from measurement errors. This is a strong assumption, but as the work deals with annual maximum values the impact of the instrumental error can be considered not significant for the aim of the analysis.

The work is carried out on a gridded 250 m x 250 m domain. The mesh size is set equal to the resolution of the adopted Digital Terrain Model (DTM). If more gauge measurements fall in the same cell, the larger one is considered.

Ordinary kriging equations are applied independently in each year. For each location, the values recorded at the nearest gauged cells are weighted according to the variogram and used to estimate the local value. Due to the neglect of the nugget effect, measurements in gauged cells are automatically preserved.

According to literature, the number of nearest gauged cells to be considered is arbitrary and it depends on the sampling pattern and the covariance matrix structure [144]. While on one hand, using the whole sample for applying the kriging equations could grant shorter computational cost, as the estimation domain is the same for all the cells of the grid, on the other hand, smaller neighborhood are preferred when there is the need to represent small-scale variability. Moreover, some authors [85] underline that the use of large neighborhoods do not lead to a significant

increase in the robustness of the estimation, as the weight associated to a distant observation quickly tends to zero because of the combined influence of the screen effect and the small to zero correlation [144]. Therefore, usually, only the stations in a neighborhood of the estimation point are considered. Some authors suggest to consider a number of sample around 10-20 [109]. Concretely, the dimension of the neighborhood should be selected according to the c_2 parameters of the variogram. In this work the significant variations of both the number and the spatial distribution of the stations along the time axis leads to the need of summarize the spatial information in a weighted mean variogram. Considering the range of this variogram for assessing the dimension of the estimation domain could affect the results, specially in years and in areas with a low density of information, leading to consider an insufficient number of rain gauges. The estimation domain is therefore limited, after a preliminary sensitivity analysis aimed at preventing the flattening of the estimated values on a global regional mean, to the nearest 10 rain-gauges, for all the cells, for all the years.

The procedure leads to the development of a set of grids (as many as the considered years), containing the estimated values of precipitation maxima for each location of the study area, configuring a "cube" of rainfall data in the $x-y-t$ space (figure 2.4a), which will be referred to as the "rainfall cube". The ordinary kriging equation provides also a "variance cube", containing the kriging variance for each cell in each year. The kriging variance is a measure of the uncertainty of the estimation for the values predicted by kriging and, due to its characteristic, will be carefully analysed in section 2.4.2.

Coring the "cubes" along the t-axis one can obtain complete "cored series" for each $x-y$ pair (figure 2.4b). This allows to have a number of uninterrupted annual maxima series equal to the number of cells in the considered domain, each related to a series of kriging variances, informing about the uncertainty of each data. All the series has the same length, equal to the length of the considered time period.

2.3.2 Weighting the L-moments

In order to guarantee a robust and data-based approach, the methodology aims at preserving as much as possible the statistics of the original series in the cored ones. This operation should be treated with caution, considering the different length of the

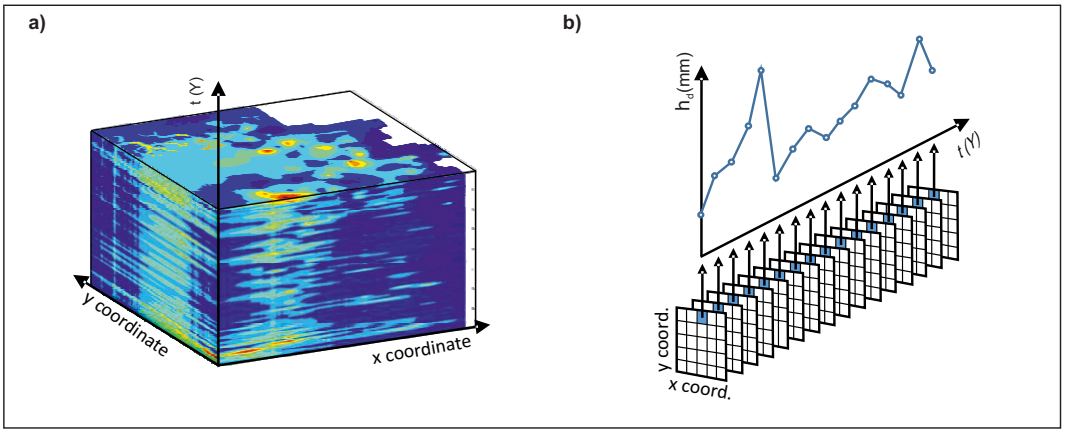


Fig. 2.4 (a) The data cube in the $x-y-t$ space. (b) Coring the data cube along the t axis a cored series for each mesh of the grid can be obtained.

original series (e.g., extracting the characteristics of a 80 years long series from a subset of 10-20 data can lead to errors, as the characteristics of the sample can be not consistent with the characteristics of the related complete series). The "patched kriging" technique helps to increase the robustness of the operation. It allows to preserve the recorded data (that is directly related with the characteristic of the original series), filling in the gaps with spatially estimated values. The estimated values are more statistically significant, the more information is available in the neighborhood of the estimation point in the considered year.

In order to take into account the different nature of the data (i.e., part of the core is measured and part is estimated by kriging) differential weight is given to each value in the evaluation of the characteristics of the cored series (i.e., more weight is given to the measured values and to the values estimated in years with more available data). The kriging variance is then considered to weight the contribution of each value to the estimation of the sample L-moments of the series. The kriging variance is a measure of the uncertainty of the estimation and is higher in cells far from gauged locations and, for a fixed cell, increases/decreases when the number of annual available station in the area decreases/increases. See, e.g., in figure 2.5a the fast increase of the kriging variance getting far from the stations, marked as red dots. In 1987 (2.5a above) it reach values up to 2500 mm^2 in the northern part of the study area, that lacks on active stations, and it shows generally lower values (up to 1700 mm^2) for year 2010, when the number of station in the study area is almost twice.

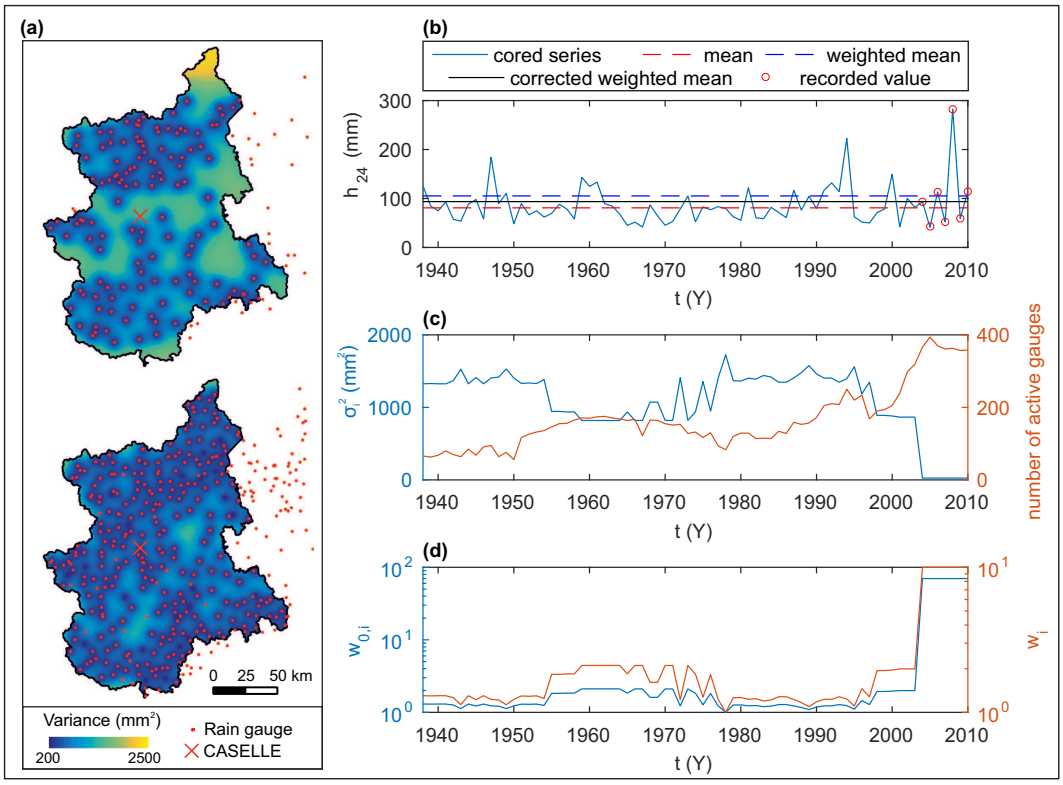


Fig. 2.5 (a) Map of the kriging variance for year 1987 (above) and year 2010 (below). The red cross shows the location of the "Caselle" rain gauge (45.19°N, 7.65°E, WGS84), installed in 2004. (b) Cored series of the "Caselle" rain gauge for 24 hours duration. The red circles mark the recorded values. All the other values are estimated with the "patched kriging" technique. The mean of the series, the weighted mean and the weighted mean with w_{max} threshold are also shown. (b) Kriging variance series for the "Caselle" location (left axis) related to the number of active gauge per year (right axis). (c) Series of the weights related to the "Caselle" series (left axis). The right axis refers to the same series, after correcting it, by setting w_{max}=10.

In the detail, for evaluating the sample L-moments of a cored series, a weight $w_i = \sigma^2_{max}/\sigma^2_i$ is assigned to the i-th value of the series, characterized by σ^2_i kriging variance (with $\sigma^2_{max} = max(\sigma^2_i)$ for the considered series).

The "Caselle" station, mentioned in section 2.2, has been installed in 2004. The cored series of the annual maxima for 24 hour duration of the cell related to its location is reported in figure 2.5b. The mean of the cored series (i.e., the red dashed line) turns out to be significantly lower than the one of the original series (i.e.,the yellow line) that is evaluated considering only 7 years of data.

Analyzing the series of the kriging variance of the "Caselle" location (figure 2.5c) one can see the sensitivity of this parameter to the number of globally available stations: as previously mentioned the kriging variance increases/decreases with the decrease/increase of the number of active gauges. It drastically decreases in the year 2004, when the station has been installed. From figure 2.5b one can also observe that the weighted mean, evaluated with the weights reported in figure 2.5d (left axis), is almost equal to the mean related to the period 2004-2010. When a station is installed in an ungauged cell, the kriging variance decrease drastically and this lead to give virtually zero weight to all the previously kriged values. Considering the lack of reliability of L-moments estimated on short series, this phenomena should be avoided, as this would undermine the benefit of the "patched kriging" methodology. A maximum threshold w_{max} is then set. For $w_i > w_{max}$, w_i=w_{max} is considered. After some sensitivity analysis, aimed at giving larger weight to the measured values without denying the contribution of the cored ones, w_{max}=10 is set. The final weights adopted for the "Caselle" cell are reported in figure 2.5d (right axis), and the resulting mean values is showed in figure 2.5a with the black line.

From a mathematical point of view, given a sample with size n sorted in ascending order $(x_{1:n} \leq x_{2:n} \leq ... \leq x_{n:n})$, and considering:

$$b_r = n^{-1} \sum_{i=r+1}^{n} \frac{(i-1)(i-2)...(i-r)}{(n-1)(n-2)...(n-r)} x_{i:n} \tag{2.4}$$

the sample L-moments can be written as:

$l_1 = b_0$

$l_2 = 2b_1 - b_0$

$l_3 = 6b_2 - 6_b1 + b_0$

and, in general:

$$l_{r+1} = \sum_{k=0}^{r} p^*_{r,k} b_k \tag{2.5}$$

with r= 0,1,...,n-1 and $p^*_{r,k} = \frac{(-1)^{r-k}(r+k)!}{(k!)^2(r-k)!}$.

As mentioned above, each value is weighted according to the estimation variance associated with it. In the detail, to the i-th value of the considered cored series,

characterized by σ_i^2 estimation variance, is assigned a weight $w_i = \sigma_{max}^2/\sigma_i^2$, with $\sigma_{max}^2 = max(\sigma_i^2)$ for the considered series. Once defined $W_i = \sum_{k=1}^{i} w_i$, each cored series (all characterized by the same length n) acquires an effective length $m = W_n$. Concretely, the weighting procedure inserts a number of virtual ties, aimed at giving more weight to some values than to others, so the effective length of a cored series equals the sum of its weights. Considering the $y_{j:m}$ elements of the series series including the virtual ties, sorted in ascending order, the equivalent of equation 2.4 for the weighted series can be written as:

$$b_r = m^{-1} \sum_{j=r+1}^{m} \frac{(j-1)(j-2)...(j-r)}{(m-1)(m-2)...(m-r)} y_{j:m} \tag{2.6}$$

with $y_{(j)}$=$x_{(i)}$ for $1 + W_{i-1} \leq j \leq W_i$.

Evaluating equation 2.6, the L-moments weighted on the estimation variances can be obtained from 2.5. For simplicity the explicit forms of 2.6 for $r = 1,2,3,4$ used in this study are reported in the following.

$$b_0 = \frac{1}{m} \sum_{i=1}^{n} w_i x_{(i)}$$

$$b_1 = \frac{1}{m(m-1)} \sum_{i=1}^{n} w_i x_{(i)} \left(W_{i-1} + \frac{1}{2}(w_i - 1) \right)$$

$$b_2 = \frac{1}{m(m-1)(m-2)} \sum_{i=1}^{n} w_i x_{(i)} \left(\frac{1}{3} w_i^2 + w_i(W_{i-1} - 1) + \frac{2}{3} - 2W_{i-1} + W_{i-1}^2 \right)$$

$$b_3 = \frac{1}{m(m-1)(m-2)(m-3)} \sum_{i=1}^{n} \frac{1}{4} w_i x_{(i)} \cdot$$

$$\cdot \left(w_i^3 + w_i^2(4W_{i-1} - 6) + w_i(6W_{i-1}^2 - 18W_{i-1} + 11) + 4W_{i-1}^3 - 18W_{i-1}^2 + 22W_{i-1} - 6) \right)$$

2.4 Application and results

2.4.1 Application of the "patched kriging" methodology

The technique is applied to the study area in the Piemonte region. Data for each duration d are considered as separate series, so as to obtain 5 different series per rain gauge.

Regression of rainfall depths with the logarithm of the elevation has been carried out considering the equation 2.1 for the 5 durations as reported in figure 2.6.

The parameters of the regression lines (figure 2.7a) are reported in table 2.1. The trends significance is evaluated with a Student's t test with a significance level α=0.05.

As shown in figure 2.7a the maximum annual precipitations of duration 1 and 3 hours show a declining trend with elevation, which loses significance for the duration 6 hours and becomes a positive trend for the durations of 12 and 24 hours. This justify the absence of the expected increasing trend of the intercept of the regression lines with the duration, and is consistent with the findings of [7] that relate the different behavior with the nature of the events typical of the different durations (mostly convective for shorter durations, stratiforms for longer ones).

Table 2.1 Parameters of the regression of precipitation versus elevation, for different durations (* indicates a significant trend at a 5% level.).

d (h)	m	m_0 (mm)
1	-3.56*	48.73
3	-2.57*	55.72
6	-0.32	54.59
12	3.52*	49.08
24	8.34*	44.54

Using the coefficients reported in table 2.1, the de-trended precipitation ($h_{d,0}$) is obtained. For d=6, $h_{6,0}$=h_6 is set, due to the lack of significance of the hypothesis $m \neq 0$.

Sample and theoretical variograms are then defined according to equation 2.3. Table 2.2 reports the coefficients of the theoretical variograms for the different

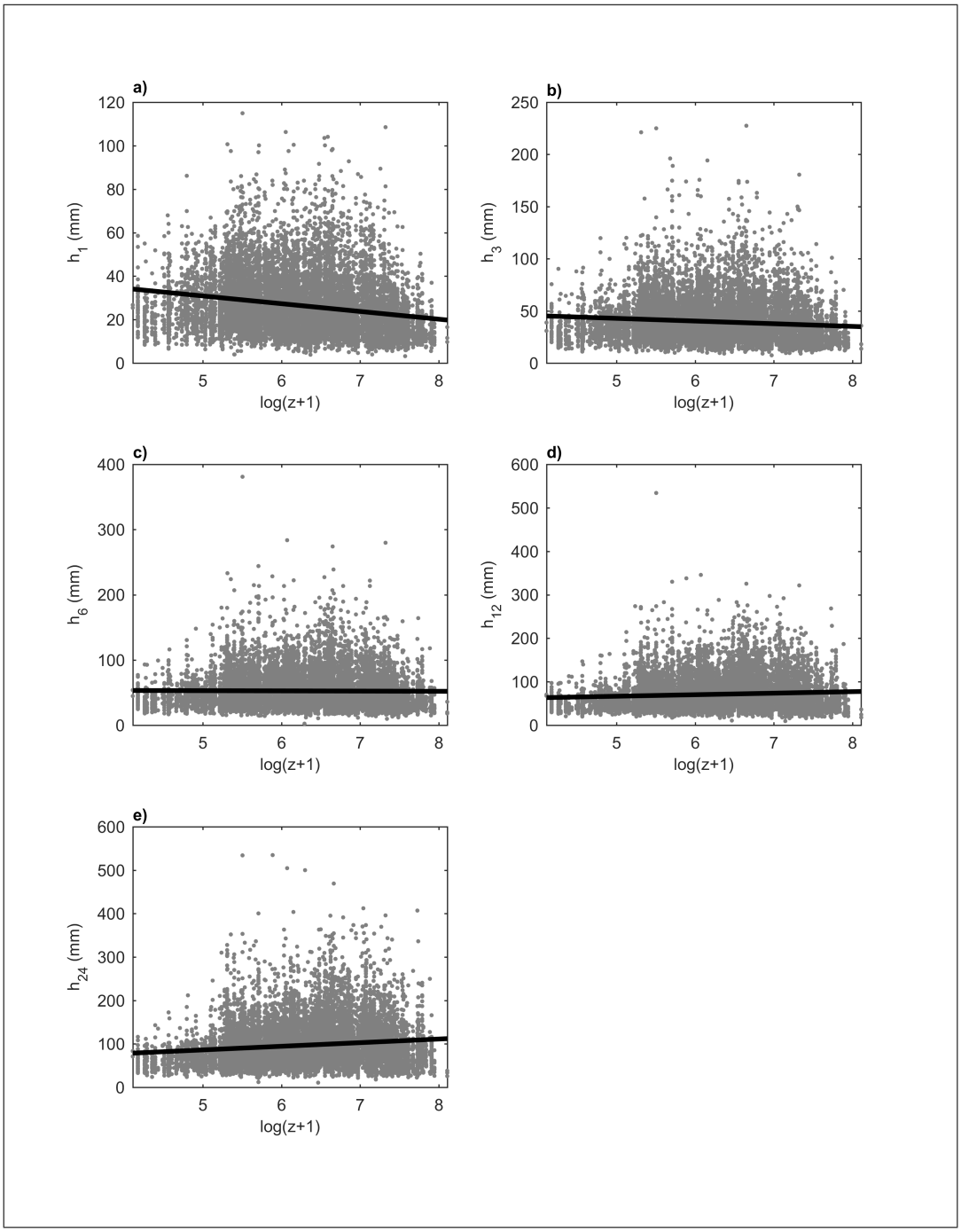


Fig. 2.6 Regression of the rainfall depths with the logarithm of the station elevations for the (a) 1 to (e) 24 hours durations. The regression line is reported in black.

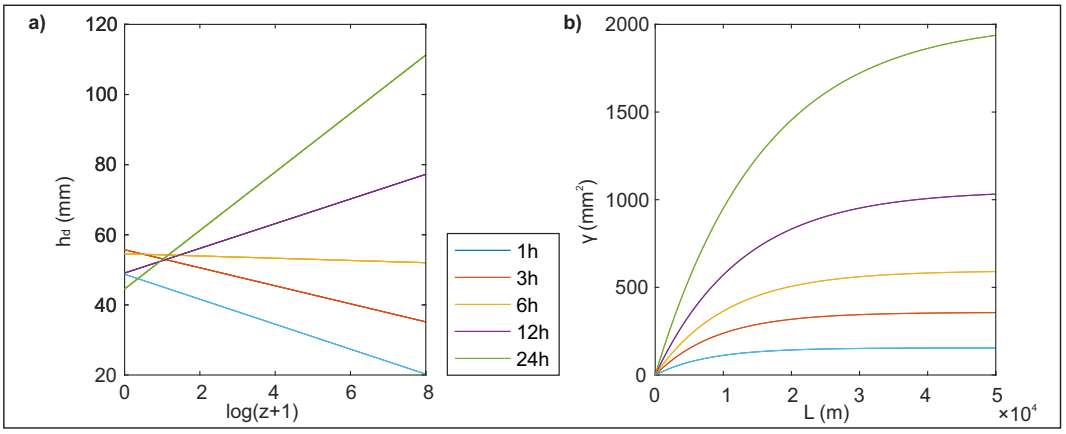


Fig. 2.7 (a) Regression lines adopted for the detrending of the data and (b) estimated theoretical variograms for the considered durations.

durations. Sample and theoretical variograms for all the durations are reported in appendix A.1. The apparently large annual variability of the sample variograms is partially ascribable to the scarce data density in the first analysed years that leads to scarcerly significant sample variograms. To avoid loss of robustness, as described in section 2.3.1, the annual variograms are weigthed according to the number of annual active stations. The obtained theoretical variograms are also reported in figure 2.7b.

Table 2.2 Estimated parameters for the theoretical exponential variograms.

d (h)	c_1 (mm^2)	c_2 (m)
1	142	6709
3	334.7	8798
6	574.2	10240
12	1051	11520
24	2028	13650

With the application of the ordinary kriging equations, as described in section 2.3.1 a set of 5 “rainfall cubes” (one per duration) with the related “variance cubes” is obtained.

2.4.2 Analysis and validation of the patched series

At first, in points where sample data are available, the cored series are validated by comparing their L-moments with the ones of the original series.

Given the lack of significance of the shorter series from a statistical point of view (i.e., as previously mentioned, the L-moments estimated from short fragmented series can be not-consistent with real characteristics of the related uninterrupted series) the analysis is restricted to the series with more than 20 years of data. Figure 2.8 reports the comparison between $L-CV$, $L-CA$ and $L-KUR$ of the measured versus the estimated series for all the durations. L-moments are measures of the location, scale and shape of the probability distributions. In the detail:

- $L-CV = l_2/l_1$, coefficient of L-variation, express the dispersion of a distribution referring to the mean
- $L-CA = l_3/l_2$, coefficient of L-skewness, give information about the symmetry of the distribution
- $L-KUR = l_4/l_2$, coefficient of L-kurtosis, give information about the "peakedness" of the distribution

with l_r, r order L-moment [87, 88].

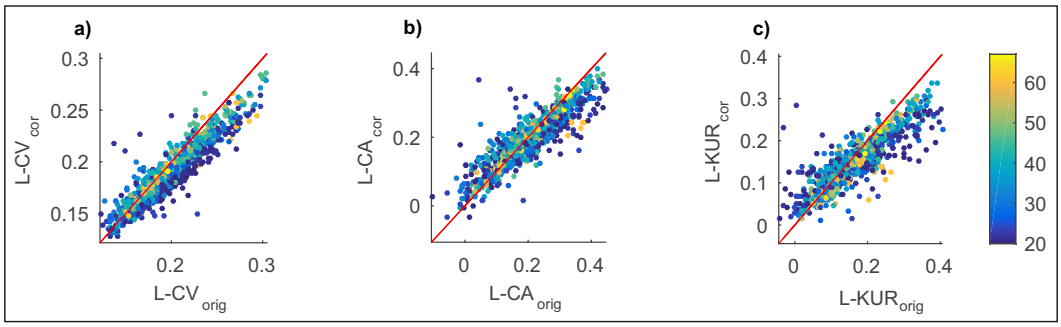


Fig. 2.8 (a) $L-CV$ (b) $L-CA$ (c) $L-KUR$ of the measured versus the cored series for all the durations. The chromatic scale refers to the length of the series.

The comparison demonstrates the ability of the methodology to preserve the L-moments reasonably well, except for a a slight decrease in the $L-CV$ of the cored series, as seen in panel (a) in figure 2.8 that compares the measured (*orig*) with the cored (*cor*) series for all the durations.

To assess the behavior of the methodology even in cell without sample data, or with a number of sample that does not allow for a robust estimation of the sample L-moments, the clouds of the sample L-moments of the cored series in the $L-CA$ versus $L-KUR$ and $L-CA$ versus $L-CV$ planes (L-moments ratio diagrams, [90])

are compared with the ones of the original series with more than 20 years of data, considering all the durations together (figure 2.9a-b).

A significant underestimation of the $L-CV$ values is evident from the analysis of panel (b), while a slight underestimation of the $L-CA$ values appears from panel (a). In the $L-CA$ versus $L-CV$ diagram, the barycenter of the cloud related to the estimated values is significantly lower than the one of the measured ones; this implies that the cored series denote smaller variability along the time axis than the original ones. This is an expected drawback for applying a spatial interpolation technique, and is consistent with what emerges from the analysis of the gauged cells in figure 2.8. As the underestimation of the $L-CV$ leads to underestimation of the design rainfall, for safety purposes, a correction factor to correct this behavior is described in the following section.

2.4.3 Variance correction

The attenuation of the variability in the cored time series can result in underestimation of the design rainfall. To correct this behavior a bias-correction factor is proposed, conceived to increase the variance of the cored series at increasing the one of the original ones.

Numerous advanced techniques for reducing the bias in estimated and interpolated time series are available, but most of them require the presence of a significant database representing the variability of the analysed variable along time (e.g., [121, 191]) or of other kind of supportive information (e.g., radar data [63]) to identify the expected variance in time for each location.

The lack of a robust database representing the variability of rainfall along the t-axis (this lack is, indeed, the main issue this chapter is trying to manage), prevents a straightforward definition of this correction factor. By using a "trading space for time approach" [150, 28] the problem is tackled by exploiting the wide spatial information to develop a correction factor for each cell and each year, based on the spatial distribution of the rain gauges, without the need to estimate the expected variability along time for each location of the domain.

One of the most influencing parameter in the spatial distribution of the bias is the distance of the considered location from the closer gauged cells. As the distribution

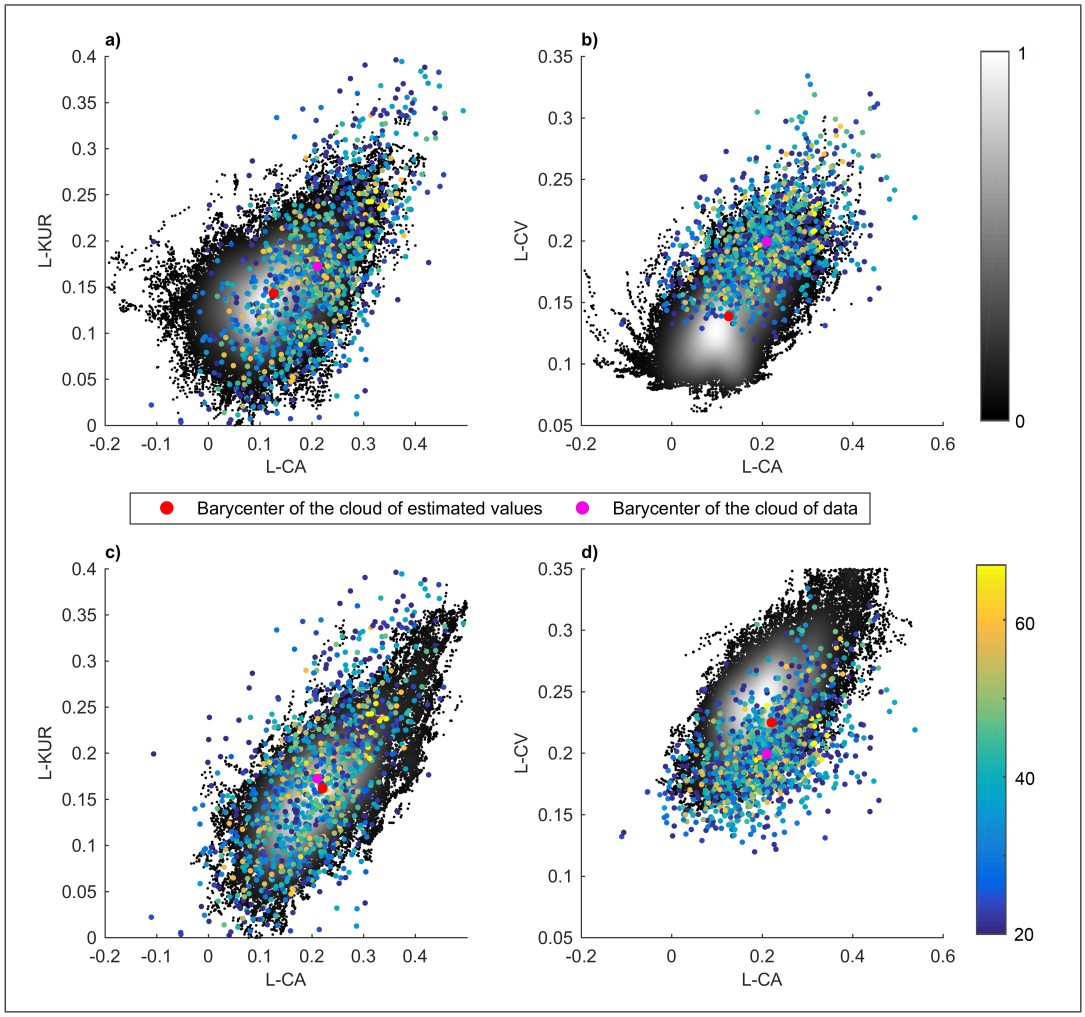


Fig. 2.9 L-moments ratio diagrams (a),(b) before and (c),(d) after the correction considering all the durations. The greyscaled cloud of points represent the cored series. The greyscale is proportional to the density of points. The colored dots represent the original series. The colorscale is related to the length of the series.

of the rain gauges changes annually, every cell is considered independently for each year. The closer the location to one or more rain gauges, the more accurate its estimation is and the less the correction that the cell requires.

In this work is supposed that all the cells with similar distance to the nearest gauged cells show similar behavior and belong to the same sub-sample. This sub-sample should have a variance consistent with the one of the sample estimated pooling together all the measured values. Why otherwise should the sample variance of the set of data estimated in cells located at a certain distance from a gauged cell differ from that of the set of the at-gauge recorded data? Apparently there are no reasons for that, excepting the existence of a bias that needs to be corrected.

The correction factor is then calibrated considering the variance of the set of the at-gauge records as the reference value. Various distance classes are considered and the variance of each class corrected to be consistent with the reference value. The spatial scale is used as a proxy for the temporal variations.

The correction procedure is applied separately for the 5 considered durations.

Concretely, for each year a D_s value is assigned to each cell, representing the inverse average distance of the cell from the nearest 10 gauged cells, evaluated as:

$$D_s = \frac{1}{\frac{1}{10}\sum_{j=1}^{10}\left(\frac{1}{dist_j}\right)} \tag{2.7}$$

with $dist_j$ number of cells separating the considered cell from the j-th closest gauged one (figure 2.10a). The inverse average distance is considered in order to assign a D_s value approaching zero to the gauged cells that does not require any kind of correction, as the recorded gauge values are to be preserved.

Equally consistent D_s classes are then considered to evaluate, for each class, the rate between Var_0 (i.e., the sample variance of the set of the at-gauge recorded data for the considered duration) and Var_{D_s} (i.e., the sample variance of the cells belonging to the D_s class - figure 2.10b). The same operation is carried out for all the years. The D_s-$\frac{Var_{D_s}}{Var_0}$ pairs belonging to all the years are pooled up together and the median value for each equally spaced D_s class is considered. They are represented as dots in figure 2.10c.

Figure 2.10c reports the results obtained for the different durations. A correction function $K(D_s,d)$ (2.8) aimed at reducing the ratio between the variance of the recorded dataset and the estimated one for each D_s class is finally found by interpolation, one for each duration. The correction factor increases with the increasing of the distance from gauged cells. For D_s >90 the distance from the nearer rain-gauges seems to loose relevance and a constant model is adopted. The obtained $K(D_s,d)$ coefficients are reported in table 2.3.

$$K(D_s,d) = \begin{cases} 1+\beta(d)\cdot D_s & D_s \leq 90 \\ 1+\beta(d)\cdot 90 & D_s > 90 \end{cases} \tag{2.8}$$

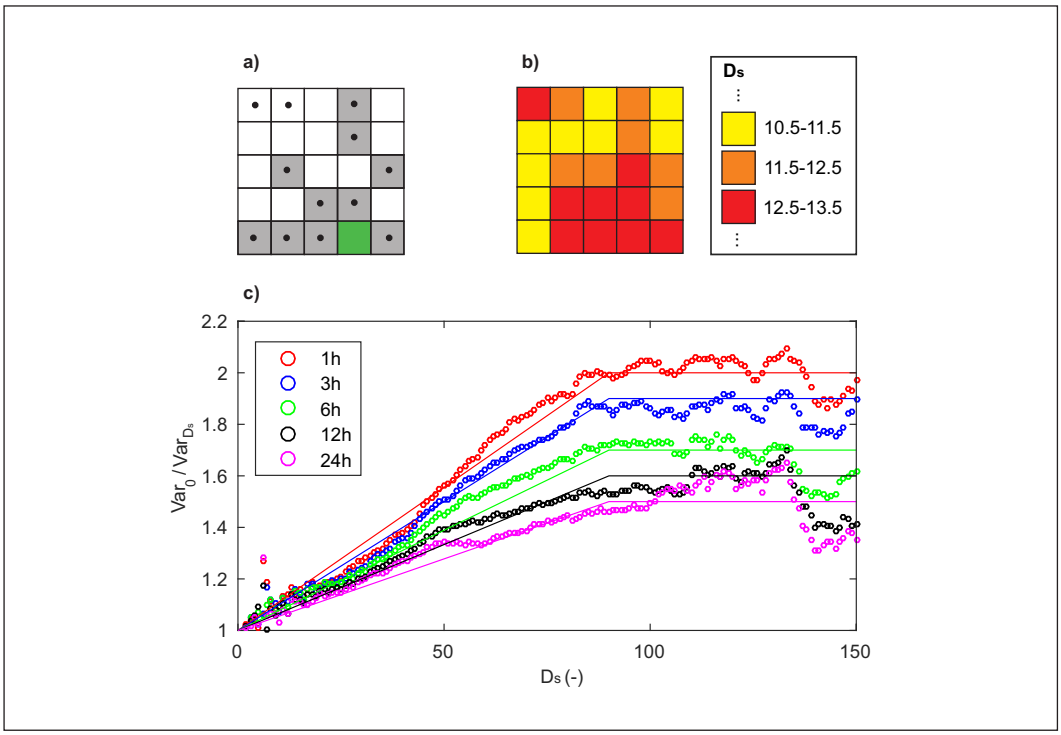


Fig. 2.10 Steps for the definition of the correction factor. (a) Considering the green cell, the distances from the 10 grey gauged ones are considered for evaluating D_s. (b) The sample variance of the set of cell in the same D_s class (e.g., same color in the figure) is considered for assessing the $\frac{Var_0}{Var_{D_s}}$ ratio of the class. The same operation is carried out for all the classes for all the years. (c) All the D_s-$\frac{Var_0}{Var_{D_s}}$ pairs are pooled up together and for each D_s class the median value is considered for defining the average correction factors for the different durations d.

Table 2.3 Coefficients β of the correction function $K(D_s, d)$ for the different durations d.

d (h)	β
1	0.0111
3	0.0100
6	0.0078
12	0.0067
24	0.0056

After the application of the variance-bias correction factor, the generic observation $\dot{x}_i(t)$ in a i cell for the generic year t is set according to the following equation:

$$\hat{x}_i(t) = \bar{x}_i \left(\frac{\dot{x}_i(t)}{\bar{x}_i} \right)^{K(D_s,d)} \tag{2.9}$$

with $\hat{x}_i(t)$ as the corrected value, $\dot{x}_i(t)$ as the original value, $\bar{x}_i$ as the mean of the cored series and $K(D_s, d)$ as the correction factor, which is a function of the distance D_s and varies for the 5 durations d. A power-law equation is adopted in order to avoid obtaining negative rainfall values.

The proposed correction helps in preserving the observed variability of the series along the t-axis, making the *L-CV* and *L-CA* values of the cored series consistent with the L-coefficients of the observed ones. Figure 2.9c-d compares the L-moments ratio diagram [90] before and after the application of the correction factor, considering all the durations together. The correction improves the performance of the technique, allowing also for more skewed cored series. Although a slight overestimation of the *L-CV* values can still be observed (more significant when considering shorter series), the main objectives of the bias correction seems acceptably reached.

2.5 Frequency Analysis

By considering the cored series, coefficients a and n of the average *IDF* are estimated for each cell in the study area. *IDF* curves are represented according to the equation:

$$\bar{h_d} = ad^n \tag{2.10}$$

with $\bar{h}_d$ average precipitation depth, for d duration and a and n parameters.

Figure 2.11 shows the parameters over the study area. To further assess the validity of the results, the relative differences between the values of the parameter evaluated with the original series and the ones estimated with the cored ones is considered. The maps of the spatial distribution of the differences (omitted) shows that no particular spatial clustering can be observed. Significant differences between the two sets of parameters are mainly related to the length of the original series. Comparing the differences with the length of the series, an increasing behavior with the shortening of the series is confirmed. This can be related to problems of significance and stability of the parameters estimated with short series. Increasing the number of rain gauges considered in the estimation phase the "patched kriging" allows for a more robust and data-based estimation procedure.

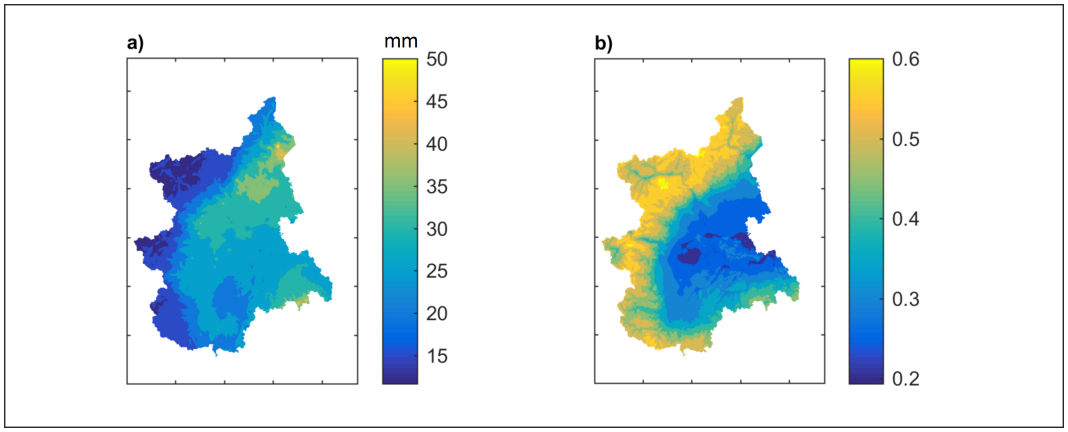


Fig. 2.11 (a) a and (b) n parameters of the mean IDF curves.

In order to define the design rainfall for a given return period for each point of the domain under analysis it is necessary to identify a probability distribution that correctly represents the annual maxima. It would be then possible to estimate the rainfall depth $\hat{h}_{d,T}$ related to a certain duration d and a given return period T using the IDF curve previously identified and the growth factor K_T:

$$\hat{h}_{d,T} = ad^n K_T \tag{2.11}$$

with a and n coefficients of the mean IDF.

Different probability distributions have been used in the literature to statistically represent rainfall extremes, with preference to the Gumbel and GEV distribution. To

assess the suitability of them to the considered dataset a preliminary analysis on the L-moments of the cored series is developed.

The procedure for the estimation of the weighted L-Moments consists in inserting a number of fake duplicate values in the sample, aimed at giving more weight to some values compared to others. This circumstance prevents the use of classic goodness of fit tests (e.g., Anderson-Darling) [64]. The goodness of fit of the distributions to the data is then discussed by analyzing results on the L-Moments ratio diagrams (figure 2.12a) [90].

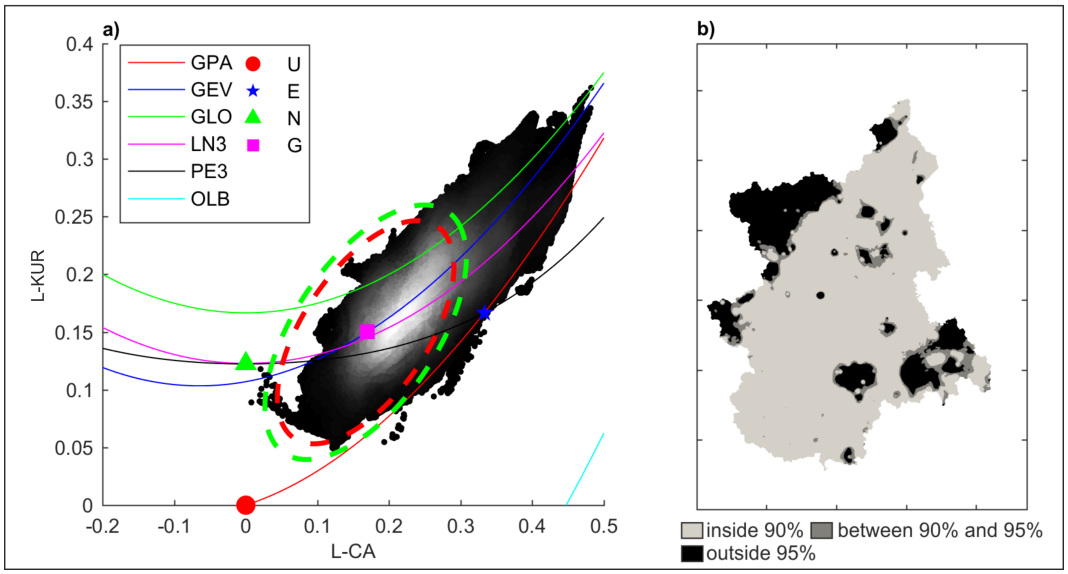


Fig. 2.12 (a) L-moment ratio diagram related to the cored series. The ellipses represent respectively the 95% (green) and the 90% (red) acceptance area, defined by bootstrapping from a Gumbel distribution. The color scale is proportional to the density of the points. Key to distributions: *E* - Exponential, *G* - Gumbel, *N* - Normal, *U* - Uniform, *GPA* - Generalized Pareto, *GEV* - Generalized Extreme Value, *GLO* - Generalized Logistic, *LN*3 - Lognormal, *PE*3 - Pearson type III. *OLB* is the overall lower bound of $L-kurtosis$ as function of $L-CA$. (b) Spatial distribution of the cells falling inside the 90% and 95% acceptance area of the Gumbel distribution.

Figure 2.12a shows the points and curves representing in the $L-CA$ vs $L-KUR$ plane different distributions commonly used in the analysis of extreme values. Plotting the L-moments of the cored series allows one to visually evaluate the global behavior of the samples.

The diagram seems to confirm that the Gumbel distribution is a good candidate to represent extreme precipitations at the regional scale, although the barycenter of

the cloud of points is slightly shifted towards higher $L-CA$ values. To identify the amount of variability due to the sample size, 25000 series with a length of 72 year has been randomly extracted from a Gumbel distribution with shape and position parameter set to 1, with a Monte Carlo procedure. This leads to the identification of the 90% and 95% acceptance areas, overlapped to the estimated points in figure 2.12a. The greater amount of points fall into the possible domain of a Gumbel distribution, although there are some series characterized by a larger kurtosis value, that fall outside the domain of the Gumbel distribution. In these cases, a *GEV* distribution can be a viable alternative, although the use of distributions with three parameters increases the uncertainty associated to the estimates. This uncertainty depends on the inherent difficulty in estimating the shape parameter of the distribution, especially in the presence of short and not evenly distributed recorded series.

Figure 2.12b shows the spatial distribution of the cells falling inside the theoretical acceptance area of the Gumbel distribution. As expected a regular pattern in the acceptance areas is hardly defined, due to the complex topography and to the different characteristics of the events generating annual maxima for different durations at the regional scale [186]. The mixture of scales involved in data generation (local and synoptic scale events) and the effect of orography on storm generation (particularly significant in the northern-west Alpine and south-eastern Appenninic areas) lead to a spatial pattern which can hardly lead to the identification of a unique regional distribution. In addition, boundaries effects may occur at the edge of the analysed domain.

For completeness, in the following, both the Gumbel and the *GEV* distribution are considered.

The *GEV* distribution can be expressed for a certain T return period by the equation [101]:

$$K_T = \frac{\theta_1}{\mu} + \frac{\theta_2}{\mu\theta_3}\left[1 - \left(-ln\left(\frac{T-1}{T}\right)\right)^{\theta_3}\right] \tag{2.12}$$

where μ is the mean, θ_1, θ_2 >0 and θ_3 are location, scale and shape parameters, respectively. When θ_3=0, the *GEV* reduces to the Gumbel distribution [113]:

$$K_T = \frac{\theta_1}{\mu} - \frac{\theta_2}{\mu} ln\left(-ln\left(\frac{T-1}{T}\right)\right) \tag{2.13}$$

For each cell of the grid the parameters of the distributions are estimated, both forcing the hypothesis θ_3=0 (Gumbel distribution) and letting the shape parameter vary.

For the estimation of the parameters of both the distributions the L-moments methodology is adopted. In the detail, at first the sample weighted L-moments are estimated for each duration and then the average L-moments among the different durations are considered for estimating the dimensionless parameters of the Gumbel and *GEV* distribution reported in figure 2.13.

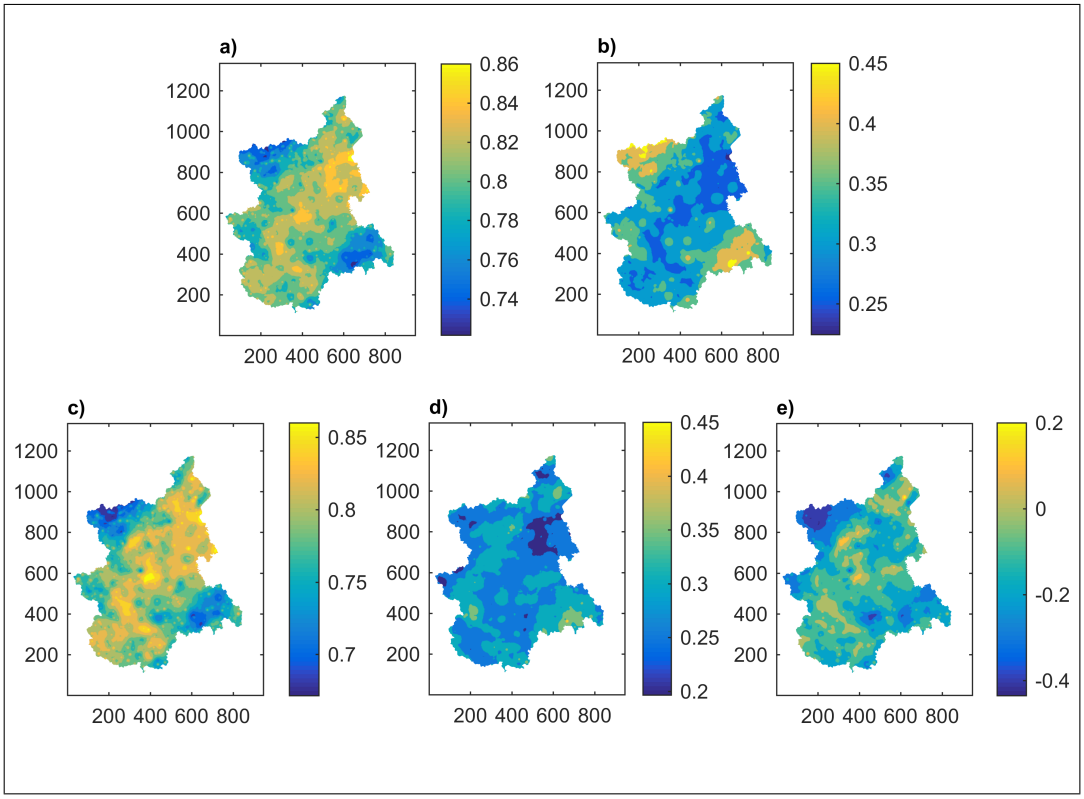


Fig. 2.13 Dimensionless position (θ_1), and scale (θ_2) parameter of the (a),(b) Gumbel and (c),(d) *GEV* distribution, normalized on the mean rainfall depth. (e) Shape parameter (θ_3) of the *GEV* distribution.

At a first sight, the use of cored series allows one to limit the noise introduced in the estimation procedure by the shorter series. This allows one to estimate more reliably the distribution parameters. For example the shape coefficient (θ_3) of the *GEV* distribution, estimated with the original series, takes on values between -1 and 3. However, according to any literature the shape parameter usually assumes

values between -0.59 and 0.76 [147], smaller or larger values being ascribable to an excessive sampling variability in small samples (see figure 2.14). In this study values between -0.4 and 0.2 are obtained.

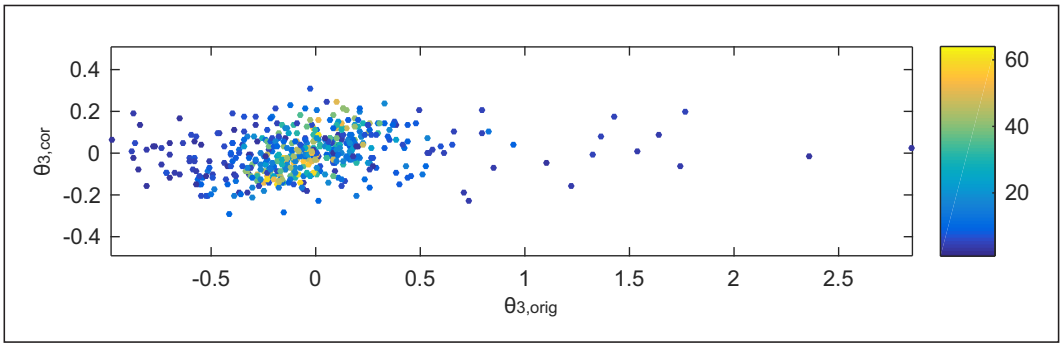


Fig. 2.14 Shape parameter from the cored time series versus θ_3 of the original samples. The color scale refers to the length of the series.

2.6 Concluding remarks

Estimation of design rainfall extremes in ungauged sites benefits of interpolation techniques usually based on datasets with good record length. Measurements networks under development, however, can usually provide only shorter series, sometimes containing important convective events, difficult to integrate with the older ones.

Treating each measurements like a point in the $x-y-t$ space, the "patched kriging" technique allows one to exploit all the available information in space and time with a non-conventional gap-filling procedure. The cored series obtained with the "patched kriging" technique provide greater robustness and less sensitivity to local artifacts during the distribution estimation phase, producing series with homogeneous lengths. The information concerning the estimation uncertainty is carried out thanks to a "variance cube" containing the kriging variance per location per year. The "best" probability distribution can be therefore defined separately at each location of the gridded domain.

The North-West of Italy has been considered as a case study given the presence of a wide rain gauge network with very fragmented records and of a complex orography. A dataset of annual maximum depths of cumulative rainfall over duration intervals of 1, 3, 6, 12 and 24 hours from 1928 to 2006 is analysed.

After the "patched kriging" application, a methodology aimed at defining the global behavior of the different distributions at the regional scale has been undertaken on the patched dataset. The methodology leds to identify the predominance of the Gumbel distribution at a regional scale, which confirms previously at-gauge studies on the available long series [13]. However, it is not possible to identify motivations behind the specific patterns of the areas of non-acceptability of the $EV1$ distribution. Based on the selection of at-site and data-driven best distributions at the local scale, the results of this study discourage the adoption of an unique distribution at the regional scale. One can argue that a GEV distribution with θ_3=0 collapses on a Gumbel distribution, so it could be the best candidate as a general-purposes model. However, a clear distinction of the cases where θ_3=0 needs validation and is very useful for the end-users. From this perspective the proposed methodology offers a powerful and expeditious procedure, suitable to grant an at-site evaluation of the best distribution and of the related quantile, aimed at carrying on a regional frequency analysis always consistent with the available data without the need to discard the information hidden in the shorter series.

An example of the potentiality of the "patched kriging" technique can be seen in figure 2.15b, which shows how the methodology allows to take into account also the information related to the "Caselle" storm, mentioned in the introduction, that would be lost with the traditional techniques. Despite the evident improvement, the spatialization of the rainfall information is probably not entirely consistent with the real behavior of the rainfall fields, as the real spatial distribution of the event is not known and the only information is the at-gauge rainfall depth. This limitation is inherent in the nature of rain gauge data and can not be tackled without inserting other source of information in the system, as, for instance, rainfall fields from weather radar.

In the following chapters, the usability of remotely sensed information to integrate rain gauge data in the spatial characterization of rainfall fields is evaluated. As remote sensing is a relatively new technology and long series of data are not yet available, the work is focused on the analysis of the potentiality of these tools in the characterization of intense precipitation events, focusing in identifying suggestions and methods to improve their applicability in this field.

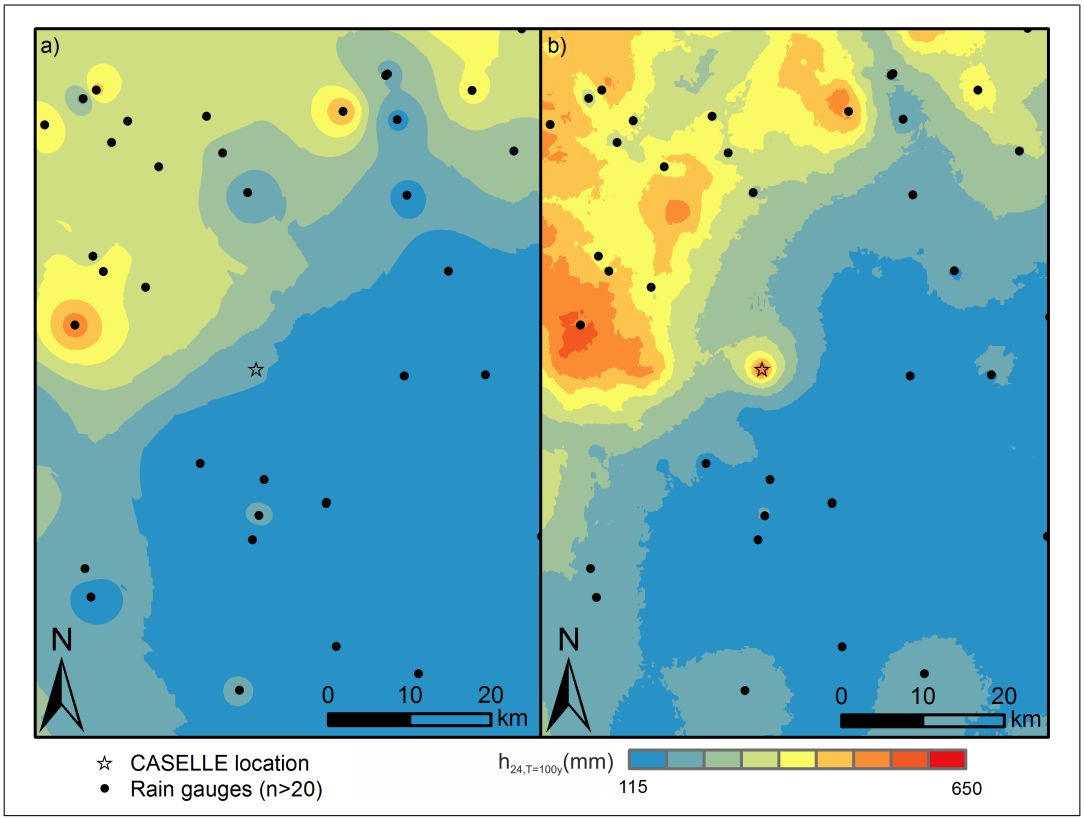


Fig. 2.15 Design rainfall quantiles predicted for the "Caselle" area (see figure 3.6) for 24 hours duration and return period T=100 Y, using a *GEV* distribution evaluated (a) with local estimation of the parameters considering only the rain gauges with record length >20 and spatial smoothing the results with the *IDW* interpolation and (b) with the "patched kriging" technique.

Chapter 3

Remote-sensing of rainfall extremes: mesoscale and weather radar

Part of the work described in this chapter has been previously published in an article in a peer-reviewed international journal [123].

As seen in the previous chapter, when focusing on local, small scale rainstorms, even good-quality high-density rain gauge networks present undoubted under-sampling of rainfall events. This means that only occasionally point measurements allow to detect the highest values in a rain burst.

In this regard, to gain a complete picture of the evolution of the rainfall field over time, the use of weather radar is a well-established practice. Weather radar is generally considered more accurate than satellite in rainfall estimation [181] and presents lower maintenance costs than traditional rain gauge networks. However, it can not offer the same spatial range of satellite data, and, as previously mentioned, at least for the Italian case, radar maps are not systematically accessible. The high potentiality of this instrument requires further efforts for improving his usability in the spatial analysis of severe storm systems.

This chapter focuses on the identification of the limits related to the classical approaches of radar rainfall estimation. After a brief analysis of the available methodology for increasing the accuracy of radar rainfall products, a technique aimed at increasing the reliability of radar rainfall estimates in the presence of

large rainfall amounts is proposed. The technique is intended for systematic and unsupervised operational applications.

3.1 Working principles, potentialities and limitations

A typical weather radar has three main components: (1) the transmitter, which generates short pulses in the microwave frequency portion, (2) the antenna, which focuses the transmitted energy into a narrow bean and (3) the receiver, that recives the backscattered radiation from the targets that intercept the transmitted pulses [22]. For further information on the working principles of the weather radar, please refer to the hydrological literature on the topic (e.g., [111]).

Radar does not provide direct measures of rainfall fields, however it is possible to estimate rainfall rates from the radar reflectivity measurements. Rainfall rate and radar reflectivity are linked by the drop size distribution (*DSD*), as recognized by Marshall and Palmer [132]. The use of standard diameter distributions leads to a power law relation between the radar reflectivity factor Z (mm^6/m^3) and the rainfall rate R (mm/h):

$$Z = a \cdot R^b \tag{3.1}$$

where a and b are coefficients related to the *DSD*. In what follows 3.1 will be referred as the *Z-R* relationship and the "radar reflectivity factor" just as "reflectivity".

As the *Z-R* relationship is inherently indeterminate, the coefficients a and b are usually estimated through an empirical approach, based on the comparison of radar and rain gauge data (e.g., [132, 102]). The estimation procedure is characterized by a high level of uncertainty, as can be recognized by the wide variety of coefficient values collected in [22, 158, 67]. A variety of papers highlights the variability of a and b according to the type of event [196, 34] and, during the event itself, in time and space [179, 119, 44]. Even with measured drop size distributions, recorded over a physically homogeneous period, the instrumental noise related to small sampling effect and the observational noise, due to the drop sorting effect, can lead to *Z-R* relationships which are only true in a statistical sense [119].

Moreover, the expected estimation quality is considerably influenced by radar characteristics, sampling methodologies, and by the characteristics of the area in which the estimation is performed [56, 117, 26]. The validity of an approach based on the use of a constant *Z-R* relationship should therefore be strictly limited to the calibration domain of the relation itself, making the relation unsuitable for any systematic operational use. The importance of distinguishing the search for a physical dependence between reflectivity and rainfall intensity in a specific context from the goal of obtaining the best radar estimation of the rainfall rate is well discussed in [47].

Optimal radar estimation of the rainfall rate, the main objective of this chapter, has been addressed by several papers in the literature. Some authors [9, 10, 31] propose to reduce the uncertainty of estimation by introducing additional information, e.g., type of precipitation, distance from the radar, etc.

Generally, a significant improvement in the estimation quality is obtained by calibrating equation 3.1 at the local scale, on the rain gauge nodes. This kind of approach reduces the problems related to orographic and climatological spatial variability. Two main categories of methodologies can be adopted: (i) a long-term (static) approach that involves the calculation of a single corrective term from an historical dataset of radar and rainfall measurements, to be subsequently applied to all the time frames in an identical way; and (ii) a short-term (dynamic) approach which requires the definition of a different calibration factor for each time step, based on the comparison between radar and rain gauge data from a definite time domain [213]. Several studies can be found describing the results of the second category of methods, with approaches suitable for different time scales: seasonal (e.g., [204, 164, 149]), monthly (e.g., [173, 120, 174]), or multi-daily (e.g., [165]).

To adapt the estimation to a single event, different methodologies have been proposed, that can be referred to as assimilation techniques of radar and gauge information to derive bias-corrected precipitation fields. Most of them stem from the pioneering work of Brandes [32] that involves the computation of a calibration matrix for each event, based on the smoothing of both the original radar field and of the subsequent correction factors. The key idea of this work has been seminal for a wide range of radar adjustment methods accounting for the spatial variability of the a coefficient of the *Z-R* relation in an operative context (see e.g., [141]).

As a conceptual evolution of the idea proposed in [32] recent methods propose adjustments of the relation 3.1 with *Z-R* data collected from moving windows at a sub-daily scale (e.g., [214, 6]). More recent studies also report on the combination of this kind of approach with appropriate interpolation techniques to significantly increase the estimation quality [207].

Some authors [80, 203] suggest another approach to increase the estimation quality, abandoning the classic radar-rain gauge comparison, and adopting different merging techniques, based on the use of geostatistical interpolators like kriging with external drift and co-kriging. More complex methodologies involve the use of advanced techniques such as statistical objective analysis [151] or Kalman filters (e.g., [195, 45]). Most of these methods are very time consuming and often not suited to operational contexts [80].

Among the most recently proposed methods, the main advancement achieved in the operational efficiency relates to the use of non-linear calibration (i.e., methods which allow a and b to vary simultaneously). As demonstrated in various studies (e.g., [35]) those methods systematically outperform linear methods (where only the a coefficient is let to vary) due to the high variability of the b coefficient in time and space [119, 41].

While most of the studies cited use either a spatial or a temporal point of view, in this chapter the problem is addressed from both the viewpoints at the same time, through a non-linear calibration procedure amenable to explain the variability of both the a and b coefficients simultaneously in space and time. A simple and operationally-ready methodology is proposed in this chapter. The technique, adaptive in time and independent of local conditions, produces an accurate radar-based estimation of rainfall intensity. The quasi real-time calibration method proposed, named Adaptive in Time and Space estimation technique (ATS) is summarized briefly in figure 3.1 and deeply described in section 3.4. Its quasi real-time nature implies that *Z-R* relationships are consistent with the evolution of the event, while the use of a spatially adaptive approach avoids the use of interpolation techniques, making the technique amenable in large areas with complex orography.

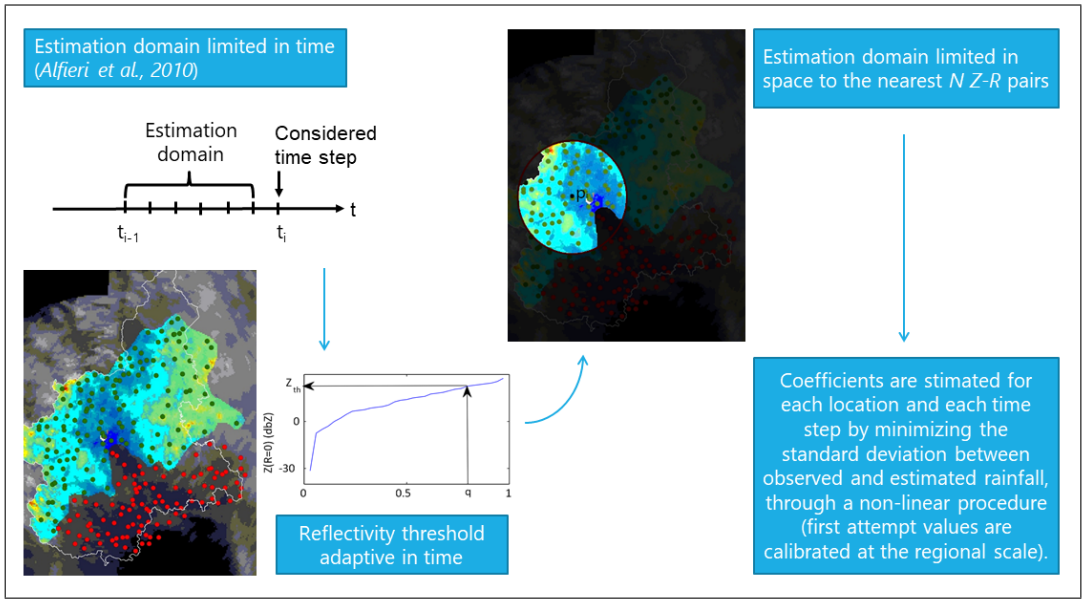


Fig. 3.1 Flow chart of the *ATS* methodology

3.2 Data and definition of the case study

Considering that in Italy weather radar data are not easily accessible, the domain in which the proposed method has been conceived and set up is an area of about 25000 km^2 in which the local agency (*ARPA* Piemonte) builds systematic radar maps. The outcomes of this analysis are easily applicable to other areas covered by C-band radars. In the detail, the area of interest is located in North-Western Italy, almost entirely within the borders of the Piemonte region. The area is characterized by a complex orography as shown by figure 3.2. The presence of orographic barriers significantly affect radar detection quality, due to the presence of echoes from the ground, occlusions of radar beam, orographic precipitation under beam, etc.

The "Bric della Croce" radar is a dual-polarization C-band doppler with digital receiver located at 736 m above sea level on the hills near Torino (7.734°E 45.035°N). *ARPA* Piemonte stores reflectivity maps on a Cartesian grid of 250 km $\times$ 250 km, with a resolution of 500 m in space and 10 minutes in time. The raw radar product is de-cluttered with a fuzzy logic algorithm [60] and the horizontal corrected reflectivity is projected on a 2D map by considering, at each point, the radar beam with lowest elevation available. This information is obtained comparing the terrain elevation for each cell of the gridded domain (from a 500 m x 500 m *DTM*) with the beam

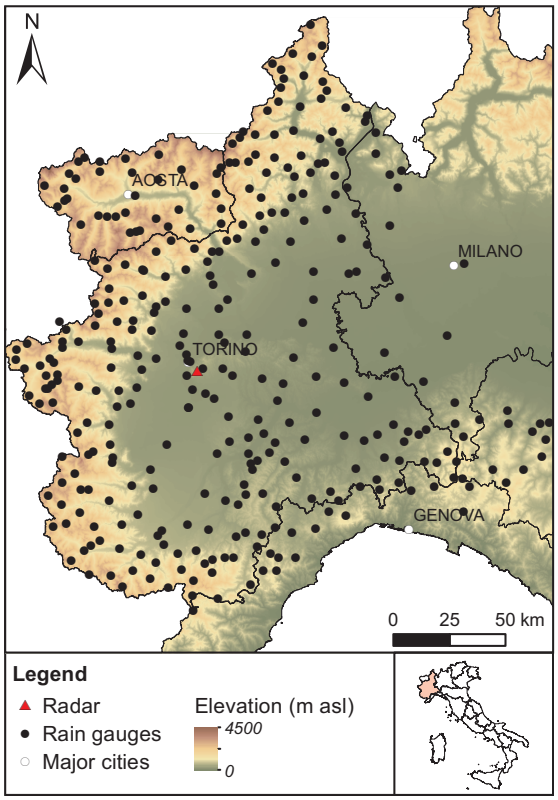


Fig. 3.2 Orography of the study area and locations of "Bric della Croce" radar and rain gauge network.

elevation. From October to May a VPR correction is applied. The mean VPR is calculated by averaging the vertical profiles falling into a neighborhood with 50 km radius from the radar site, during the previous 1-hour period.

ARPA Piemonte also manages the rain gauge network, which consists of 378 tipping bucket gauges distributed over the Piemonte region (25000 km^2). Rain data have a 10 minutes resolution in time and a minimum threshold of rainfall detection of 0.2 mm/10 min [83].

As the aim of this chapter is the definition of an adaptive calibration method particular attention should be paid to the quality of input data. The magnitude of the rain gauge errors is highly dependent on the local rainfall intensity and the timescale. For example, at moderate rainfall intensities of 10 mm/h [46] report relative standard errors of 4.9% for the 5-min and 2.9% for 15-min timescales. The available 10-minutes time resolution therefore results in a good compromise between relatively low standard errors and a good robustness at the hourly scale.

Much has already been said on the quality of the radar data in the previous sections. Supplementary analysis of the radar product quality, referred to the study case, is performed in the following, after the selection of the events.

The criteria that lead to the definition of the set of rainfall events to be used in this analysis comply with the need of representing the most intense precipitation events recently occurred in Piemonte that have involved a significant part of the regional territory. If on one hand the last assumption can lead to exclude very intense but localized events, on the other it is essential, at this stage, for a robust definition and verification of the methodology. Further refinements for very localized events can be pursued in a second stage. Two different categories of phenomena can be defined: (i) convective events, typically very localized and characterized by high intensities and short durations, and (ii) stratiform events, covering wider areas and characterized by lower intensities and longer durations. The most intense significant events occurred in the region from 2003 to 2008 are considered, with different criteria for convective and stratiform events, based on the duration and spatial distribution of the rainfall fields. Convective events are identified as those registering the annual maximum value (of hourly duration) in, at least, 22 rain gauges. Stratiform events, instead, are selected when the extension of the area registering the highest annual areal precipitation in 24 or 48 hours equals, at least, 1/4 of the regional area.

The 18 selected events are listed in table 3.1.

Table 3.1 Dates and codes of the selected events.

Convective		Stratiform	
ID	Dates	ID	Dates
1	27/07/2003	9	31/10-01/11/2003
2	02/08/2005	10	25/10-02/11/2004
3	20/08/2005	11	15/04-17/04/2004
4	06/07/2006	12	06/09-12/09/2005
5	12/07/2006	13	14/09-15/09/2006
6	08/08/2007	14	01/05-04/05/2007
7	30/08/2007	15	25/05-28/05/2007
8	29/05/2008	16	28/10-06/11/2008
		17	01/12-04/12/2003
		18	16/12-17/12/2008

To further evaluate the quality of the used radar products table 3.2 reports the rate between invalid radar record at gauge location and total number of records during the event. Events 1 and 6 report an anomalous high numbers of invalid radar records. They have not been excluded from the case study to define the impact of these anomalies on the results.

Table 3.2 Number of invalid radar records divided by the total number of records for each event (n_{inv}).

Convective		Stratiform	
ID	n_{inv}	ID	n_{inv}
1	0.254	9	0.001
2	0.003	10	0.004
3	0.003	11	0.005
4	0.004	12	0.051
5	0.003	13	0.003
6	0.461	14	0.003
7	0.002	15	0.073
8	0.050	16	0.011

3.3 Preliminary analysis and regional relationship

Before proceeding with the implementation of the adaptive methodology a regional analysis for calibrating a local *Z-R* relationship with the classic radar-rain gauge static approach has been developed. This is aimed at both assessing the limitations of the traditional methodologies and providing a solid background for the subsequent analysis.

At first, for limiting the amount of noise into the following calibration procedure, the radar visibility area is identified. This involves the identification of a relation between the estimation error and some descriptors of the radar beam geometry, aimed at excluding from the calibration domain areas where the beam attenuation lead to a scarcely significant radar information. In this phase the relationship considered for the *Z-R* conversion does not significantly affect the results. The coefficients commonly adopted for the estimation of radar rainfall from the local authority (a=200 and b=1.5 [102]) are then adopted as first-attempt values. A preliminary bias-correction factor based on the analysis of the estimation quality in the proximity of the radar, where the beam attenuation influence can be neglected, is calibrated

Once identified a proper visibility region, the regional relation is calibrated according to the procedure described in the second subsection, minimizing both the bias and the absolute error over the whole area.

3.3.1 Radar visibility map

A first evaluation of the estimation error was obtained by calculating the cumulative rainfall (R_{cum}) at the 378 rain gauges over a two-year time window for which continuous radar-rain gauge data are available. According to some preliminary tests, considering longer periods does not significantly improve the accuracy of the results. The values are compared with the estimated cumulative rainfall $\hat{R}_{cum}$, obtained by converting the radar reflectivity time series at the measuring points (adopting the coefficients suggested in [102]), and by cumulating the estimated precipitation over the whole period.

The differences between R_{cum} and $\hat{R}_{cum}$ are analyzed at first within a high-visibility area, including stations within 50 km from the radar with $h_A \leq 1500$ m

in order to assess the existence of a systematic bias, independent from the beam attenuation. Analyzing the dispersion of $\hat{R}_{cum}$ versus R_{cum} for the 58 rain gauges falling in this area, it is evident that all estimated values fall below the 1:1 line (see figure 3.3a). Therefore, before proceeding with further investigations, a bias-correction factor F, defined as the ratio between the average of the R_{cum} and $\hat{R}_{cum}$ for the 58 high-visibility stations, is applied. F=1.65 is obtained. Once removed the systematic bias, the errors still affecting the radar estimations (clearly shown in figure 3.3b) can be attributed to the different radar beam geometry in the correspondence of the stations. The relative errors (ε_{rel}) at each station are consequently obtained as:

$$\varepsilon_{rel} = \frac{\left| R_{cum} - F \cdot \hat{R}_{cum} \right|}{R_{cum}} \tag{3.2}$$

For each point of the region four descriptors are considered: the distance from the radar, the radar beam height (absolute height, h_A), the terrain elevation and the difference between these last two quantities (relative height, which is the minimum height of the beam above the ground). Relative errors are significantly correlated with 3 of the 4 above-mentioned descriptors. For example, using the absolute height of the beam, the spatial variability of the relative error can be described as:

$$\varepsilon_{rel} = -0.08 + \frac{0.15}{1000} \cdot h_a \tag{3.3}$$

that is the best relation found between errors and spatial covariates (with R^2=0.78).

Equation 3.3 allows one to associate an error level to each location as a function of the beam height. This relation enables one to define which stations to use during the calibration phase, using the beam height as a threshold to define the visibility area. In order to preserve as much information as possible, while excluding areas with poor visibility, a threshold value ε_{rel}=0.51, corresponding to a beam height of 4000 m a.s.l is selected. According to this threshold, the acceptance area shown in figure 3.4 is defined.

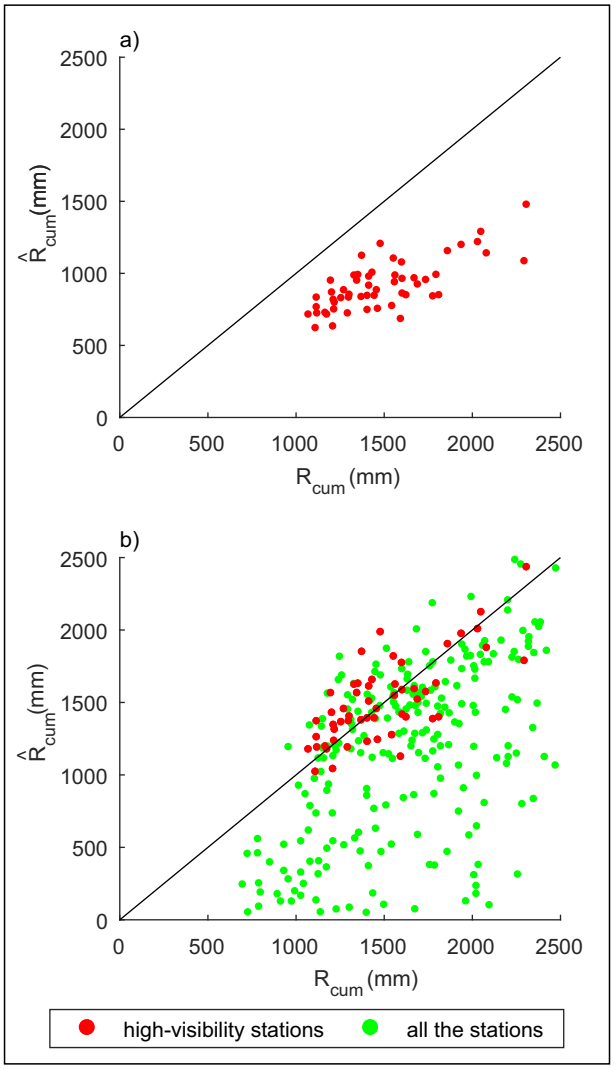


Fig. 3.3 (a) Cumulative rainfall recorded at-gauge (R_{cum}) and radar-estimated with the coefficients reported in [102] ($\hat{R}_{cum}$) over a two year time windows in the high-visibility area. (b) BIAS-corrected cumulative rainfall for all the considered stations.

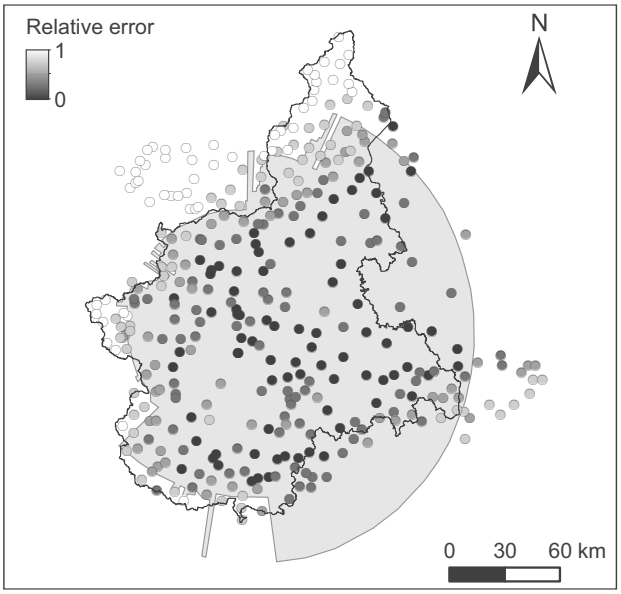


Fig. 3.4 The defined acceptance area for the estimation of the *Z-R* relationship (gray background). The grey scale refers to the stations relative errors.

3.3.2 Regional *Z-R* static relationship

A local version of 3.1 is then calibrated over the previously identified visibility area, to define a robust static (in time and space) relationship suitable for the considered region. At this stage, in order to reduce the processing time, preliminary "data binning" is applied; i.e. radar reflectivity data are grouped into classes, with minimum width equal to the resolution of the radar data (0.5 dBZ). The median value is chosen as representative of each class. Classes with less than 10 items (i.e., the tails of the distribution) are assembled, to increase the robustness of the estimators. To each reflectivity class Z, a value of R equal to the average of the corresponding measured rainfall was associated.

The calibration procedure, described in the following of this section, involves reconsidering equation 3.1 on the whole acceptance area in order to obtain regional estimates of the a and b coefficients. The calibration methodology is targeted at minimizing both the absolute estimation error ε_{abs}:

$$\varepsilon_{abs} = \sum_{j=1}^{n} \sum_{i=1}^{d_n} \left| \hat{R}_{i,j} - R_{i,j} \right| \tag{3.4}$$

and the bias:

$$bias = \frac{1}{n}\sum_{j=1}^{n}\left|\sum_{i=1}^{d_n}(\hat{R}_{i,j} - R_{i,j})\right| \tag{3.5}$$

in the parameter space ($1< a <1000$, $1< b <4$). In both equation 3.4 and 3.5, $R_{i,j}$ and $\hat{R}_{i,j}$ are the rainfall observed and estimated using 3.1 by varying a and b coefficients, respectively at t_i in p_j; n is the number of rain gauges and d_n the number of records per rain gauge.

The two are combined for each event to define the index $I_{3,ev}$, that takes into account the quality of the estimation both at the hourly and event-scale.:

$$I_{3,ev} = I_{1,ev} + I_{2,ev} \tag{3.6}$$

where:
$I_{1,ev} = \left(\frac{\varepsilon_{abs}}{min(\varepsilon_{abs})} - 1\right) \cdot 100$
and
$I_{2,ev} = \left(\frac{bias}{min(bias)} - 1\right) \cdot 100$
in which $min(\varepsilon_{abs})$ and $min(bias)$ are the minima of ε_{abs} and $bias$, evaluated in the considered parameter space.

The optimal pairs of a and b values, reported in figure 3.5a, are obtained by minimizing $I_{3,ev}$ for 8 convective and 10 stratiform events (see table 3.1). Events 17 and 18, occurred in December 2003 and December 2008, are excluded from the subsequent processing since they appear isolated from the others, probably due to the presence of snow. No significant between convective and stratiform emerges from figure 3.5a.

The global optimal values of a and b are then obtained by applying the described procedure to the Z-R pairs of the whole set of events (i.e., merging convective and stratiform events).

The regional static relation reads:

$$Z = 40 \cdot \hat{R}^{2.5} \tag{3.7}$$

where the (*a,b*) pair is represented by the big black dot in figure 3.5.

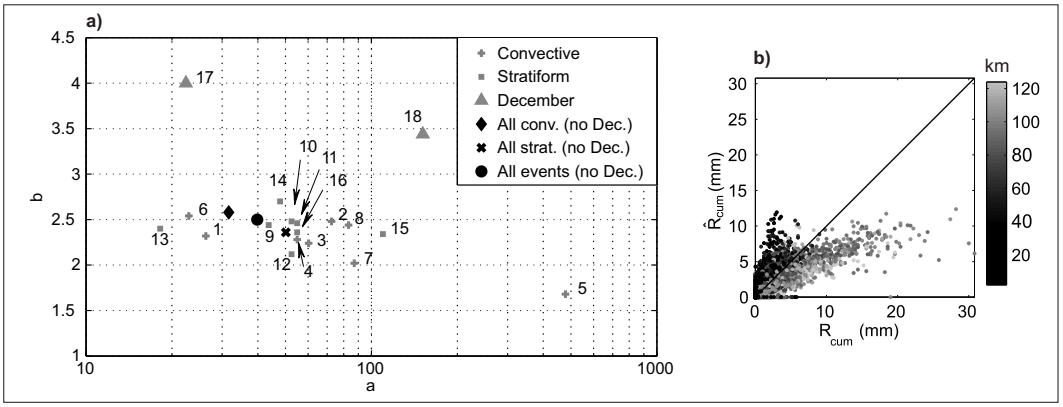


Fig. 3.5 (a) Optimal $(a-b)$ values of the *Z-R* relation for each single event, for the different event categories and for the whole set of events. (b) Comparison between observed precipitation R_{cum} and rainfall estimated with the regional formula $\hat{R}_{cum}$ for the event occurred on 10/31-11/1/2003 at the event scale. (The grey scale refers to rain gauge distance from radar in km).

Estimates obtained for the selected events with the regional relationship 3.7 systematically outperform those obtained with commonly used relations (e.g., [102]) as reported in figure 3.8. However, the improvement is not consistent for all the events and it seems to be significantly related with the nature of the analyzed rainfall system. Limitations due to the use of a single relation on a vast and complex territory, leading, in some cases, to poor reconstruction of the precipitation volumes are yet to be overcome. Figure 3.5b shows an example of the presence of over/underestimation clusters as a function of the distance of the location from the radar.

3.4 The Adaptive in Time and Space estimation technique

A methodology aimed at accounting for the spatio-temporal variability of the *Z-R* relation by means of an adaptive estimation of the two coefficients of 3.1 is then proposed. The methodology, called Adaptive in Time and Space estimation technique (*ATS*) relies on the definition of calibration domains limited in time (section 3.4.1) and space (section 3.4.2).

In the following the rationale for the selection of the domains and the related calibration procedures in time and space are presented.

3.4.1 Definition of the time domain

To face the variability of the *Z-R* relation in time a quasi real-time calibration window adapted from [6] is proposed. Coefficients a and b are estimated for each time step t_i considering the *Z-R* pairs belonging to a calibration window of duration D, $[t_{i-D};t_i]$, as shown in figure 3.6a.

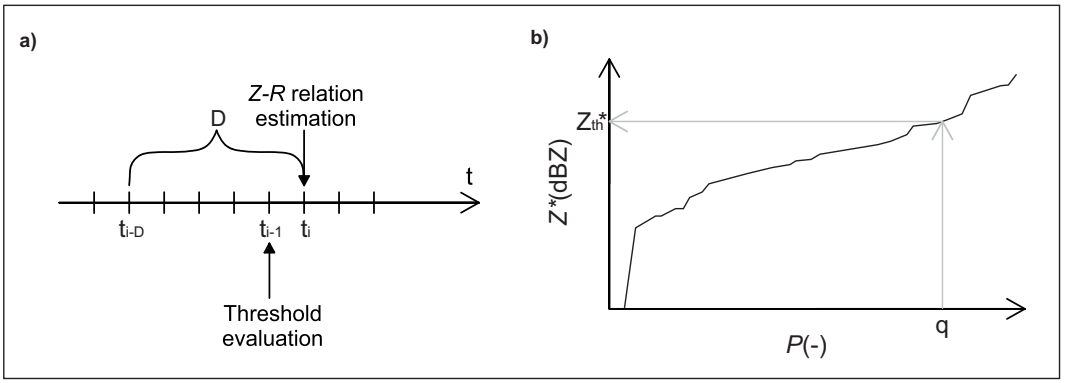


Fig. 3.6 (a) Schematic representation of the temporal estimation domain and (b) identification of the reflectivity threshold for a generic time step.

Such an approach allows one to follow the evolution of the event in quasi real-time, calibrating a different power-law relationship 3.1 at each time step. The inconvenience of this approach is that the high amount of noise can lead to highly variable estimates of the a and b coefficients due to the small size of the estimation domains.

To overcome this problem a methodology involving a systematic "cleaning" of the data is introduced. This involves the use of a reflectivity threshold variable in time to exclude from the estimation domain *Z-R* pairs with rainfall rate at or near zero. Working at a low rainfall rate scale, small differences between observed amounts and uncalibrated radar data can lead to spuriously large or small calibration factors [32]. Furthermore, the threshold is used to represents the amount of noise of the radar measurements due to instrumental and sampling uncertainly. The threshold is also used to discriminate between the presence and absence of rainfall at the considered

time step: the rainfall rate is systematically set null for the locations characterized by under-threshold reflectivity values.

The threshold is calibrated for each t_i by selecting the rain gauges that have registered no rainfall at the instant t_{i-1}. The threshold value is then defined as the quantile with cumulative probability q of the empirical distribution of the reflectivity data associated to the selected rain gauges. Considering that the empirical distribution changes at each time step, fixing a value of q, the threshold values vary in time according to the evolution of the event. The choice of the optimal q value is discussed in section 3.5.1. The threshold identification procedure for a generic instant t_i is exemplified in figure 3.6b. The graph shows the cumulative empirical distribution of the reflectivity values related to absence of rainfall for a generic instant t_{i-1}. The threshold value Z^*_{th} for the instant t_i is evaluated considering the empirical quantile related to the cumulative probability P=q.

After some preliminary tests, for a 10 minute resolution, a time window of 60 minutes is assessed to provide sufficient information for a robust *Z-R* calibration. To reduce the degrees of freedom of the procedure, the width of the time window is fixed.

3.4.2 Definition of the spatial domains and *Z-R* relationship estimation

To account for the spatial variability of reflectivity, the calibration domain is further confined, for each location p_j, to the N nearest rain gauges, using only above-threshold reflectivity values. In other words, for each pixel of the gridded study area, a specific calibration domain of N station is defined, N being therefore an important pre-determinate parameter in the procedure. The selection of the optimal N is widely discussed in section 3.5.1. If the valid number of *Z-R* pairs (i.e., the above-threshold pairs into the whole estimation domain) is less than N, an estimation with the available pairs is attempted.

Considering that the threshold value and the spatial distribution of rainfall change continuously, the procedure requires to re-consider the number and position of valid stations for each location p_j at each time step t_i.

Minimizing the sum of the squared differences between the observed and the estimated rainfall for each local domain and in each time step a pair of optimal a and b coefficients is obtained. The estimated rainfall $\hat{R}$ is evaluated according to the following relationship, obtained from 3.1, cleaning the recorded reflectivity with the threshold value Z^*_{th} :

$$\hat{R} = 10^{\frac{Z^* - Z^*_{th}}{10 \cdot \hat{b}} - \frac{log_{10}\hat{a}}{\hat{b}}} \tag{3.8}$$

where $\hat{a}$ and $\hat{b}$ are the estimated coefficients of the *Z-R* relationship, Z^* (dBZ) and Z^*_{th} (dBZ) the radar reflectivity and the reflectivity threshold respectively, both expressed in decibels ($Z^* = 10 \cdot log_{10}Z$). As the definition of the local domains refers to the nearest rain gauge, local domains referred to near locations partially overlap one another, granting a smooth variation of the coefficients and the continuity of the radar Probabilistic Quantitative Precipitation Estimates (*QPE*).

The estimation procedure is carried out in the *Z-R* plane by non-linear regression. The adopted optimization algorithm is a subspace trust region method and is based on the interior-reflective Newton method [54]. The optimized coefficients are estimated iteratively using as initial values the "static" coefficients calibrated at the regional scale (see section 3.3.2). If the number of *Z-R* pairs is not sufficient for a robust estimation (i.e., if the algorithm does not converge after 400 iterations), the backup regional relationship is used (see section 3.3.2).

Calibration and verification procedures are carried out in cross-validation mode for each considered rain gauge, i.e., excluding one rain gauge at a time from the evaluation of the *Z-R* relationship and then comparing the estimated rainfall depth with the actual measurement at the excluded rain gauge.

3.5 Application and results

3.5.1 Calibration of the *ATS* technique

As stated in section 3.4.1, the width d of the calibration window is set to one hour, in order to limit the influence of the temporal variability in the adaptive search for

optimal parameter values. The use of *Z-R* data recorded with a 10 minutes frequency grants a good estimation robustness (5 samples for each window, each characterized by a number of observations equal to the number of considered rain gauges).

The *ATS* methodology is characterized by two calibration parameters: N and q, indicating the number of rain gauges in the local domain and the probability used for the definition of the zero-rainfall threshold respectively. In order to define the best N and q parameter values the estimation efficiency at both hourly and event scale on the considered set of events is maximized, by considering jointly the absolute estimation error (eq. 3.4) related to the hourly scale and the bias (eq. 3.5), related to the event scale. The calibration of N and q is therefore carried out by exploring the parameter space $0 \leq q \leq 0.9$ and $0 \leq N \leq 80$ (q=0 corresponding to the application of the methodology without any threshold) and minimizing the performance index I_3 previously defined (eq. 3.6), for each event (*ev*). Optimal q and N are then chosen for each event as the values corresponding to the I_3 index minima.

The mean of the event patterns, each normalized to the maxima of $I_{3,ev}$ for the event, is plotted in figure 3.7. The minimum of the $\bar{I}_3$ index in the *N-q* parameter space identifies the optimal pair of values for the whole set of events. To evaluate the presence of different behaviours among the different event typologies, the calibration procedure is carried out at first on the whole set (figure 3.7a), then on the convective ones (figure 3.7b) and finally on the stratiform ones (figure 3.7c).

For convective events, due to the high spatial variability of rainfall fields, a rather high q threshold is required, in order to exclude "false positive" values (i.e., large reflectivity values associated with absence of rainfall). The number of rain gauges in the local domain (N) seems to play a less important role, if a minimum of 5 rain gauges (corresponding to 25 *Z-R* pairs) is ensured. Below this value, the quality of the estimations rapidly declines, as shown in the panel (b) of figure 3.7.

For stratiform events the number of rain gauges to be included in the domain becomes crucial and should be appropriately set (usually between 10 and 30) in order to avoid including in the estimation procedure stations with low information content. In this case the choice of the quantile q is less influential (see panel (c)), due to the low variability of the reflectivity values.

The use of different *N-q* pairs for each category of events does not improve significantly the estimation quality. Therefore, in order to propose a methodology

Table 3.3 Mean and standard deviation (*Std*) of the threshold values for the analysed events.

Convective			Stratiform		
Event	Mean (dbZ)	*Std* (dbZ)	Event	Mean (dbZ)	*Std* (dbZ)
1	-2.06	6.51	9	20.94	6.96
2	14.94	7.52	10	11.77	10.74
3	9.55	10.70	11	11.88	9.85
4	5.17	11.51	12	5.59	9.55
5	2.16	8.66	13	16.73	7.99
6	2.58	9.97	14	9.78	10.38
7	4.45	9.54	15	9.71	8.36
8	19.73	8.17	16	17.70	6.30

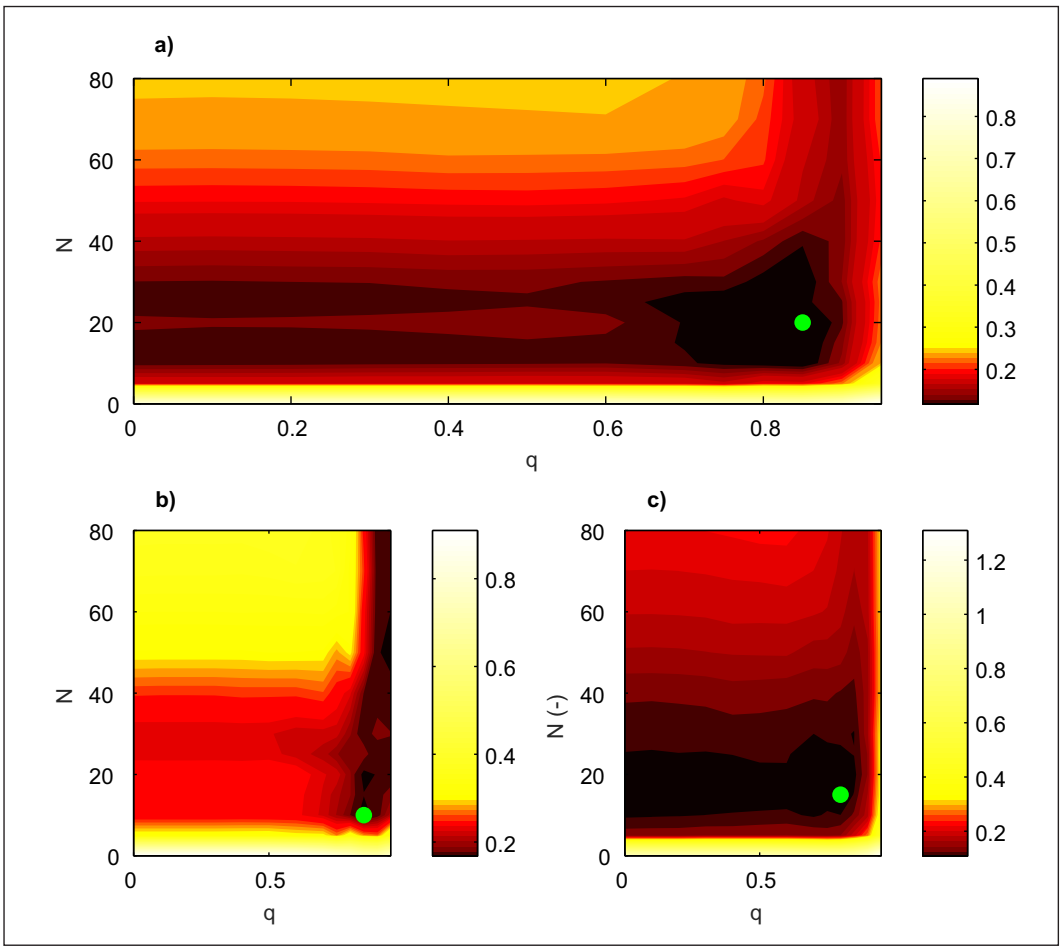


Fig. 3.7 Pattern of the $\bar{I}_3$ index into the parameter space q-N (a) for all the selected events, (b) for convective only, (c) and for stratiform only events. $q = 0$ implies the application of the methodology without any threshold. The green dots indicate the location of the minima.

robust and easily applicable in real-time, the suggested optimal values for both convective and stratiform events are $N = 20$ and $q = 85\%$, corresponding to the best global mean value of the $\bar{I}_3$ index (figure 3.7a).

The limit of $N = 20$ in the spatial domain can grant a good robustness at the adopted time resolutions, and the relatively high quantile q for the definition of the threshold makes the methodology amenable to different kind of events. To support this statement, table 3.3 shows that the adopted quantile lets the threshold vary consistently even in the considered case study, without significant loss of performance with respect to the event-variable q values displayed by panels (b) and (c) of figure 3.7.

3.5.2 Results

The *ATS* procedure is applied to the 16 events listed in table 3.1 using the q-N pair defined in the previous section. The performances of the procedure are compared with two static methodologies: the Joss-Waldvogel formula [102], routinely used for rainfall estimation over the study area, and the regional relation 3.7. Our procedure is also compared with the adaptive methodology proposed by Brandes [32], that entails the estimation of a corrective factor at each rain gauge site, with a radar-gauge comparison carried out in real time. All factors are then interpolated on the whole radar field with the procedure described in [80]. Operatively, as first-attempt relation the regional one 3.7 is adopted. To avoid the problems related to the low rainfall rate scale, only the pairs for which both recorded and estimated rainfall exceed 2.5 mm are considered.

The quality indicators are defined at both the hourly and the event scale. Figure 3.8a,b show the coefficients of determination obtained with the four methodologies for all the events. Figure 3.8c shows the bias values.

The results confirm the validity of the proposed methodology that allows us to obtain lower volumetric errors for almost all of the events, with respect to both the static formulations. In addition, the correlation coefficients, not subjected to optimization during the calibration phase, show a general improvement in the estimation quality at both the hourly and the event scale.

As for the comparison with the adaptive methodology, the ATS technique shows remarkable improvements at the event scale, due to the contribution of the adaptive threshold to discriminate between presence and absence of rainfall, that leads to a more affordable estimation of the cumulative rainfall at the event scale. The improvement is greater for stratiform events, thanks to the spatial uniformity of these events, allowing an easier identification of the local spatial domains. For convective events the improvements are less marked, as the spatial variability of rainfall fields often leads the ATS technique to work as a regional estimator, similar to the one adopted in [32].

Events 1 and 6 are the only exceptions to the generally good performances. They are characterized by a generalized deterioration of the estimation quality, partly attributable to the poorer quality of the available data. The number of invalid radar record is indeed higher in events 1 and 6 than in all the other events, as reported in table 3.2. This reduces the quality of results, as the efficiency of the proposed method, that involves a dynamic calibration, is quite sensitive to the quality of the input data.

To underline the validity of the proposed technique, in figure 3.9 the measured precipitation with the regional estimate (top graphs) and with the ATS estimate (bottom graphs) are compared for 3 different events.

Graphs in column (a) and (b) show the comparison between the cumulative rainfall obtained with the regional relationship and the one obtained with the ATS technique, for the events 15 and 13, both considered at the event scale. The greater ability of the ATS technique to retrace the event is clearly apparent.

The scatter plots in column (c) refer to the event 2, where data is aggregated at the hourly scale. In this case there is a deterioration of the determination coefficient, essentially due to the underestimation of a (limited) number of high values (that are underestimated also by the regional relationship).

The results demonstrate that the proposed technique is particularly suitable for stratiform events. The wide spatial scale allows for an easy identification of the local spatial domain. Figure 3.9b shows the clearly improved estimation quality obtainable with the ATS technique with respect to a unique Z-R relation for a large-scale event.

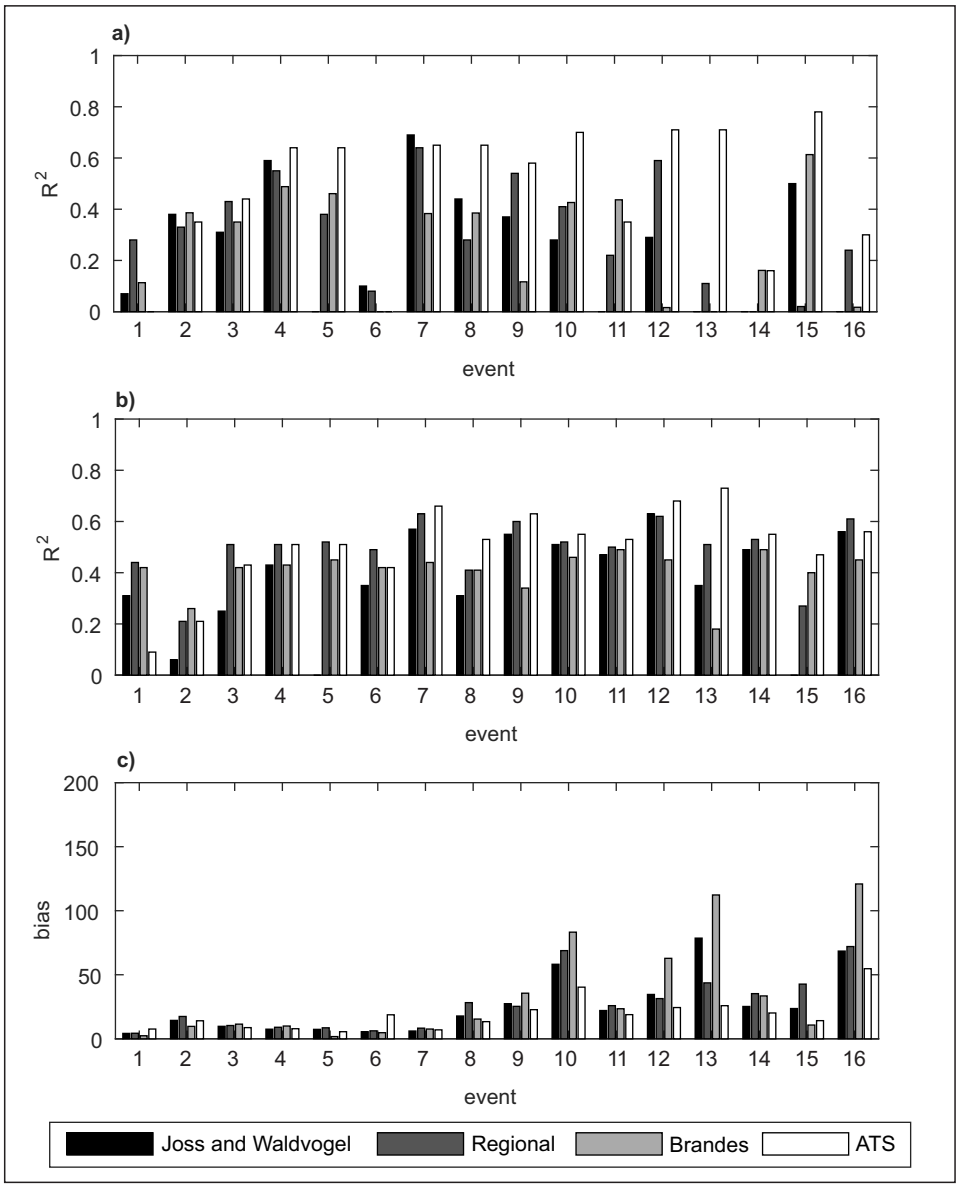


Fig. 3.8 Comparison between the coefficients of determination obtained with the Joss-Waldvogel formula, the regional formula, the methodology proposed in [32] and the ATS technique (a) at the event scale and (b) at the hourly scale. Correlation coefficient R^2 falling outside the range (0,1) have not been reported. (c) reports the comparison between the bias obtained with the four above-mentioned methodologies at the event scale.

The increase in the estimation quality is less significant for convective events, where the localized nature of the high values of reflectivity signal contrasts with the uneven spatial distribution of the rain gauges, preventing the identification of a uniform and numerically robust spatial domain. In these cases, to reach a sufficient number of representative station the *ATS* procedure would require to use distant *Z-R* pairs, that can be poorly representative of the event core. Moreover, due to the rapid evolution in time, an hourly calibration window can be still too wide, leading to the use of non-representative *Z-R* pairs.

Given that the methodology has different performances for different spatial scales of the events, its validity for convective rainstorms needs to be further assessed.

A more dense rain gauge network and the availability of radar-rainfall data with higher temporal resolution may facilitate the use of the adaptive approach, increasing the estimation quality, also for convective events.

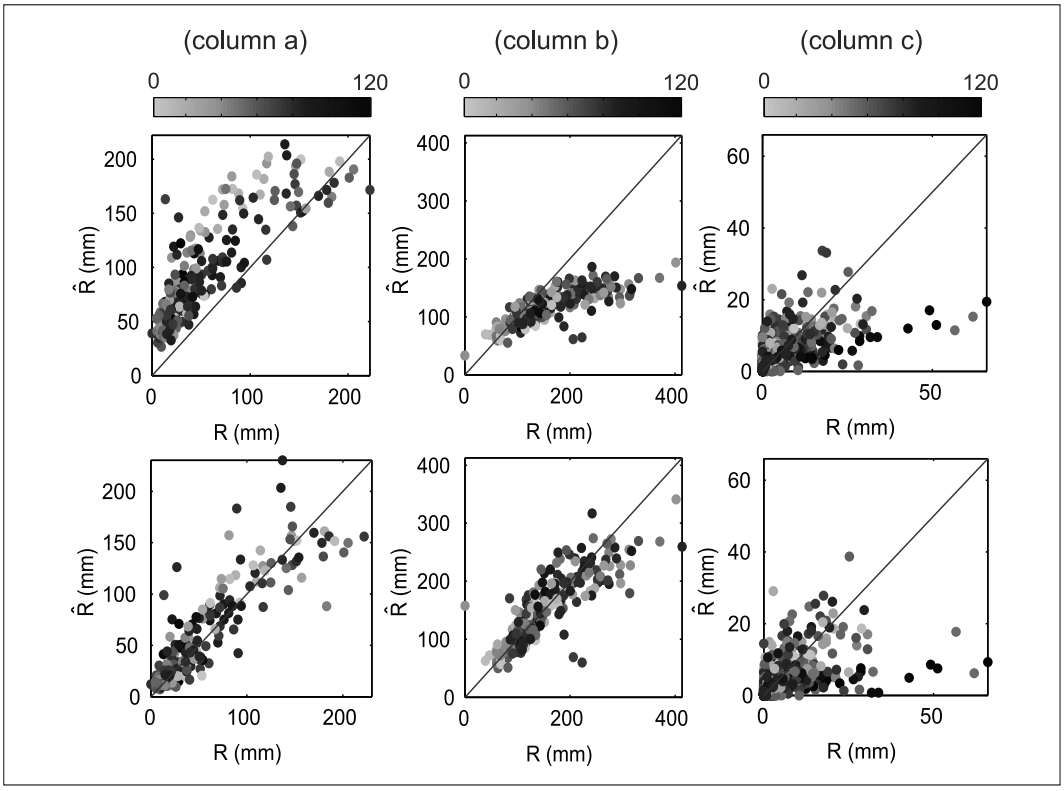


Fig. 3.9 Comparison between observed precipitation and rainfall estimated with the regional formula (top graphs) and with the *ATS* technique (bottom graphs). The comparison is made at the event scale for event 15 (column a) and 13 (column b). Comparison in (column c) refers to event 3 at the hourly scale. (The grey scale refers to the distance from radar, in km).

3.6 Concluding remarks

The proposed methodology adopts confined spatial and temporal domains for a quasi real-time calibration of the relation between radar reflectivity and rainfall rate at a local scale. By doing so, the spatial and temporal variability of the *Z-R* relation is considered, making the technique suitable for systematic operational use and building systematic high-resolution rainfall maps, regardless of local conditions, characteristics of the radar, sampling methodologies and spatio-temporal distribution of the events under analysis.

An accurate pre-processing of both radar and rainfall measurements is required in order to maximize the quality of the reconstruction. Even though the technique has been proposed for a real-time use it can be a good candidate for applications finalized to building up long records of rainfall depth maps. After refinements needed to improve the performances on convective events, this technique could provide valuable information for several applications of rainfall spatial analysis.

Among the possible further refinements for the proposed methodology, particularly interesting is the possibility to vary the parameters of the calibration windows in time (e.g., time window amplitude, numbers of rain gauge into the local domain, percentile to be used to define the threshold), according to the evolution of the event. In order to make the procedure more robust and accurate, it could be also helpful to use multiple regression techniques during the calibration phase, by considering other radar variables (e.g., specific differential phase, differential reflectivity, etc.).

Chapter 4

Remote-sensing of rainfall extremes: synoptic scale and satellite data

Part of the work described in this chapter has been previously published in an article in a peer-reviewed international journal [124].

Even when a well designed rain gauge network exists, it is not always sufficient to catch very localized and severe thunderstorms, as the uneven spatial distribution of the networks and the complex orography reduce the chances of fully observing extreme rainstorms. This criticality increases at higher elevations, where the gauge density is smaller. As seen previously, weather radars can effectively provide high-resolution coverage at the mesoscale domain, but with limited coverage and availability. For the purpose of this research, the access to systematic measurements, without gaps and operating at the synoptic scale as those provided by meteorological satellites, is of great interest.

This chapter documents an explorative analysis that investigates both the consolidated product of the Tropical Rainfall Measuring Mission ($TRMM$) and the first products of a new generation of weather satellites, which can have significant potential in the coming years. In the detail, the focus is on the usability of data provided by the $TRMM$ and Global Precipitation Measurement (GPM) missions in the study of the spatio-temporal characteristics of severe rainfall. Once the significant bias affecting the rainfall depth estimation is assessed, the ability of satellite products to represent the timing of severe storms is explored.

Considering the wide nature of satellite data, and the shortness of the available time series, the Italian scale is temporally abandoned for carrying out our analysis considering the whole domain of *TRMM* data. By using a "trading space for time approach" [177, 150], such a wide and global scale allows us to overcome the lack of robustness, which is due to the limited length of the available *TRMM* and *GPM* series and provide useful information that can be then applied to a wide range of spatial scales.

Even if a statistical characterization of rainfall with large return periods from satellite data is yet unfeasible, this analysis can help exploiting all the potential of those data for the analysis of severe storm. Satellite-derived timing information can help in reconstructing the synoptic influences on the local spatial variability, allowing a systematic analysis of the synoptic configurations concurrent to severe rainfall patterns. This could benefit the estimation of intensity-frequency-duration relationships in ungauged areas and the characterization of extreme rainfall in poorly gauged regions of the world, helping to clarify the connections between large-scale meteorological systems and actual rainfall distribution in space. Furthermore, satellite products can help connecting recorded rainfall with the tracks of severe storms, allowing a high resolution analysis over wide areas, including the seas (e.g., [75]). Moreover, large scale timing information can be useful in many other fields of the hydrological sciences which focus on the joint occurrence of severe rainfall and other phenomena (e.g., soil erosion, landslides, etc. [206, 53]). In fact, by expanding the analysis at the global scale, satellite data can drastically improve the sample size of these studies with benefits for the robustness of the outcomes.

The following questions are investigated: (a) Does the accuracy of the timing of extreme rainfall estimated by satellite vary with latitude? (b) Does this accuracy change with improvement in satellite spatio-temporal resolution and, more specifically, does rainfall data from *GPM* perform better than data from *TRMM*? And, (c) how does accuracy vary with *GPM* spatial resolution?

4.1 An exploratory case study: the Liguria region

The Liguria region is considered as a challenging introductory case study, as it possesses numerous characteristics that prevent a direct use of *TRMM* data (e.g.,

high latitudes, complex morphology, etc.). At first, the ability of satellite products to reconstruct the magnitude of extreme rainfall is tested considering the three largest daily rainfall depths recorded each year between 1998 and 2014 by the rain gauges of the Liguria regional network. Results are reported in figure 4.1b. An average value of 11% underestimation can be pointed out. Indeed, while a slight overestimation can be recorded for the smaller amounts, the curve [52] shows an increasing tendency starting from at-gauge values (R) of around 100 mm. The underestimation reaches values near 100% for the largest rainfall quantities. The average underestimation increases to 16% when only considering gauges above 400 m *asl* (the average elevation of the gauge network). More significant differences would be expected if a larger number of gauges were to be available at higher elevations. This dependence on the elevation of the station prevent the adoption of a straightforward bias correction of $TRMM$ data.

It is therefore evident that $TRMM$ data are still far from being useful when intense rainfall estimation is the target. They may however provide useful information on the spatio-temporal distribution of extremes.

The timing of the rainfall events considered in the Liguria case study (i.e. the three largest daily rainfall depths recorded each year between 1998 and 2014 by each rain gauge) is shown in panel (c) of 4.1. Apparently, the timing is distributed throughout the year, with a low density in the central dry season (i.e. from June to August). In about 25% of the cases, $TRMM$ identifies the correct date of occurrence of each considered event, confirming the findings of other studies (e.g., [184]). If only the rain gauges above 400 m *asl* are considered there is no significant variation in the match percentage; this finding possibly denotes a behavior unrelated to elevation. Figure 4.2 shows the time series of the "Genova Quezzi" rain gauge during year 2011 along with the signal from the overlying $TRMM$ cell: the satellite correctly identifies the days with intense precipitation, even if underestimating the rainfall amount.

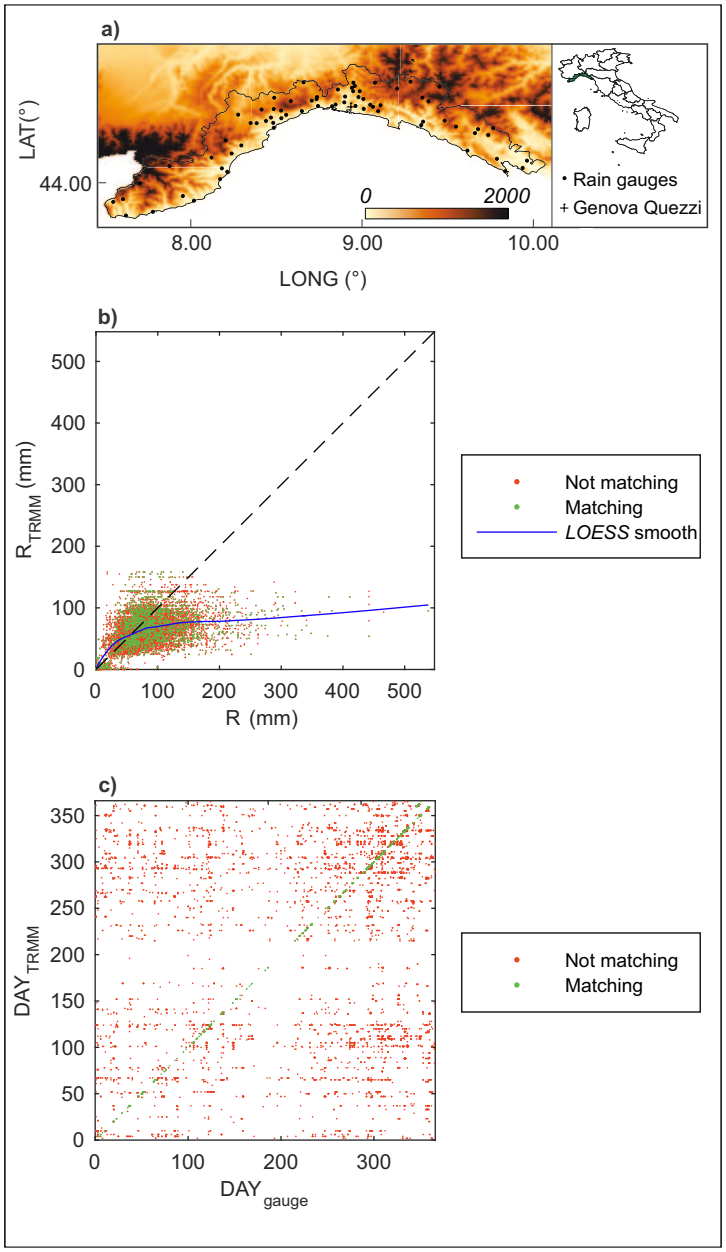


Fig. 4.1 (a) The Liguria region, in North-Western Italy. Orography and localization of the regional rain gauge network. The chromatic scale refers to the altitude in meters above the sea level (b) 3 largest annual daily rainfall recorded at each rain gauge (R) compared with the values of the overlying $TRMM$ cells (R_{TRMM}) in the 1998-2014 period. (c) Temporal distribution along the year of the matching and not-matching 3 major events. Green dots refer to days ranked likewise in the rain gauge and $TRMM$ series.

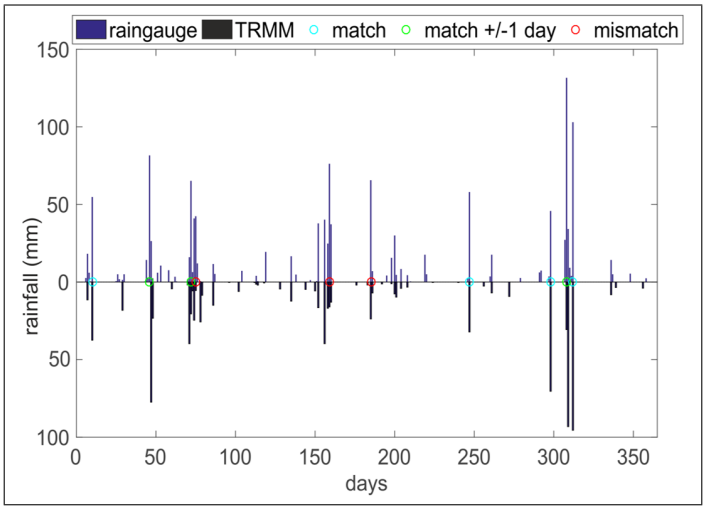


Fig. 4.2 The 2011 daily rainfall series of the rain gauge "Genova Quezzi" (44.237°N, 8.9726°E, 200 m *asl*, WGS84) and of the overlying *TRMM* cell.

4.2 Assessment of the timing of extreme rainfall using satellite data

4.2.1 Data

The assessment of the ability of *TRMM* and *GPM* to evaluate the date of occurrence of intense rainfall events is performed on data spanning from 1998 onwards. The period coincides with the duration of the Tropical Rainfall Measurement and Global Precipitation Measurement missions.

TRMM provides a range of rainfall products. In this work the *TMPA* (*TRMM* Multi-satellite Precipitation Analysis) 3B42 v.7 precipitation dataset is analysed. *TMPA* rainfall estimates are obtained by combining *TRMM* Precipitation Radar (*PR*), Passive MicroWave (*PMW*), and InfraRed (*IR*) estimates within a 3-hour window centered on a synoptic time (0, 3, 6, 12, 21 UTC) over the 50°S-50°N area (for more details, see [96]). From October 2014, with the decommissioning of the *PR*, a climatologically calibrated/adjusted research *TMPA* is available [20]. *TMPA* rainfall has high spatial (0.25°) and temporal (3 h) resolution and is widely applied in different branches of the earth sciences, especially in data-sparse regions (e.g., [20, 105, 18, 27]).

From 1 April 2014 the analysis also includes the final post-real-time run of 3*IMERGHH* product from the *GPM* mission. This product is characterized by a finer spatial resolution (0.1°) and is generated on half-hourly intervals (0, 0:30, 1, 1:30, …, 23:30 UTC), over the 60°S-60°N region (for more details, see [95]). At the time of writing the product is available until 30 June 2015.

To evaluate the impact of the varying spatial resolution, products are considered at both their natural spatial resolution (0.1°x0.1°) and at coarse resolutions of 0.2° x 0.2° and 0.3° x 0.3°.

The $NOAA-GHCND$ rainfall dataset v 3.22 [135] is adopted; it contains daily records from over 75000 stations in 179 countries [134].

Before being aggregated to the daily scale, satellite data at the original time resolution are shifted to best match the rain gauge data, considering the combined effect of time zones and national sampling practices. When the latter is not available, the best match compatible with the time zone shift is selected through a robust statistical approach, carried out at the national scale. Figure 4.3 displays the locations of the considered stations. Only those stations which have recorded at least 30 days of data in the 1998-2015 period are considered. Figure 4.3 also shows the distribution of stations per latitude interval.

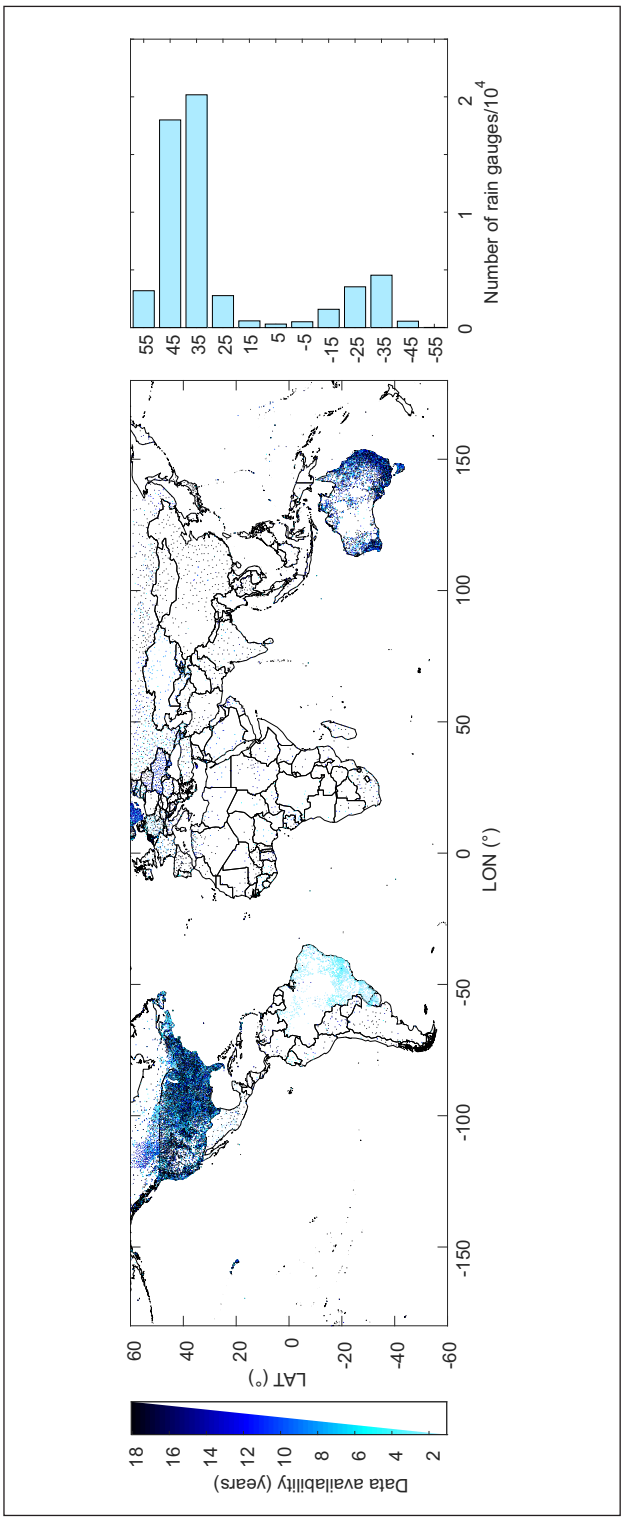


Fig. 4.3 Subset of the *GHCN – DAILY* database considered. Only stations that have recorded at least 30 days of data in the 1998-2015 period are reported. The chromatic scale refers to the length of the time series per rain gauge. The bar graph reports the distribution of stations per latitude interval.

Table 4.1 Periods of the analysis and characteristics of the satellite products. The "N.cells" field refers to the number of cells considered (i.e. the number of cells containing at least 1 rain gauge). The 3*IMERGHH* dataset is analyzed at the original spatial resolution (0.1°) and at coarser resolution of 0.2° and 0.3°.

Period	Dates	Product	Source	Resolution	Coverage	N.cells
(I)	01.Jan.1998 31.Dec.2013	*TMPA* 3B42v7	*TRMM*	0.25 - 3h °	50°S-50°N	17877
(II)	01.Apr.2014	*TMPA* 3B42v7	*TRMM*	0.25° - 3 h	50°S-50°N	13808
	30.Jun.2015	3*IMERGHH*	*GPM* 0.1	0.10° - 0.5 h	60°S-60°N	24762
		3*IMERGHH*	*GPM* 0.2	0.20° - 0.5 h	60°S-60°N	17746
		3*IMERGHH*	*GPM* 0.3	0.30° - 0.5 h	60°S-60°N	13140

In order to consider the different temporal coverage of the analyzed products, the analysis is carried out for two different time periods: (I) 1 January 1998 to 31 December 2013, for which only *TRMM* products are available; (II) 1 April 2014 to 30 June 2015 for which *GPM* products are also analyzed. A brief summary of the characteristics of the analyzed products is reported in Table 4.1 for each period.

4.2.2 Methodology

The methodology adopted for the comparison is summarized in figure 4.4.

The analysis is carried out on a gridded domain. The cell size is set to the spatial resolution of the considered satellite product (see Table 4.1). Each rain gauge is assigned to the related cell, according to its position. In case of cells with multiple records in a day, the largest value is considered (i.e., if multiple rain gauges belong to one cell, only consider the maximum rainfall depth recorded each day is considered).

For each year and cell, the N_{tot} most significant daily events identified in the satellite annual time series are compared with the N_g most significant ones recorded by the rain gauges. The dates of the occurrence of these events are now referred to as "satellite significant dates" and "rain gauge significant dates", respectively.

The value of N_{tot}=5 is set through a preliminary analysis on a subset of the data, aimed at increasing the robustness of the results and limiting the noise in the

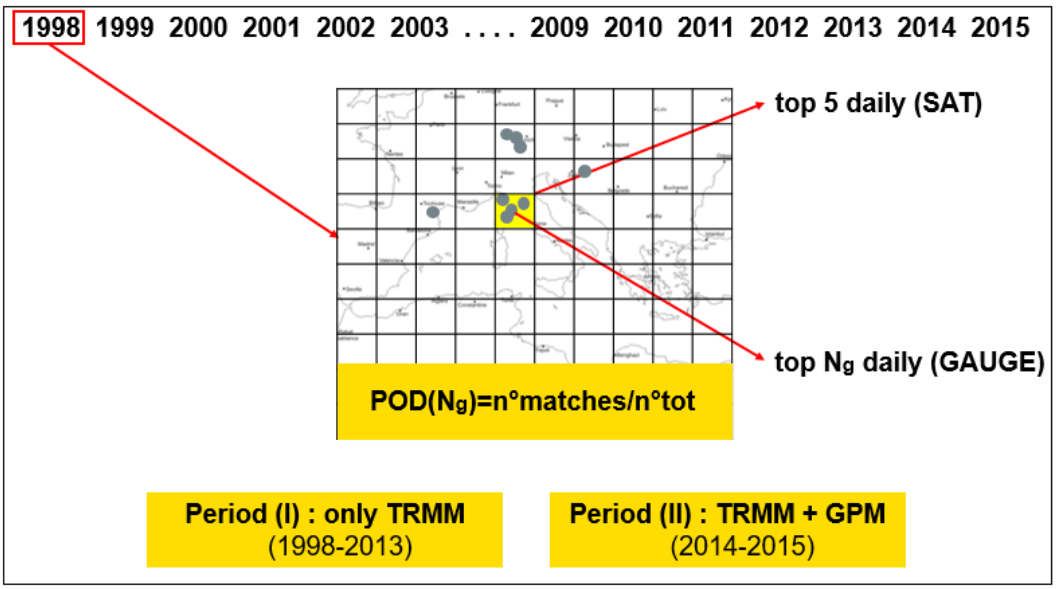


Fig. 4.4 The methodology adopted for the evaluation of the time-matching ability of the different satellite products.

procedure. On the other hand, the number N_g varies, simultaneously for all cells, between 5 and 15 in order to test the sensitivity of the results.

The agreement between the two sets of dates is then evaluated by checking how many satellite events among N_{tot} find a counterpart in the gauge-recorded significant events. In order to check the quality of the timing assessment, the probability of detection [212], defined as the fraction of significant dates correctly matched over the total, is considered:

$$POD_{N_g} = N_{match}/N_{tot} \tag{4.1}$$

where N_{match} is the number of satellite significant dates matching the rain gauge significant dates, and N_{tot} is the number of satellite significant events considered. The probability of detection varies with the number N_g of gauge significant events considered. The range is 0-1 and the ideal value is 1. This value can not be reached in the real case, due to the different scales to which the compared measurements are referred. This involves many complications and would require advanced considerations on the areal distribution of the rainfall inside each considered cell. However, those advanced high spatial resolution analysis are beyond the scope of this work,

given the amplitude of the considered domain and the preliminary nature of the study.

The is evaluated for each N_g value. N_g is used as a proxy to represent the precision of the matching. The increase of N_g leads to an increasing of the probability of a match and, for a given probability, to a decreased ability of the instrument to identify the right timing. The assessment of this sensitivity is deemed useful to explore the potential of satellite data and to provide the best grounds for improvements.

4.3 Results

4.3.1 Evolution of the estimation accuracy over time

Figure 4.5a shows the relationship between the average POD_{N_g} among all cells and N_g for the different analysed products and according to the periods described in Table 4.1. Results relating to period (I) are presented as box plots, representing the variation within the time span of the average POD_{N_g} among all cells. For period (II) lines representing the average POD_{N_g} among all cells along the whole period are plotted for the different products. The central mark of the box plots is the median, the edges of the boxes are q_1 and q_2 (i.e. the 25th and 75th percentiles), the whiskers extend to the most extreme data points not considered outliers (i.e. values out of the range $[q_3 + 1.5(q_3 - q_1), q_1 - 1.5(q_3 - q_1)]$). Outliers are plotted individually as red crosses.

For period (I) the use of $TRMM$ data provides a match between "satellite significant dates" and "rain gauge significant dates" for approximately 35% of the dates. The match rate increases, as expected, with the increase in N_g, exceeding 50% when the 15 most significant events are considered.

For period (II), over the whole range of considered N_g, GPM shows a match rate which is circa 5% greater than that of $TRMM$. This is a significant result, considering that the $3IMERGHH$ product is at the preliminary stage of its development and that, similarly to the $TRMM$ example, refinements of the conversion algorithms and improvement in the satellite constellation equipment, will increase quality as the GPM mission moves forward [96, 126, 118]. Indeed, while the shape of the box

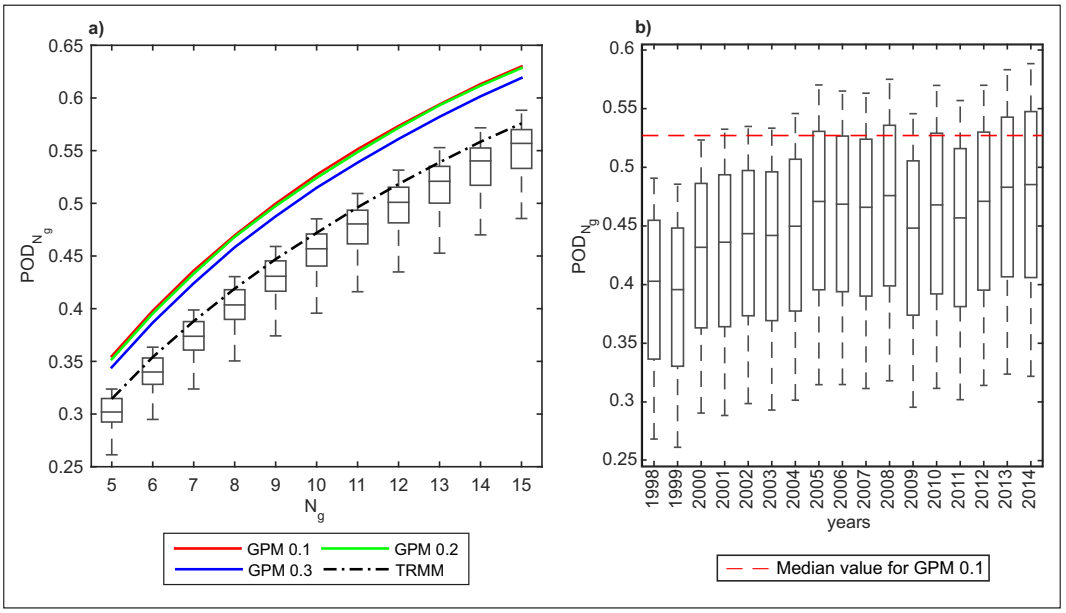


Fig. 4.5 (a) Values of POD_{N_g} for different satellite products varying N_g. The box plots refer to the distribution of the average POD_{N_g} among all the considered cells for $TRMM$, along period (I). (b) Box plots representing the distribution of POD_{N_g} with N_g, for $TRMM$ along the 1998-2014 period. The red dotted line represent the median value of the distribution of POD_{N_g} on N_g for GPM over the 2014-2015 period.

plots of $TRMM$ annual performances in figure 4.5a seems to suggest little variation over time, a substantial improvement can be detected in $TRMM$ performance over the years. The line referring to $TRMM$ performance in period (II) is always above the 75th percentile of the box plots, which suggests an improvement over time. The increasing linear trend in the mean POD_{N_g} value is confirmed by a Student-T test with a 5% significance level. This leads to an average of 10% increase in the ability to detect the timing of severe rainfall from the beginning to the end of the $TRMM$ cycle. This improvement is singularly due to the evolution of the $TRMM$ constellation, since the impact of the refinements in the algorithm is null given that only consider the latest version (i.e. version 7) is considered. The variation of the distribution of the average POD_{N_g} value along the 1998-2014 period is illustrated in figure 4.5. Visible improvements can be noticed corresponding to major additions and upgrades in the $TRMM$ satellite constellation (e.g., the start of Sounding Units in 2000, etc. [96]). The comparison with the median value of POD_{N_g} for GPM in period (II) highlights the potential of the instrument. If further improvements were

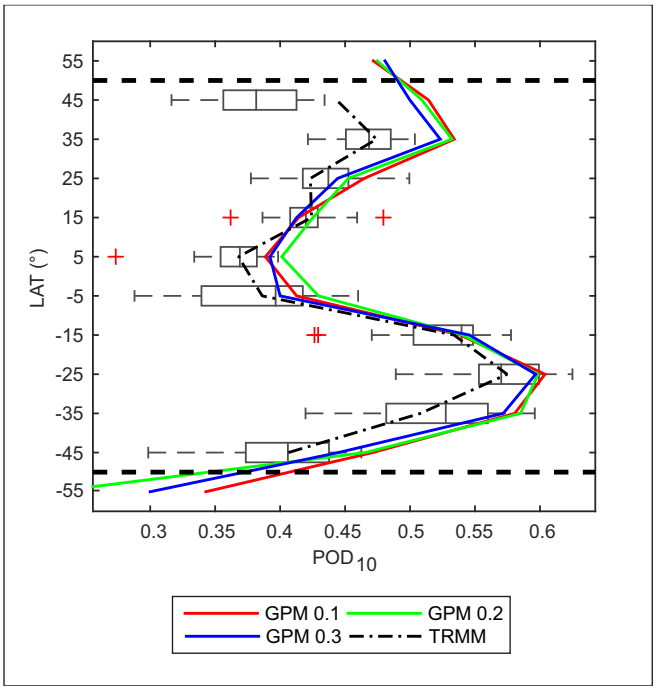


Fig. 4.6 Variability along latitude of POD_{10}, considering 10° latitude bands. The box plots refer to the distribution of the average POD_{10} among the cells of each band for $TRMM$, along period (I). The lines refer to $TRMM$ and GPM in each latitude interval for period (II).

to be obtained during the GPM mission, the probability of detection for N_g=5 would draw near to a 50% accuracy.

Moreover, the comparison between the different analyzed GPM resolutions (figure 4.5a) shows that the degradation in the accuracy due to the increase in spatial resolution is negligible. This allows one to use the finest available resolution and exploit all capabilities of the instrument.

4.3.2 Estimation accuracy on a spatial scale

The second stage of the analysis is focused on the variability of the timing accuracy on a spatial scale. In order to reduce the complexity, the value of N_g is fixed to 10. At first, the domain is divided into latitude bands of 10°. The POD_{10} is then evaluated for each interval. Box plots representing the variation within period (I) of the average POD_{10} among the cells in each latitude band are shown in figure 4.6. For period (II) the mean value of the POD_{10} for each latitude band is reported.

Both *TRMM* and *GPM* show an evident variability with latitude, with maxima in the correspondence of latitudes 35°N and 25°S and minima at the Equator and at the borders of *GPM* and *TRMM* domains. The fluctuation of the POD_{10} with latitude shows significant similarities with the distribution of the rain gauges (figure 4.3). Both results show a concurring double-peaked behavior; the areas characterized by greater rain gauge densities seem to display larger POD_{10}. This outcome can be partially attributed to a greater robustness of the results in areas with higher rain gauge densities, where POD_{10} is less sensitive to outliers. A higher density coincides with a larger probability of having multiple rain gauges in the same cell. This provides longer gauge series for comparisons, and facilitates the identification of the significant events. Moreover, the areas with higher rain gauge densities coincide with the countries that provide more frequent updates of rainfall data (e.g., USA, Australia). Low density areas (e.g., at the Equator) are more sensitive to outliers, and the lack of complete series complicates the identification of the top events by adding noise to the system. The Equatorial band in figure 4.6 seems to present a lesser degradation, even if minimal, with the coarser *GPM* resolution compared to the original one. A coarse spatial resolution can probably allow for a more robust verification by attenuating the negative effects of the smaller gauge density (i.e. it allows for multiple rain gauges in the same cell, even in the presence of low densities). On the other hand a finer spatial resolution involves a lower underestimation of intense rates when averaging on the cell area, making satellite measurements more comparable with rain gauge ones. The main issues identified could therefore be also partially attributed to the dataset used in the verification phase, but due to the complexity of the issue related to the different resolution of the measurements, further investigations are thus recommended as soon as a denser dataset will be available.

The current status does not allow for further analysis that are undoubtedly needed, also considering that climatological aspects could further affect the spatial distribution of the results (e.g., the dependence with latitude could also be due to the different types of dominant events, etc.).

The spatial distribution at the global scale of the POD_{10} for *TRMM* on period (I) and for *TRMM* and *GPM* on period (II) is presented in figure 4.7. The resolution, at 5°x5°, is coarse because of the difficulty to distinguish isolated cells at the original

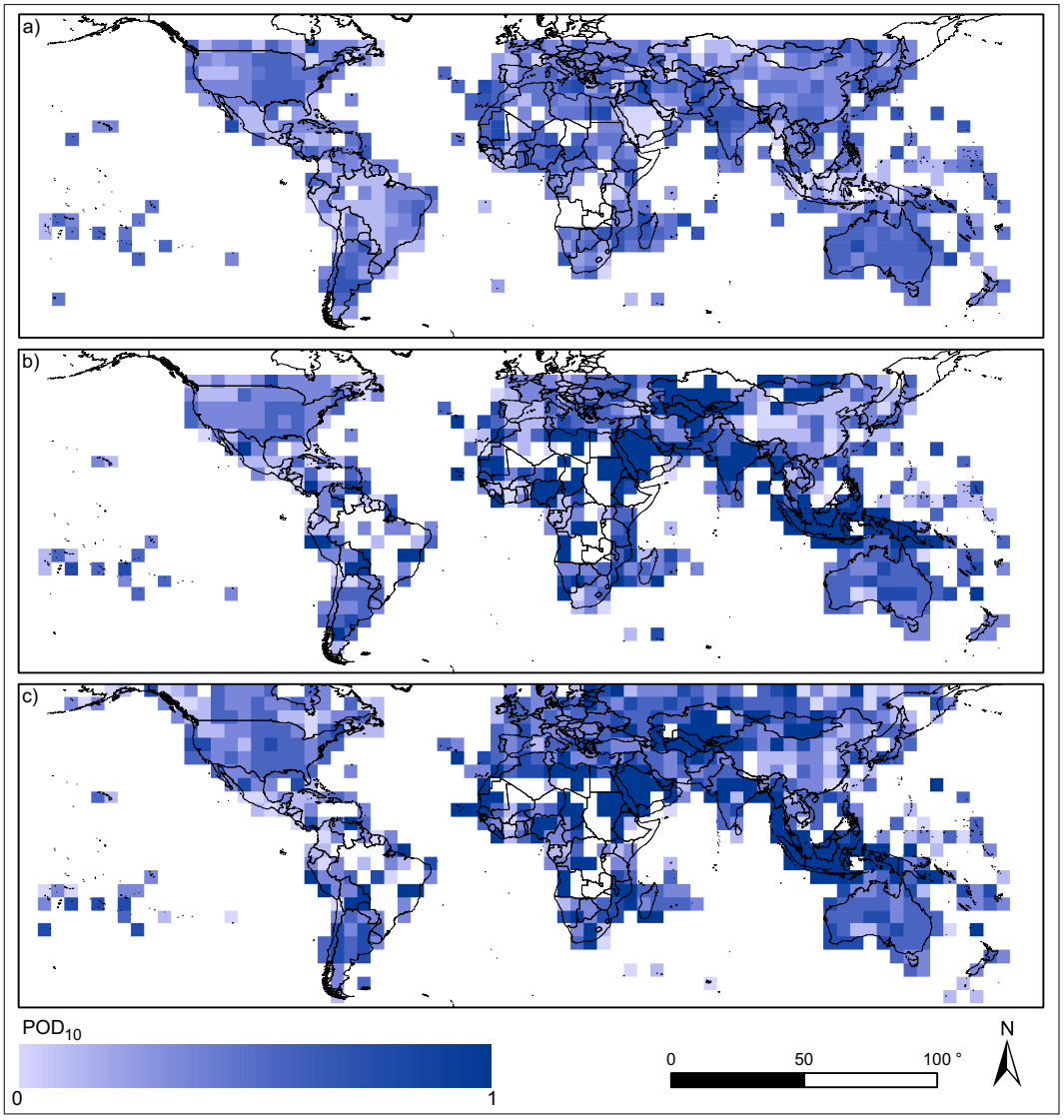


Fig. 4.7 Average POD_{10} on a 5°x5° gridded domain. (a) *TRMM* over period (I) and (b) over period(II). (c) *GPM* at the original resolution over period (II).

resolution. In period (I) (panel (a)), the results are consistent with what emerges from figure 4.6, presenting a cluster of larger POD_{10} values around Oceania. POD_{10} values are of the order of 0.45 even in areas characterized by deep convection, in which *TRMM* products are known to be unreliable (e.g., South American Andes [159]). The cluster of small POD_{10} values located in Brazil is probably due to the worse local quality of the rain gauge data. In fact, from figure 4.3, one notices that the data availability for the region is limited to a single year. Once again, Table 4.2 confirms that *TRMM* results show large improvements when considering solely period (II).

Panel (b) and panel (c) of figure 4.7 represent respectively the performance of *TRMM* and *GPM* over period (II). For ease of comparison figure 4.8a presents a map of the differences between the POD_{10} with *GPM* and the POD_{10} with *TRMM*. The results of the two instruments seem generally consistent, with large areas with coincident POD_{10} values. Some areas in which *TRMM* seems to exceed *GPM* are clearly identifiable (e.g, Bolivia, Indian area). On the other hand, noticeable improvements can be recognized in areas where global satellite products are known to be poorly reliable (e.g, Europe and Mediterranean Area). As the comparison is difficult due to the complex spatial variability, results are summarized for some areas of interest in Table 4.2 reporting mean and standard deviation *Std* of the distribution of the POD_{10}. This confirms remarkable improvements for *GPM* at the global scale, allowing for an increase in the POD_{10} up to 9% for the Mediterranean basin and Africa. The general improvement are also summarized at the continental scale in figure 4.8b.

The value of N_g was set beforehand at a fixed value; this assumption has proven acceptable because, in the context of the analysis of the spatial variation, N_g turns out to be a less significant parameter. Even though different values of N_g produce different absolute values of POD_{N_g}, the global trend in the spatial pattern is preserved. For completeness, the analogues of figures 4.5b and 4.7 for N_g=5 are reported as figure 4.9 and figure 4.10. Despite the differences in the magnitude, the spatial distributions are similar.

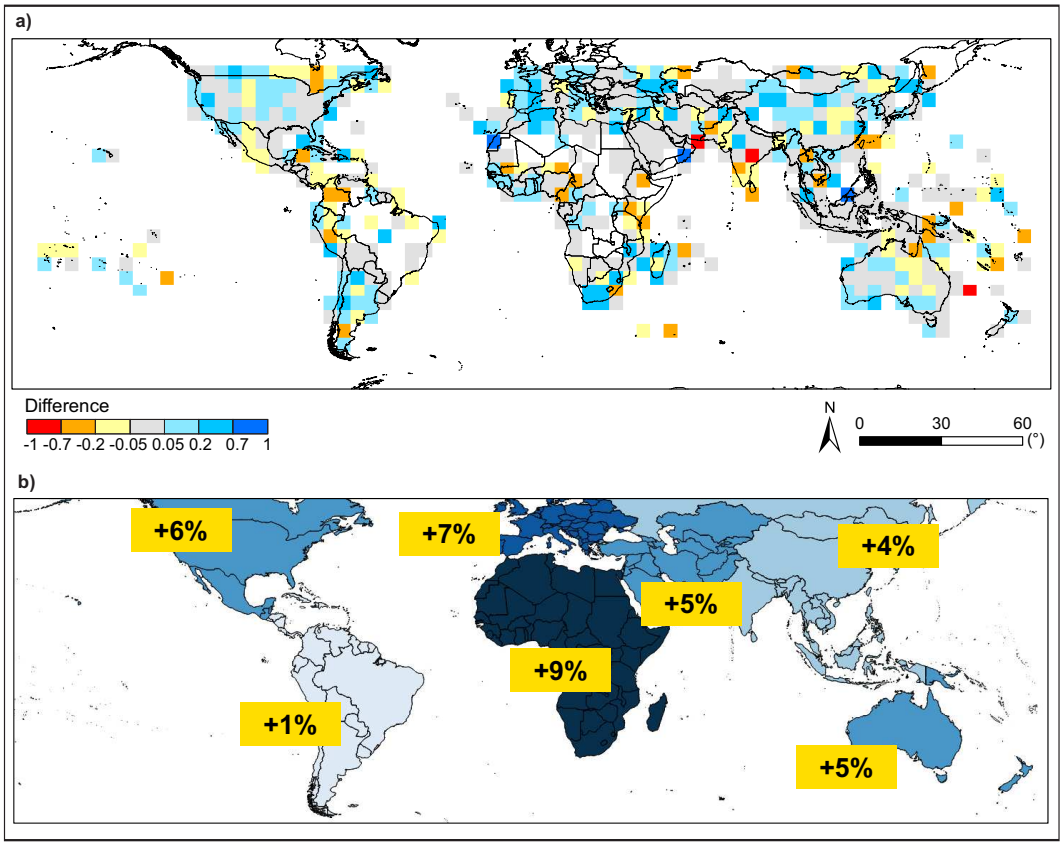


Fig. 4.8 (a) Differences between the POD_{10} with *GPM* and the POD_{10} with *TRMM* on a 5°x5° gridded domain. (b) Schematic representation of the improvement in the POD_{10} from *TRMM* to *GPM* at the continental scale.

Table 4.2 Mean and standard deviation of the distribution of the POD_{10} for some areas of interest

	Period(I)		Period(II)			
	TRMM		*TRMM*		*GPM* 0.1	
	Mean	*Std*	Mean	*Std*	Mean	*Std*
Africa	0.44	0.17	0.53	0.33	0.62	0.31
America (North)	0.44	0.15	0.47	0.25	0.53	0.26
America (Center-South)	0.38	0.18	0.49	0.30	0.49	0.31
Asia (West)	0.43	0.15	0.55	0.32	0.59	0.32
Asia (Middle-East)	0.39	0.20	0.61	0.35	0.66	0.32
Europe	0.36	0.16	0.44	0.27	0.51	0.24
Oceania	0.54	0.11	0.54	0.24	0.59	0.24
Andes	0.45	0.18	0.52	0.31	0.54	0.32
Australia	0.54	0.11	0.55	0.24	0.59	0.24
Brazil	0.37	0.18	0.56	0.31	0.65	0.34
Mediterranean basin	0.37	0.18	0.43	0.29	0.53	0.27
USA	0.45	0.14	0.48	0.24	0.54	0.26

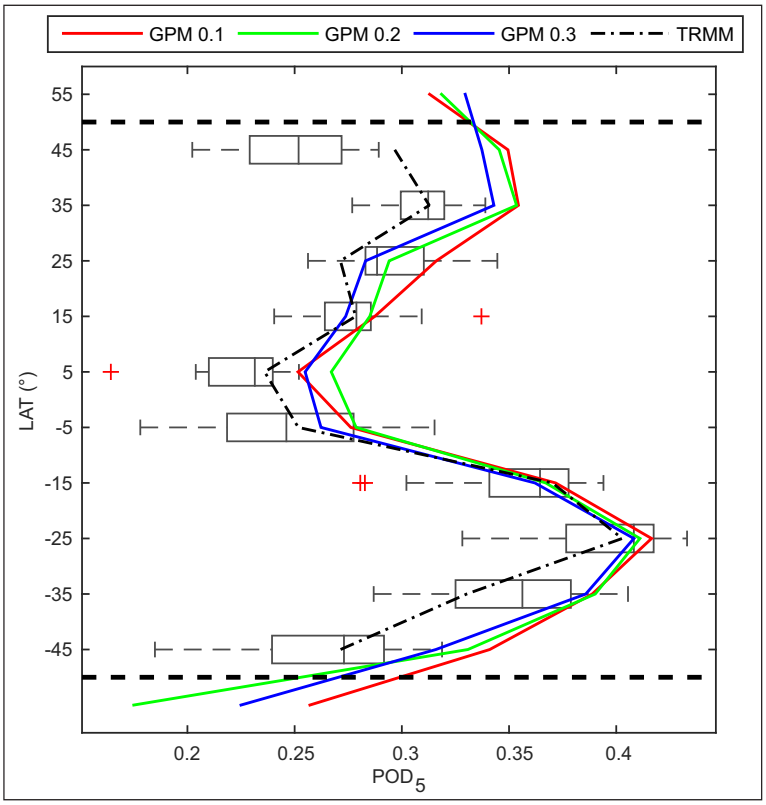


Fig. 4.9 Variability along latitude of POD_5, considering 10° latitude bands. The box plots refer to the distribution of the average POD_5 among the cells of each band for *TRMM*, along period (I). The lines refer to *TRMM* and *GPM* in each latitude interval for period (II).

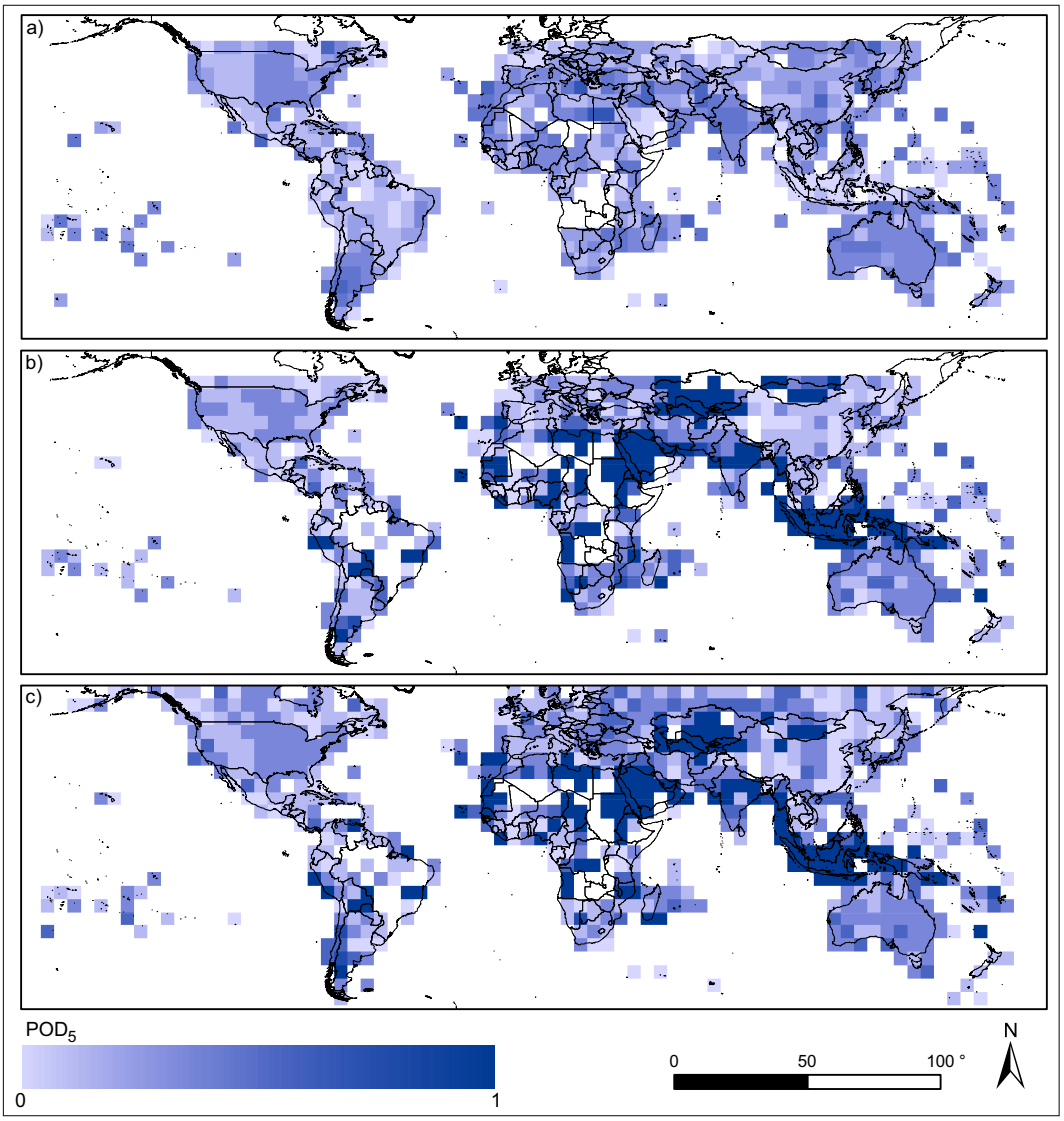


Fig. 4.10 Average POD_5 on a 5°x5° gridded domain. (a) $TRMM$ over period (I) and (b) over period(II). (c) GPM at the original resolution over period (II).

4.4 Concluding remarks

Two satellite rainfall products have been analyzed in order to assess their potential in recognizing severe precipitation events at the global and synoptic scale. These products are known to underestimate rainfall rates for deep convective systems. However, their interesting spatio-temporal resolution and their quasi-global coverage

make them useful in the definition of spatial precipitation features, with particular reference to the occurrence.

A preliminary analysis carried out in the Liguria region (North-West of Italy) confirms a marked underestimation of extreme rainfall depths, but highlights the ability of *TRMM* to identify the dates of occurrence of severe rainfall events. The analysis is expanded to the global scale, over the 1998-2015 period. The performance of *TRMM* in identifying the timing of global extreme precipitations is found consistent with that obtained in the preliminary analysis, matching nearly 35% of the dates of occurrence. The matching capability registers a 10% improvement over the *TRMM* life-cycle, due to the evolution of the satellite constellation.

The results obtained with the Global Precipitation Measurement Mission products after only one year of operation, seem promising. The finer spatio-temporal resolution and the increased measurement range (to include light-intensity precipitation and falling snow [138]) allow *GPM* results to be quite more accurate than the *TRMM* ones. At the global scale *GPM* shows a greater ability in matching the day of occurrence of intense rainfall, with a probability of detection in the order of 0.6-0.7. These results hold also for the areas in which *TRMM* faces issues due to the flattening of the rainfall peaks (e.g., the Mediterranean region).

The great abundance of data considered here leads to significant outcomes concerning the limits of *TRMM* products in the analysis of intense rainfall events and the high potential that *GPM* shows after only one year of operation. Nevertheless, further in-depth analysis and testing are needed before using the results in areas with low gauging density.

Satellite data can play an important role in the analysis of the spatio-temporal connection between severe rainfall systems at the global scale even when dealing with extreme and extraordinary rainfall systems. The timing identification approach can provide new perspectives for the use of satellite products for the analysis of rainfall pattern, considering that the lack of accuracy in the estimation of the rainfall amounts at the daily scale still limits their direct systematic use.

Going back to the regional scale, satellites can be useful in the analysis of very localized and severe rainfall systems from different point of views. Timing information from satellite can provide systematic informations on the framework in which extreme precipitations have been generated, allowing for the development of

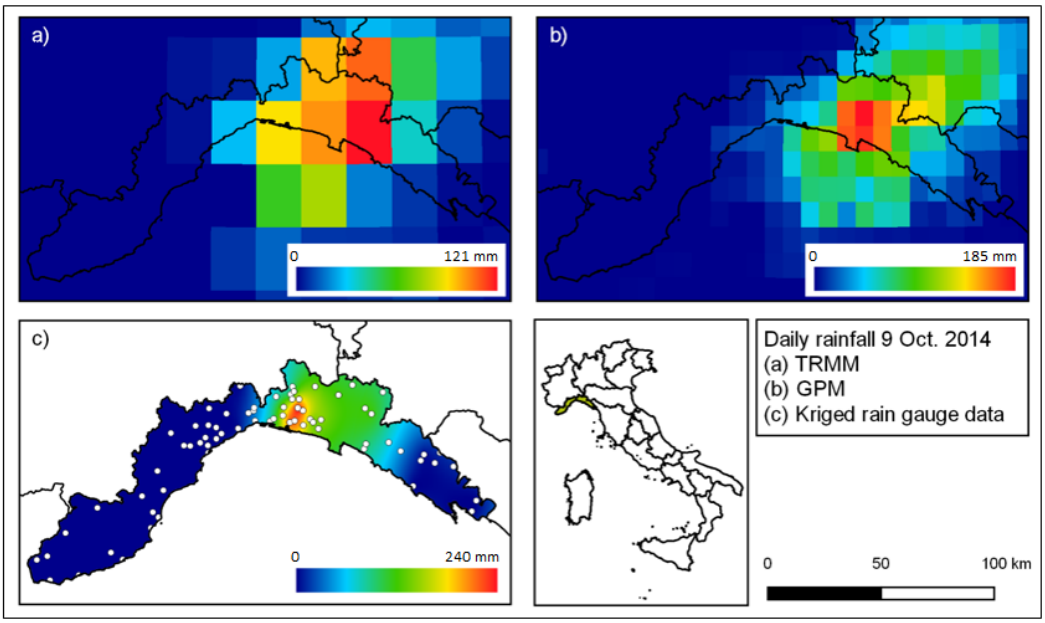


Fig. 4.11 Daily rainfall over Liguria region on the 9th October 2014 with (a) *TRMM*, (b) *GPM* 0.1 and (c) kriged rain gauges data.

prediction methods based on the synoptic/mesoscale information. Moreover satellite products can help representing the spatial rainfall fields in ungauged areas. The *TRMM* and *GPM* maps of the super-extreme rainfall event occurred in Genova on 9 October 2014 reported in figure 4.11, shows that *TRMM* helps to coarsely identify the area where the event occurred, and that *GPM* is able to provide a high resolution representation of the spatial rainfall. Once longer series will be available, this kind of information could be useful, particularly if merged with the rain gauge data, for defining the border of the storm systems and help spatializing the storm-hazard assessment in ungauged areas.

Chapter 5

New insights in rainfall hazard assessment in Italy

In Italy most of the hydrological analyses are based on the rain gauge data. However, a systematic and updated national database of extreme rainfalls is still missing. For this reason, updated rainstorm hazard assessments are actually only available at the regional scale. Various studies present different metodologies and are sometimes based on very different data densities and record lengths (e.g., [50]). For dealing with this complex background information, in this study a new dataset has been developed. The first sections of this chapter are dedicated to the description of this dataset, built using several different sources of data [122]. Exploratory statistical analyses for providing information on the climatology of extreme rainfall at the national scale are then carried out, considering the spatial variability and stationarity in time of the larger quantiles.

The last part of the chapter is focused on the analysis of the highest empirical quantiles of the annual maximum rainfall occurred in Italy in the last century. Many of these events have been studied as individual extraordinary events (e.g., [161], [73]), due to the large recorded intensities and/or to their severe consequences, but they have been seldom expressly addressed as a definite population [169]. The analysis aims at assessing the conditions for possible consideration of a super-extreme component in the probability distribution. This separate component would present a statistical behavior significantly different from that of the rest of the extremes, and should include the "Black Swans", introduced in section 1.6. Subsequently, the spatial

distribution of the data overcoming the 0.999 quantile of the empirical probability distribution is analysed, taking into account the density of the rain gauge network in time and space. Finally, morphological features are investigated at the country scale, for possible explanation of the spatial distribution of the highest quantiles.

5.1 The Italian Rainfall Extremes Dataset (*RED*)

In view of the assembling of the first comprehensive dataset of extreme rainfall of short duration in Italy (see appendix B for details) several major sources of data have been analysed. The resulting dataset, referred to as *RED* (i.e., Rainfall Extremes Database), includes data from more than 4500 stations across the country, spanning the period between 1916 and 2014, and refers to annual maximum rainfall recorded in 1 to 24 consecutive hours (exact durations available are 1-3-6-12 and 24 hours). Observations dating before 1916 have been discarded as considered not significant and too much unevenly distributed. The number of data available per year is reported in figure 5.1 compared with that of the *CUBIST* database [49] compiled in 2007 in the framework of a national project funded by the Italian Ministry of Education and Research (*MIUR*). The figure shows the minimum number of data per year among the 5 available durations. For more details, refer to appendix C.1.

As every station is related to an unique value of annual maxima for a given duration, the presence of a measurement implies the presence of a station. The number of available stations increases with time, and drastically grows after the dismissal of the *SIMN* and the development of the local agencies (see section 1.4). The decrease after 2010 can be attributed to the fact that not all the regions have published the data for the most recent years. The smaller size of the *RED* compared the *CUBIST* database in some years can be attributed mainly to two reasons:

- The presence, before 1945, in the *CUBIST* database of data from territories lost by Italy after World War II (e.g. , Istria) or from neighboring countries, not included in the *RED*;
- The fact that regional agencies could have decided for different reasons not to include data or station from the *SIMN* dataset in their database. Part of

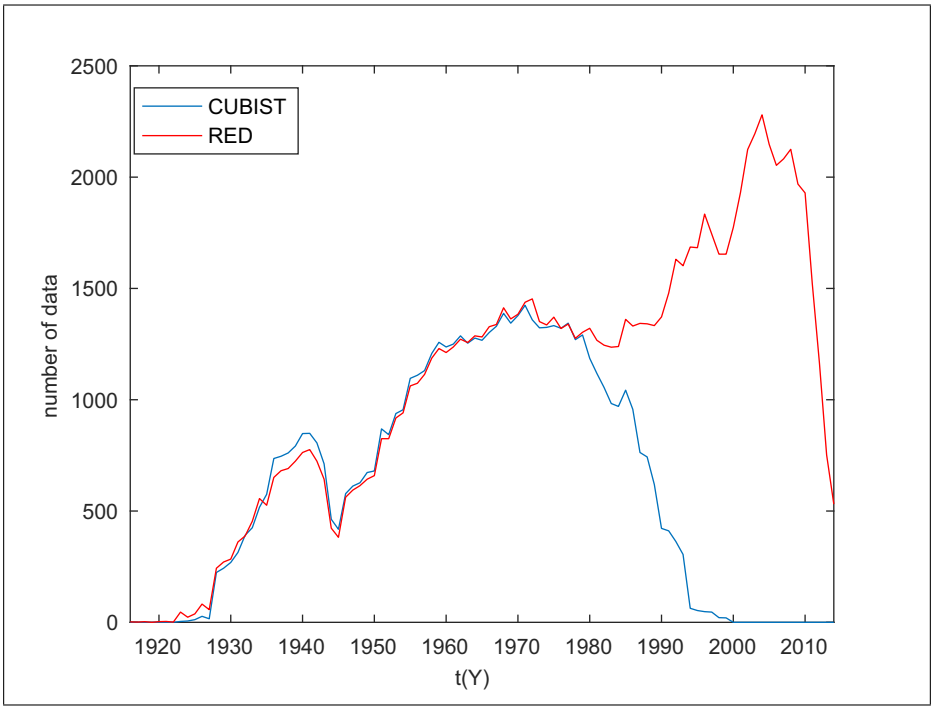


Fig. 5.1 Data availability per year in the *RED* and *CUBIST* databases (the smallest value across the 5 considered duration is reported per each year)

these data could therefore be lost not considering the *CUBIST* database for that regions.

Considering the limited relevance of the information loss, it was decided that further corrections to the *RED* are not required for the objectives of this study.

All the time series are classified according to their length. Results are shown in Figure 5.2. Obviously, considering the short life of the regional operational centers, most of the new series are shorter than 20 years, but the merge with the *CUBIST* series provided also a significant amount of longer series (see, e.g., the 60-70 class).

The spatial distribution of the stations is shown in figure 5.3. The color scale refers to the number of the available data per each series. The minimum number across the 5 duration is considered. One can clearly distinguish that, even if all the national territory is represented, the density of the station is varying widely across the nation. In particular, some areas are characterized by a prevalence of short and very short series (e.g., Lazio, Campania). The variability in the spatio-temporal distribution of the stations will be discussed in the following.

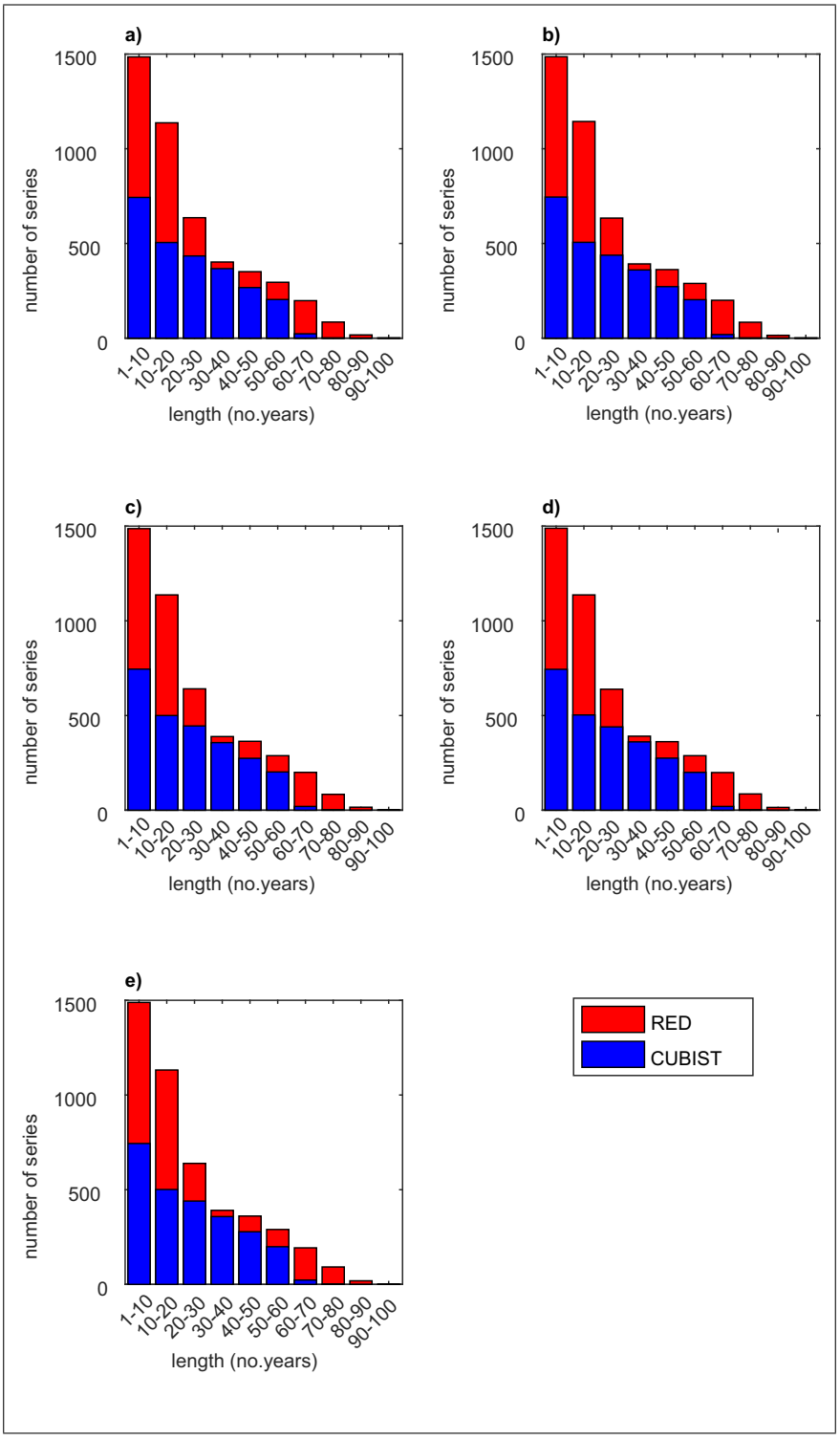


Fig. 5.2 Number of series per length class in the *RED* and *CUBIST* databases for durations from 1 (a) to 24 (e) hours

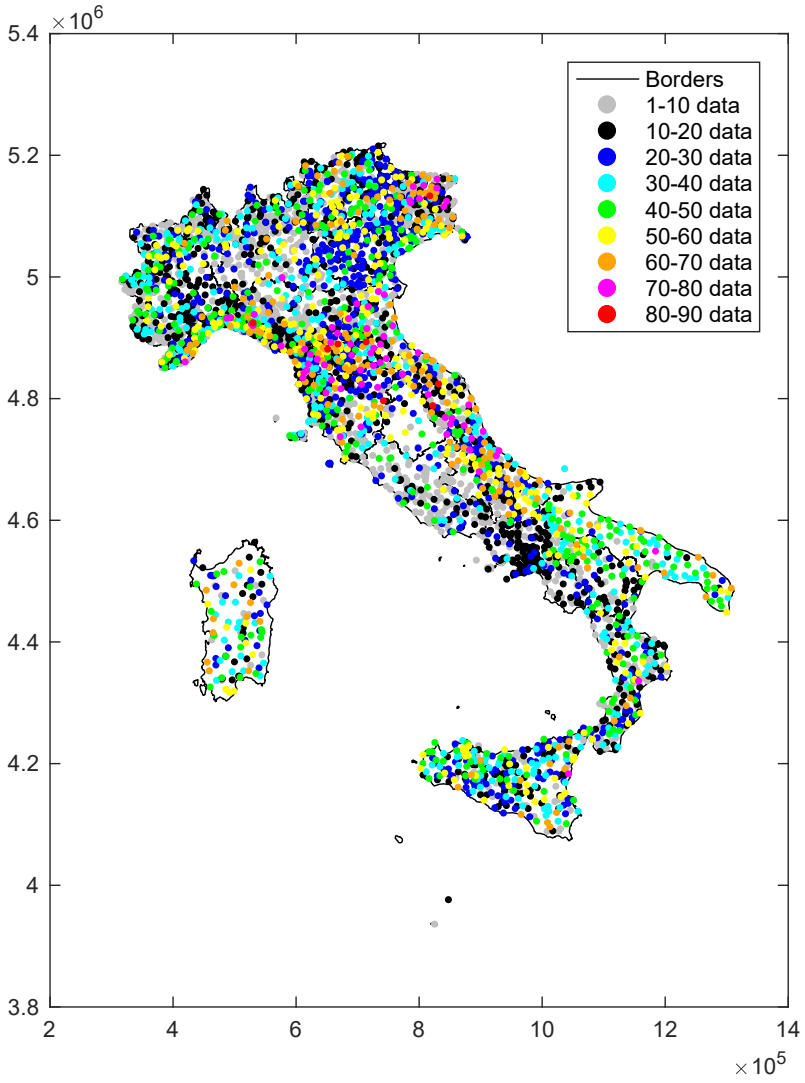


Fig. 5.3 Length of the series in the *RED* database represented in space. The color refers to the minimum length among the 5 available durations. If more stations overlap due to the resolution of the picture, the one with the longer series appears on top

5.2 Descriptive statistical analysis of rainstorms in Italy

A preliminary descriptive analysis of the characteristics of extreme rainfalls at the national scale has been carried out on the newly developed *RED* database,

considering that very little information is available at the contry scale at the time of this study. Series with a minimum length of 20 years of data have been considered in this analysis. This constraint leads to a subset of 1974 series available for the analysis, out of the original 4686. For each duration the median of the series is reported in figure 5.4. The median is used as a robust estimator of the central tendency of a series, less sensitive than the mean to the presence of outliers.

When short durations are considered, stations with larger median values seems to be evenly distributed across the country. Moving to longer durations, a marked spatial clusterizations of the series with larger median value seems to arise. The Alps and the Appennine area, as well as the eastern Adriatic Coast, and the western parts of Sicilia and Sardegna are characterized by low median values for all the considered durations. The northern shore of the Adriatic sea, the pre-Alps in Piemonte and Lombardia, the Liguria region and the northern Toscana, as well as the eastern coasts of Calabria, Sicilia and Sardegna shows the highest median values.

For each series, the sample L-moments have then been computed according to equation 2.5 [89]. As previously mentioned, L-moments are measures of the location, scale and shape of the empirical distribution of the records. The mean L-moments ratios among the different durations are considered. More in the detail the analysis refers to: $L-CV$, $L-CA$ and $L-KUR$, introduced in section 2.4.2, as they can give information respectively on the dispersion, skewness and "peakedness" of the empirical distribution of the values. The spatial distribution of the first 4 order of L-moments of the series is represented in appendix C.2.

L-moments ratios are mapped in figure 5.5. Panel (a) shows that the coastal areas and the islands are generally characterized by a higher variability in the annual maxima series, presenting larger $L-CV$ values. The northern part of the peninsula, even if characterized by large median values, shows lower $L-CV$, which is typical of areas with large average rainfall values. It is harder to identify a precise spatial pattern in the distribution of the skewness and kurtosis values. Coastal and island areas seems to generally show larger skewness values, confirming the influence of the Mediterranean sea on the climate of these areas.

The coefficient of variation (CV) and the coefficient of skewness are evaluated for each station with more than 20 years of data. For each station the durations with larger CV and skewness are selected. Results are reported in figure 5.6. Shorter

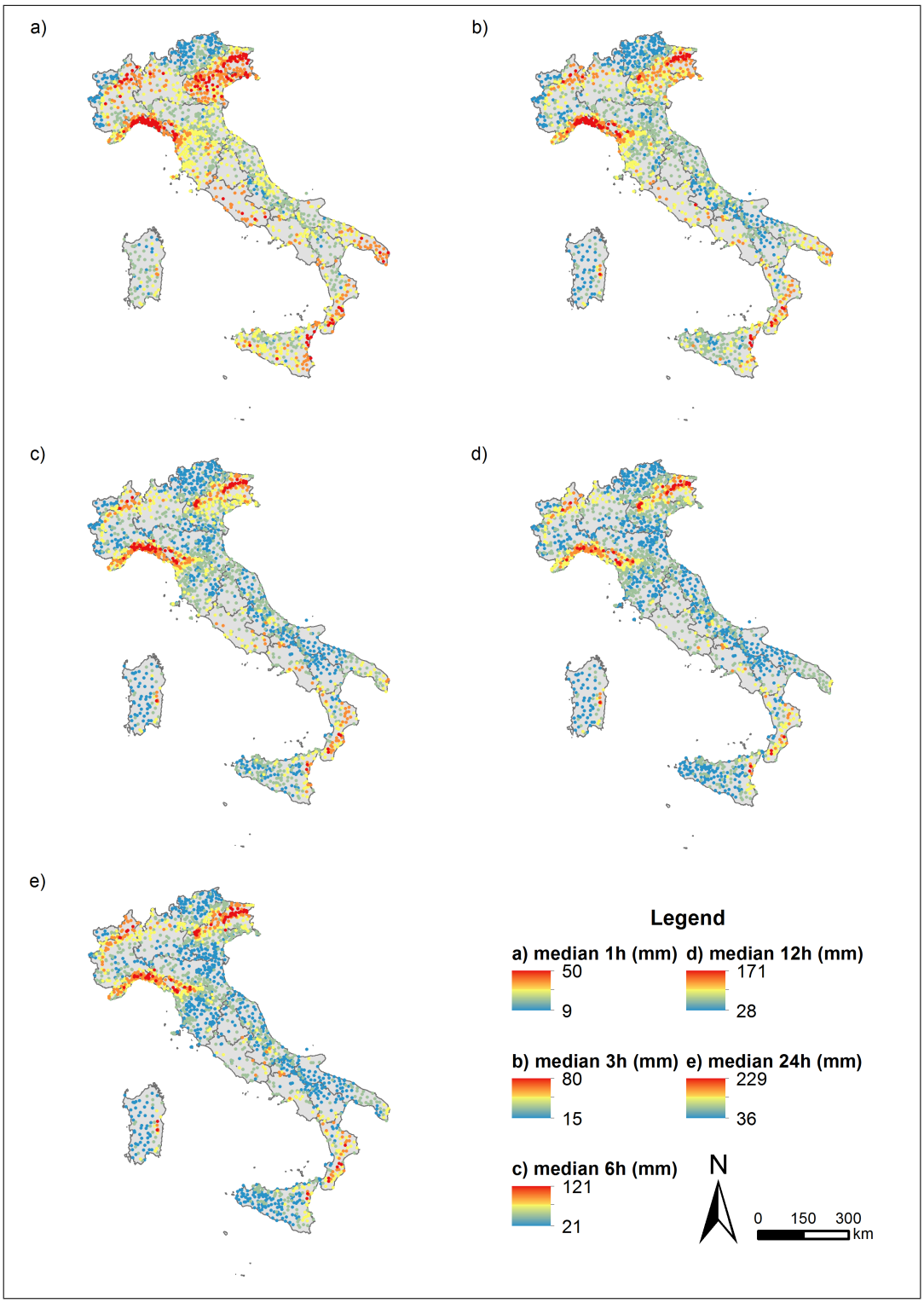


Fig. 5.4 Median values of the *RED* series with more than 20 data for the durations from 1 (a) to 24 (e) hours.

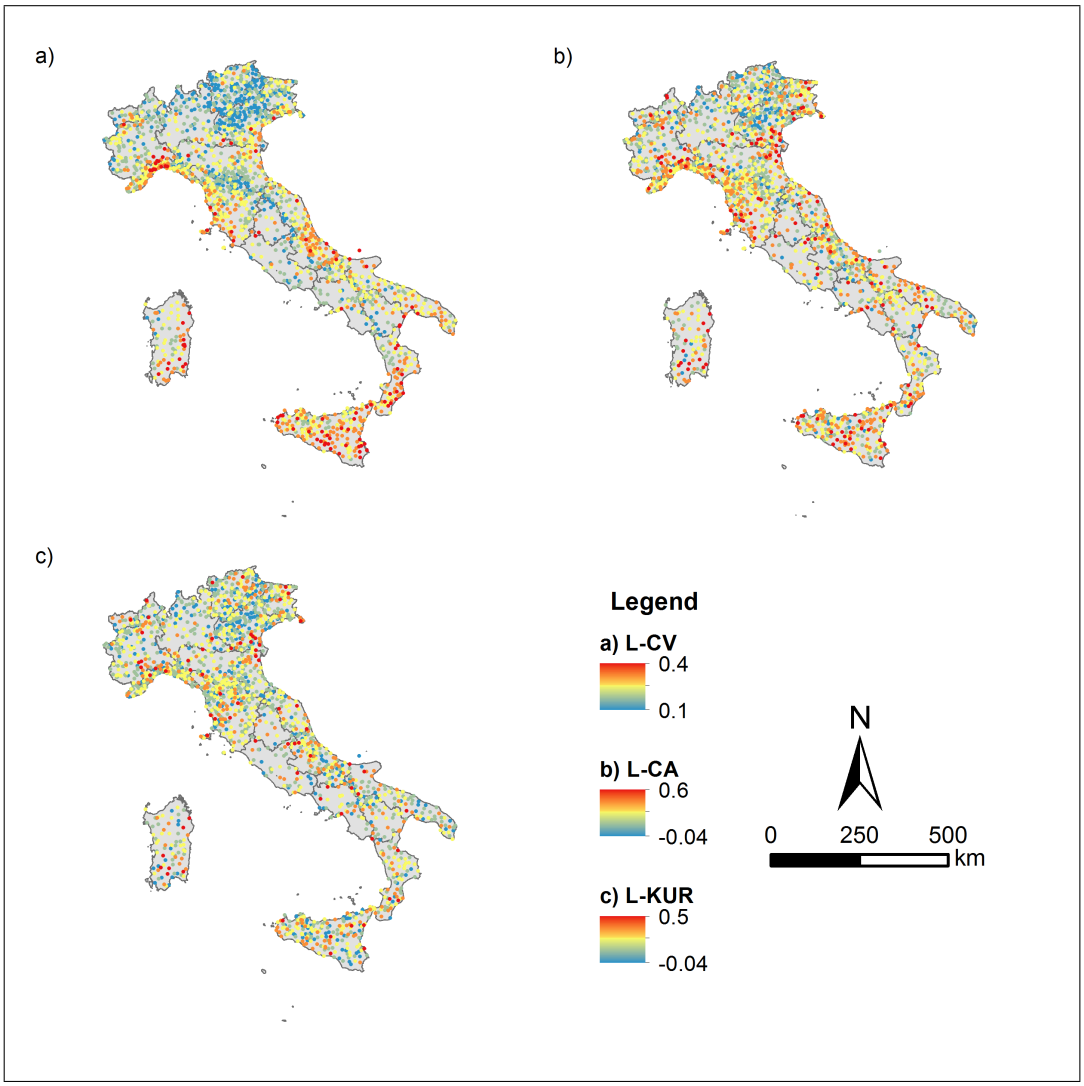


Fig. 5.5 Average statistics for the five durations considered: (a) $L-CV$, (b) $L-CA$ and (c) $L-KUR$. Series with more than 20 data are considered.

durations are generally characterized by larger variability. The skewness is quite homogeneous among the different durations.

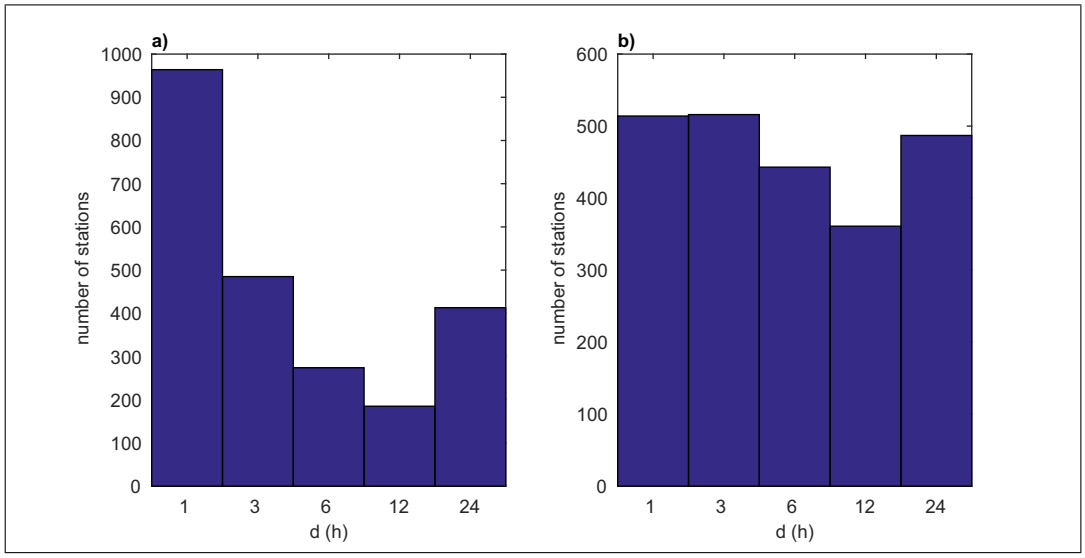


Fig. 5.6 Number of station showing the highest (a) *CV* and (b) coefficient of skewness for a certain duration. Stations with more than 20 records are considered.

To conclude this introductory section, the a and n coefficients of the mean *IDF* curves in the form of the equation 2.10 have been computed for the series with more than 20 years of data. The spatial distribution is shown in figure 5.7.

To take into account the spatial variability of the mean of the extremes, related to the different extreme rainfall regime at the national scale, all the analyses have also been performed on normalized data, obtained dividing each value by the median value of the series which it belong to. The normalized rainfall h_d^* is then defined as:

$$h^*_{Y,i} = h_{Y,i}/median(h_i) \tag{5.1}$$

with $h_{Y,i}$ rainfall depth related to the Y-th year of the i-th series and $median(h_i)$ as the median value of the i-th series.

In the following the upper tail of the distribution of the annual maxima at the national scale is analysed by pooling together all the data for each duration. 10 series composed of about 80000 values are then obtained, 5 with absolute and 5 with normalized values. All the analyses have been carried out on both the absolute and the normalized rainfall: as the analysis does not entail the selection of a probability

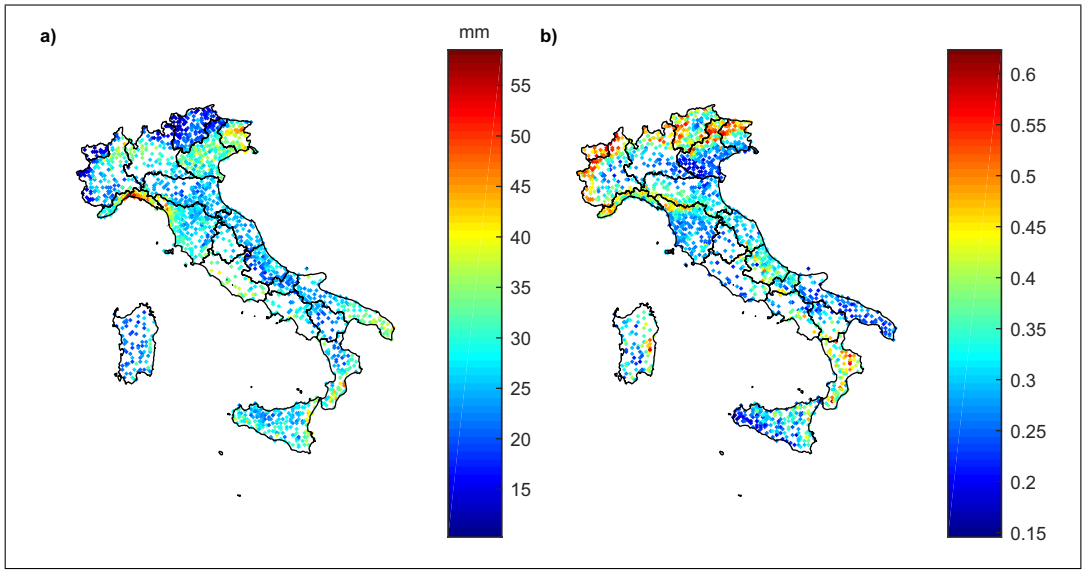


Fig. 5.7 (a) a and (b) n coefficients of the IDF curves for the series of the RED with more than 20 data.

distribution for predicting quantiles all over Italy, further evident sources of non-uniformity or dishomogeneity have not been removed at this stage. In the detail, the analyses referred to the absolute values are performed in order to give an idea of the order of magnitude of the considered events and to show the influence of the local climatology on the identification of the "extraordinary" component, explored later in this chapter.

5.3 Stationarity assessment of the largest rainstorms

As a second-level exploratory statistical analysis, the possible existence of trends in the annual maximum rainfall has been explored. This is aimed at assessing whether the higher percentiles of the empirical distribution of the pooled up data are to be considered as stationary in the about 100 years of observations.

To take into account the different length of the considered series, the nonparametric Mann-Kendall test has been adopted [131, 104, 79]. The Mann-Kendall test is a widely used non-parametric test for evaluating the presence of monotonic trends in time series data, adopted in a number of recent studies testing for assessing changes in rainfall extremes (e.g., [5, 211]). The existence of positive autocorrelation in the

data increases the probability of detecting trends when actually they do not exist, and vice versa. Although this is a well-known fact, few studies have addressed this issue, and autocorrelation in the data is often ignored. To assess the impact of autocorrelation on the data under analysis both the traditional Mann-Kendall test and a modified version conceived in [84] for autocorrelated data are considered.

The tests have been applied to the series obtained by considering annually the quantile related to a certain cumulative probability P. Growing probability values from 0.5 (i.e., the median) to 0.999 (the extreme tail of the distribution) are explored. The procedure, involving in each year the definition of the empirical quantiles, requires that the number of data per year is sufficient for their correct identification (e.g., for identifying the 0.999 quantile, at least 1000 data have been required). The number of data per year, reported in figure 5.1 refers to the whole dataset. Figure 5.8 shows the number of available stations (coinciding with the number of data) per year considering different classes of record length per station. A 20-year minimum length of the series has been confirmed for the robustness of the outcomes. The selected threshold grants a quite stable number of station in the 1958-2007 period, in which, therefore, at least 1000 stations per year are available.

The quantiles are then evaluated for each year in the considered period. They have been plotted in figures 5.9 and 5.10, for the absolute rainfall depth and the normalized ones, respectively. The Mann-Kendall tests has been then applied considering the aforementioned quantiles and two sided significance levels of 5% and 10%. The p-values resulting from the test are reported for (a) the absolute rainfall and (b) the normalized rainfall in table 5.1 and 5.2 for the traditional and modiefied test respectively. If the p-value is less than the significance level, the null hypothesis (absence of trend) is rejected. Larger p-values implies a more marked absence of trend.

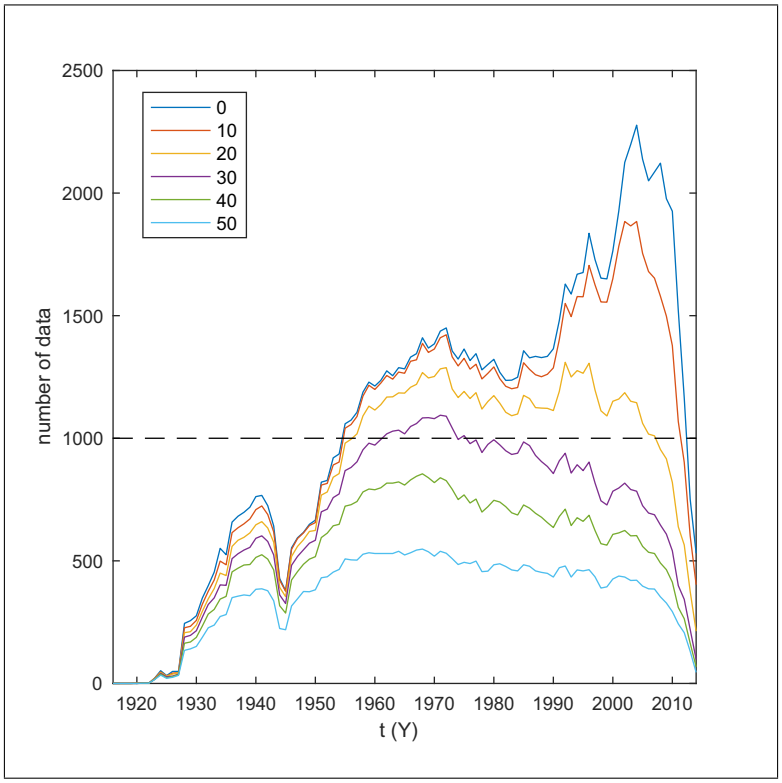


Fig. 5.8 Number of available data per year considering different minimum threshold for the length of the station records. Zero refers to no-threshold.

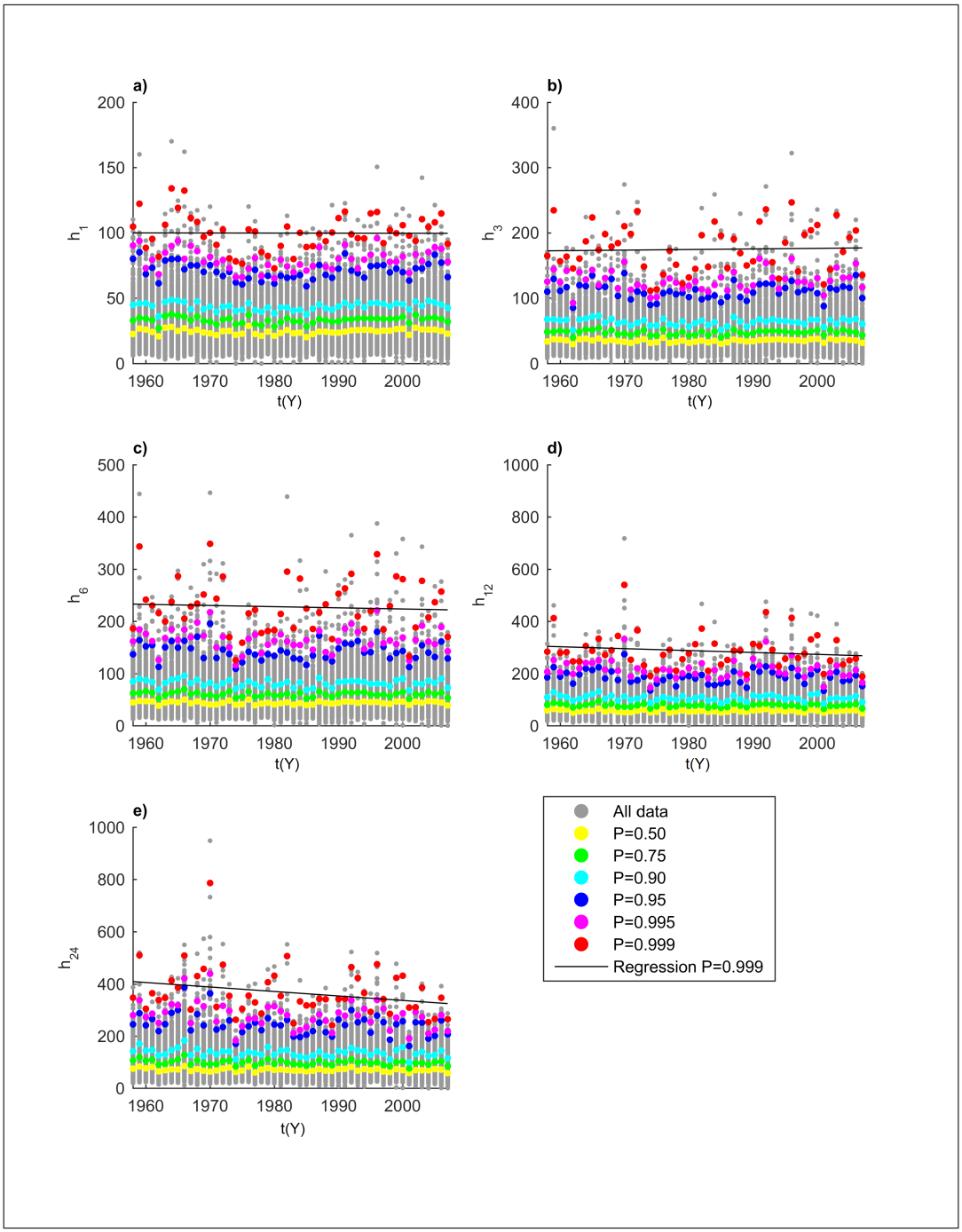


Fig. 5.9 Annual quantiles considered in the trend analysis of the annual maximum depths for durations from 1 (a) to 24 (e) hours. The regression lines related to the quantiles with 0.999 cumulative empirical probability is also reported.

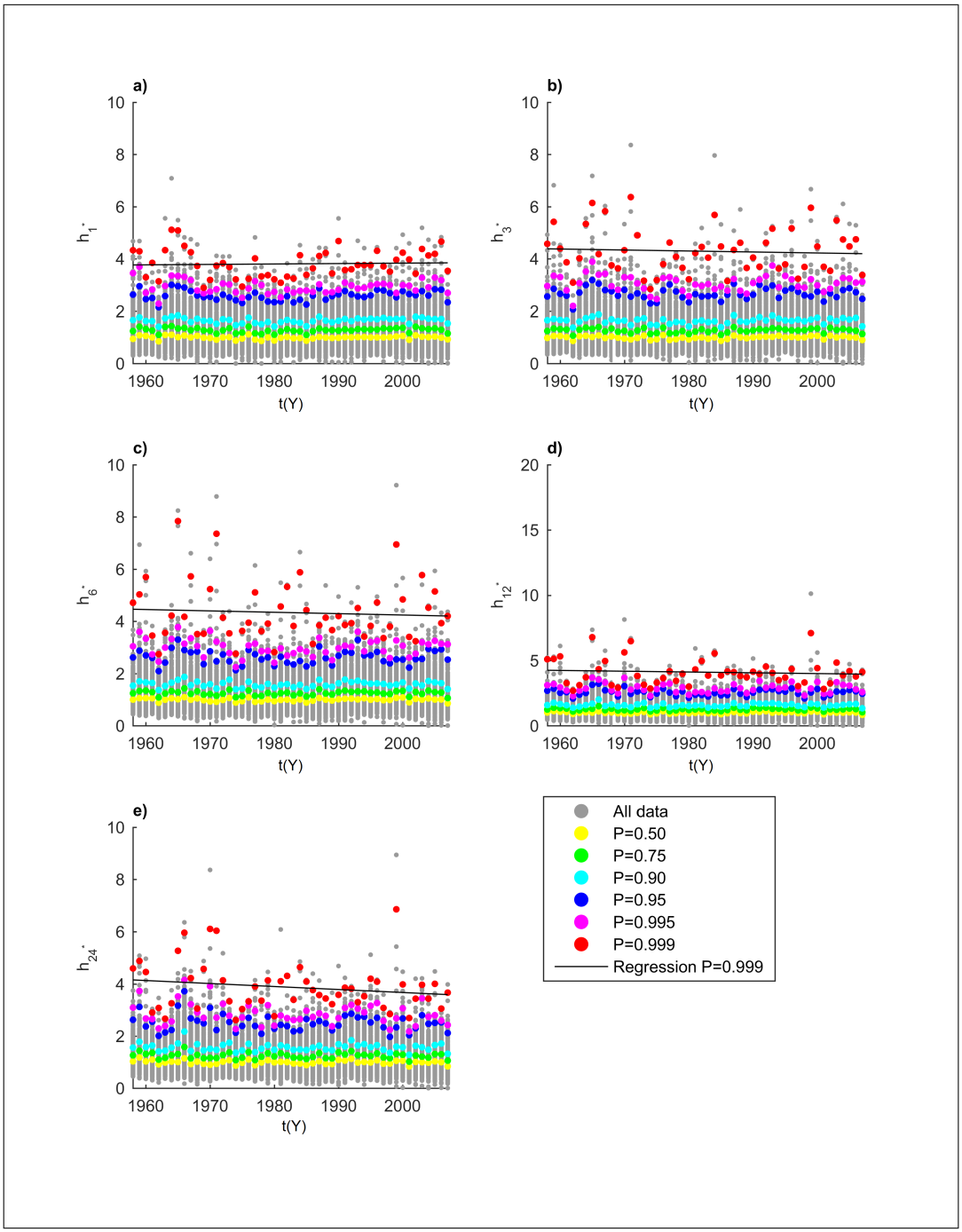


Fig. 5.10 Annual quantiles considered in the trend analysis of the normalized rainfall for durations from 1 (a) to 24 (e) hours. The regression lines related to the quantiles with 0.999 cumulative empirical probability is also reported.

Table 5.1 P-values of the Mann-Kendall test (a) for the absolute rainfall depths and (b) for the normalized rainfall. A $^{-}$ symbol indicates a downward trend. Bold values refer to a significant trend under a 10% significance level, values in square brackets refers to a trend significant under a 5% significance level.

(a)

P	1 h	3 h	6 h	12 h	24 h
0.5	0.18	0.28	0.38^{-}	0.30^{-}	$\mathbf{0.07}^{-}$
0.75	0.28	0.40	0.28^{-}	0.22^{-}	$\mathbf{0.06}^{-}$
0.9	0.38^{-}	0.34^{-}	0.24^{-}	0.23^{-}	$\mathbf{0.06}^{-}$
0.99	0.30^{-}	0.40^{-}	0.25^{-}	0.27^{-}	$\mathbf{0.09}^{-}$
0.995	$\mathbf{0.07}^{-}$	0.35	0.33^{-}	0.13^{-}	$\mathbf{[0.04]}^{-}$
0.999	0.32	0.39	0.39^{-}	0.28^{-}	$\mathbf{0.09}^{-}$

(b)

P	1 h	3 h	6 h	12 h	24 h
0.5	0.38	0.37	0.40	0.39^{-}	0.28^{-}
0.75	0.37	0.38	0.40	0.40	0.36^{-}
0.9	0.40	0.30	0.40	0.40^{-}	0.39^{-}
0.99	0.39^{-}	0.39^{-}	0.40^{-}	0.38^{-}	0.29^{-}
0.995	0.28^{-}	0.37^{-}	0.39^{-}	0.38^{-}	0.26^{-}
0.999	0.32	0.40^{-}	0.40	0.40^{-}	0.20^{-}

Table 5.2 P-values of the Mann-Kendall test for autocorrelated data [84] (a) for the absolute rainfall depths and (b) for the normalized rainfall. A $^-$ symbol indicates a downward trend. Bold values refer to a significant trend under a 10% significance level, values in square brackets refers to a trend significant under a 5% significance level.

(a)

P	1 h	3 h	6 h	12 h	24 h
0.5	0.16	0.35	0.77^-	0.44^-	**0.06^-**
0.75	0.48	0.90	0.40^-	0.28^-	**[0.05]**$^-$
0.9	0.37^-	0.57^-	0.32^-	0.30^-	**[0.05]**$^-$
0.99	0.93^-	0.96^-	0.22^-	0.26^-	**[0.03]**$^-$
0.995	0.81	0.62	0.52^-	0.14^-	**[0.03]**$^-$
0.999	0.81	0.80	0.85^-	0.40^-	**0.08^-**

(b)

P	1 h	3 h	6 h	12 h	24 h
0.5	0.75	0.68	0.89	0.85^-	0.39^-
0.75	0.50	0.74	0.92	0.93	0.64^-
0.9	0.27	0.37	0.91	0.98^-	0.88^-
0.99	0.45^-	0.83^-	0.92^-	0.73^-	0.43^-
0.995	0.87^-	0.75^-	0.85^-	0.76^-	0.35^-
0.999	0.41	0.54^-	0.93	0.97^-	0.24*

Results for the traditional and modified test are consistent. This show the scarce influence of the autocorrelation of the data on the results of this analysis.

Concerning the absolute values, an increasing but non-significant trend emerges for the higher quantiles of the shorter durations. The trend tends to vanish with the increase of the duration and becomes a significant downward trend for the larger ones. This is evident also analyzing the regression lines related to the larger quantiles in figure 5.9 and is consistent with most of the outcomes reported in [210] for the European area. The normalized rainfall shows a general absence of significant trends.

The comparison between the two tables shows the significant influence of the data intrinsic inhomogeneity on the results. Once normalized with the median of the series, the trends seems to weaken for all the durations, with general increase of

the p-values. Also the 24-hours duration trends loose their significance. This is a significant clue on the importance of considering the non-uniformity problems when dealing with these kind of large and spatially distributed database. The relevance of the topic is also underlined in the following of this chapter, when dealing with the spatial distribution of the most intense rainfall events recorded at the national scale.

5.4 Insights from the empirical distributions

In this section, the empirical distribution of annual maxima for the whole country and for each duration separately is analysed in search of clues of the existence of a family of "uncommon" extremes, with similar characteristics at the national scale. These extreme events are usually referred to as "extraordinary", according to their intensities (e.g., 300 mm/day in [168]) and their frequent severe consequences (e.g., flash floods with devastating damages in small basins, rainfall-triggered landslides, human losses), but, as aforementioned, they have been seldom expressly explored as a definite population [169, 168].

Aim of this section is to provide elements for the selection of the super-extremes as a possible definite category within the annual maxima database. Implicit long-term objective is to revise the assessment of the probability of occurrence of these events in space, regardless of the information contained in the individual series of the stations which have already recorded an extraordinary event.

5.4.1 Spatial distribution of the extraordinary events

An exploratory way for selecting only the extreme rainfall amounts is to rank them all and consider only the very top ones. Between 1930 and 2010 there are 80 years of observations, and a viable selection could be to consider in average the highest event per year. Given the whole amount of 80000 data for each duration this involves to select the top 0.001 of the empirical cumulative frequency distribution. Table 5.3 reports the quantiles (in absolute and relative terms) related to the 0.001 frequency of exceedance of the cumulative empirical distribution.

Table 5.3 Quantiles and number of exceedances per duration

d (h)	Number of exceedences	Absolute quantile (mm)	Relative quantile
1	80	101.81	3.79
3	81	175	4.09
6	80	225.88	4.19
12	80	299.67	4.07
24	81	394.95	4.07

The rain gauges with records exceeding the 1/1000 quantile for the duration 1 and 24 hours are mapped in figure 5.11. The number of exceedances is evaluated considering the absolute records (panel (a) and (b)) and the normalized rainfall (panel (c) and (d)). For other durations, please refer to appendix D.1.

Considering the absolute values, for all the durations the over-threshold values are mainly clustered in the Liguria and Friuli Venezia Giulia regions and, in the south, in Calabria and on the eastern coast of Sicilia.

Comparing the spatial distribution of the over-threshold values with the median values of the series in figure 5.4 one can see that there is a significant overlapping between the two sets of maps. The clusters identified correspond to the area with high medians. The selection procedure carried out with the absolute values shows the areas with more intense rainfall in absolute terms. This represents just a partial picture of the hazard related to the development of extreme systems, as the ground effects usually strongly reflect the deviation from the average rather than the absolute value, underlying one more time the need of considering normalized values.

Panels (c) and (d) show a more homogeneous spatial distribution, even if some clusters still persists, specially for the 24 hours duration.

Also the absolute amounts of the values exceeding the selected thresholds show a very diversified pattern (see figure 5.12 for the 1 and 24 hours duration and appendix D.2 for the others). It is common knowledge that most of the exceptional rainfall events occurred in the last century in Italy interested the Liguria region, that holds 4 of the 5 record events reported in table B.2. Judging the exceptionality of the rainfall amount related to the local climatology shows that “more extreme” events seems to have occurred also elsewhere in the country.

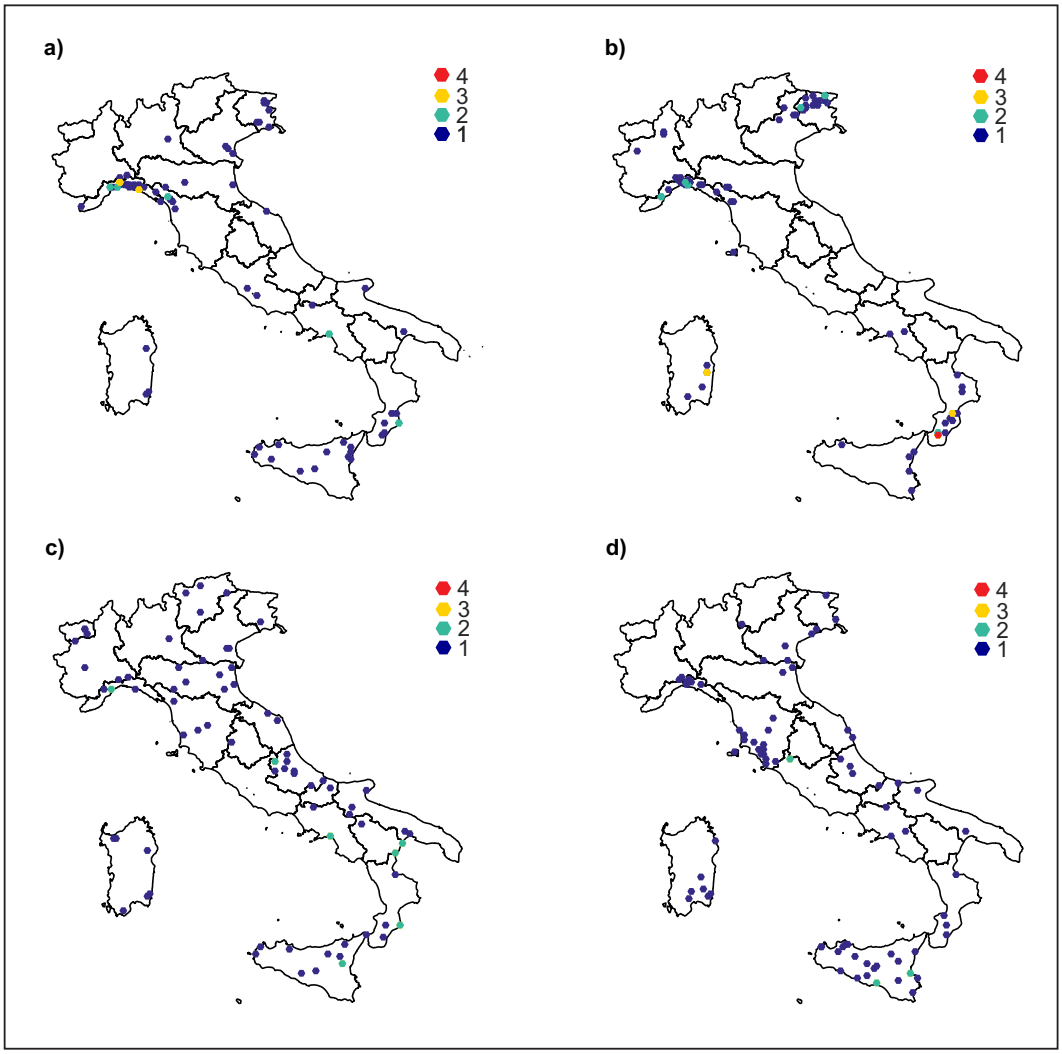


Fig. 5.11 Number of exceedances of the 0.999 cumulative probability of the absolute values for the duration (a) 1 hour and (b) 24 hours and of the normalized rainfall for the duration (c) 1 hour and (d) 24 hours. The color scale refers to the number of exceedances per series.

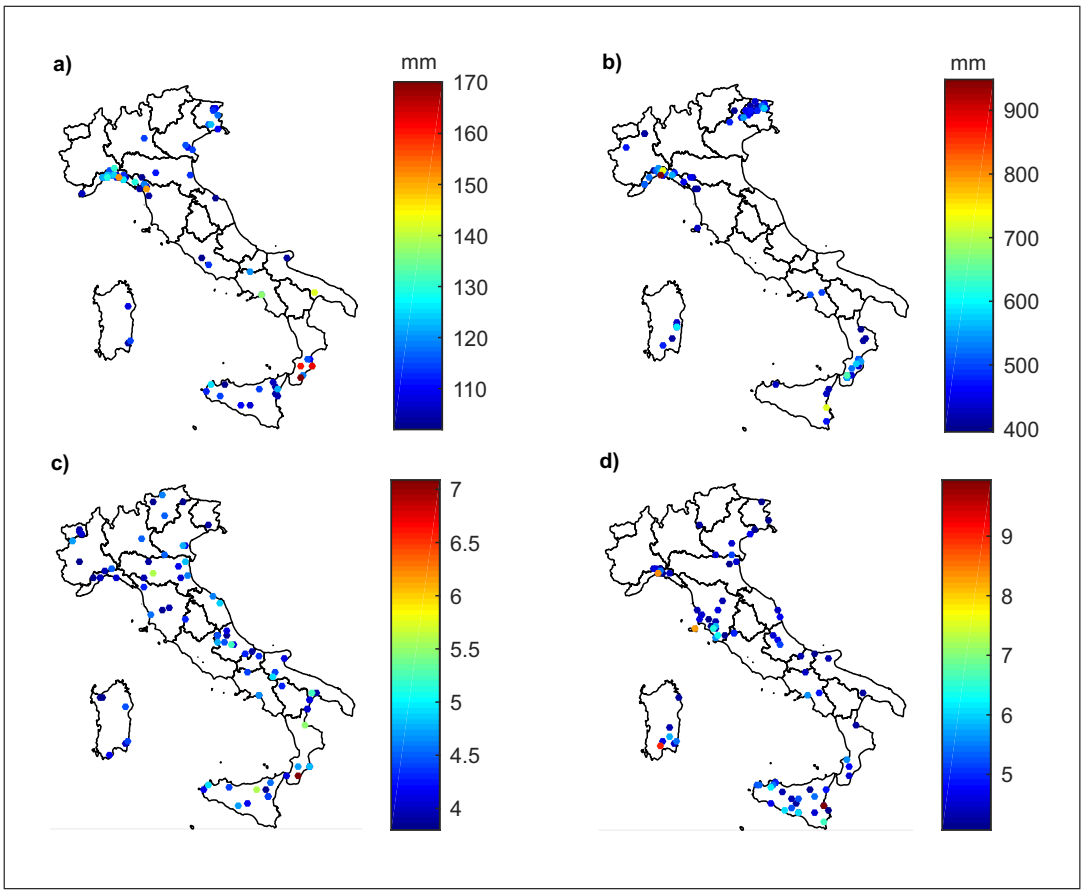


Fig. 5.12 Maximum rainfall (mm) recorded at the stations with at least one value over-threshold for (a) 1 hour and (b) 24 hours. Maximum h^* recorded at the stations with at least one value over-threshold for (c) 1 hour and (d) 24 hours.

5.4.2 Exploring the cumulative empirical frequency distribution

It is interesting to see if some hints for super-extremes selection arises from the empirical frequency distribution. Figure 5.13 reports the empirical frequency distribution of the normalized dataset for the different durations. Figure 5.14 focuses on the upper tail of the distribution of the normalized values on bi-logarithmic scale.

The shape of the empirical distributions seems to suggest a common behavior for the annual maxima of durations greater than 3 hours. The hourly maxima series only seems to show a more exponential shape, while the larger durations clearly deviate from the exponential for the larger values, possibly denoting a sub-exponential ("heavy tailed") behavior [193, 107, 108]. For instance, in [146] the authors compare the upper part of empirical distributions of thousands of daily records from around the world with four common theoretical tails: those of the Pareto, Lognormal, Weibull and Gamma distributions in order to assess the behavior of the tails. They underline that the adoption of "light tail" models, when "heavy tailed" are required would result in a significant underestimation of the quantiles for a given return period. This clue requires particular attention, as it has practical implications whenever a national statistical analysis of the extreme rainfall would be attempted.

The existence of two distinct populations of extremes has been tackled first for Italy in [169], with the definition of the $TCEV$ distribution. The authors identified a second $EV1$ distribution representing the occurrence of a sub-sample of large values in a national database of annual maximum floods. More recently, some authors have focused on the need to consider all the available data at the national scale together, in order to identify a unique threshold [168] that would allow to discriminate a national sub-set of super-extreme rainfall events. The availability of a wide national database, increasing the sample size for the analysis of the upper tail of the distributions, would allow to overcome the limitations in the robust estimation of a specific component for the representation of the super-extremes (as in the $TCEV$ distribution). In practical terms, with very few manifest super-extremes available so far, the estimation of the $TCEV$ second component would result with a too high estimation variance at whatever single-series or even regional estimation exercise. Said that, two main operational problems prevent a simple identification of a threshold to discriminate between common and uncommon extreme events:

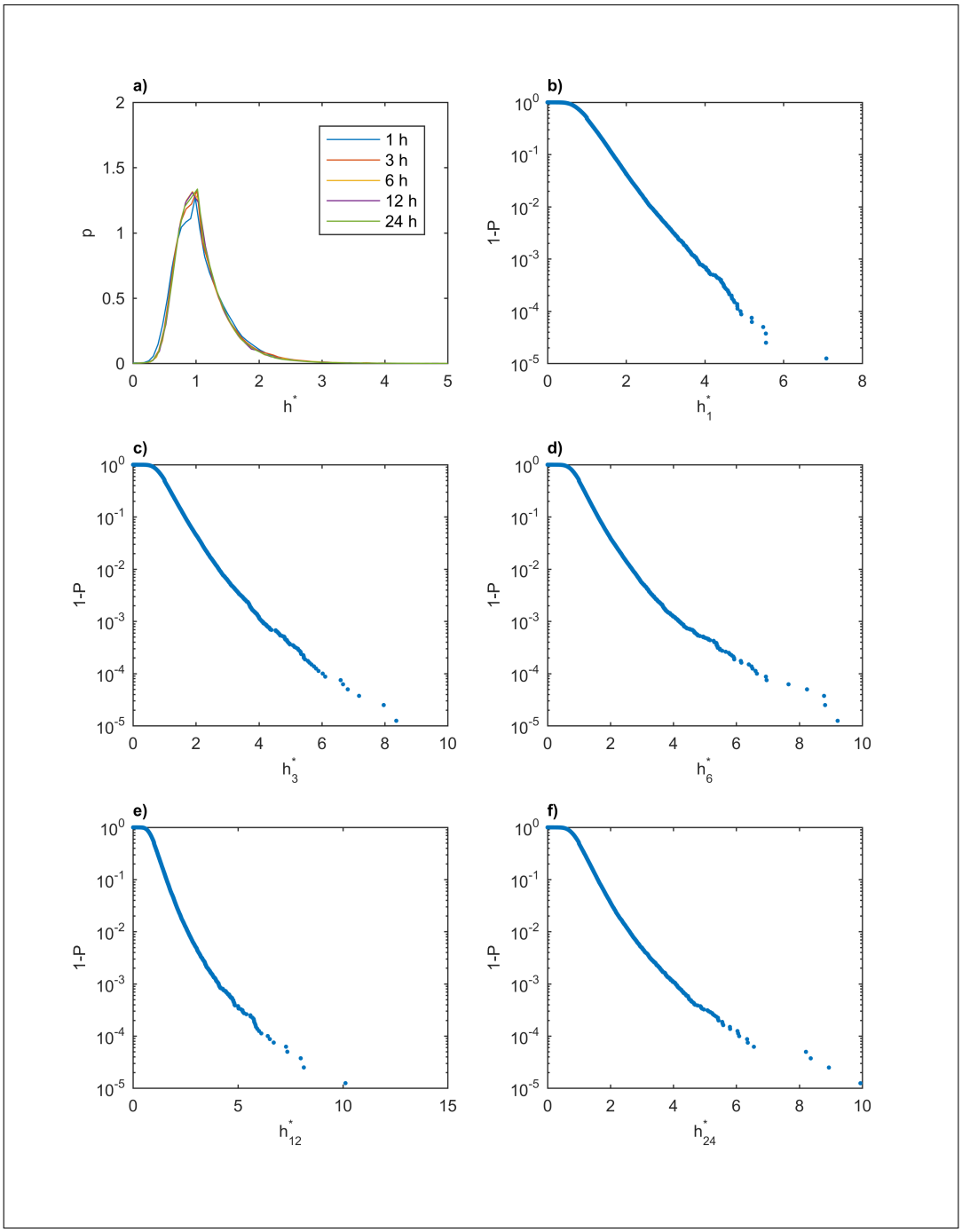


Fig. 5.13 Normalized rainfall depths: (a) empirical probability density functions for the 5 durations (the x-axis is upper limited at 5 for clarity) and and empirical frequency exceedances (1-*P*) for durations from 1 (b) to 24 (f) hours.

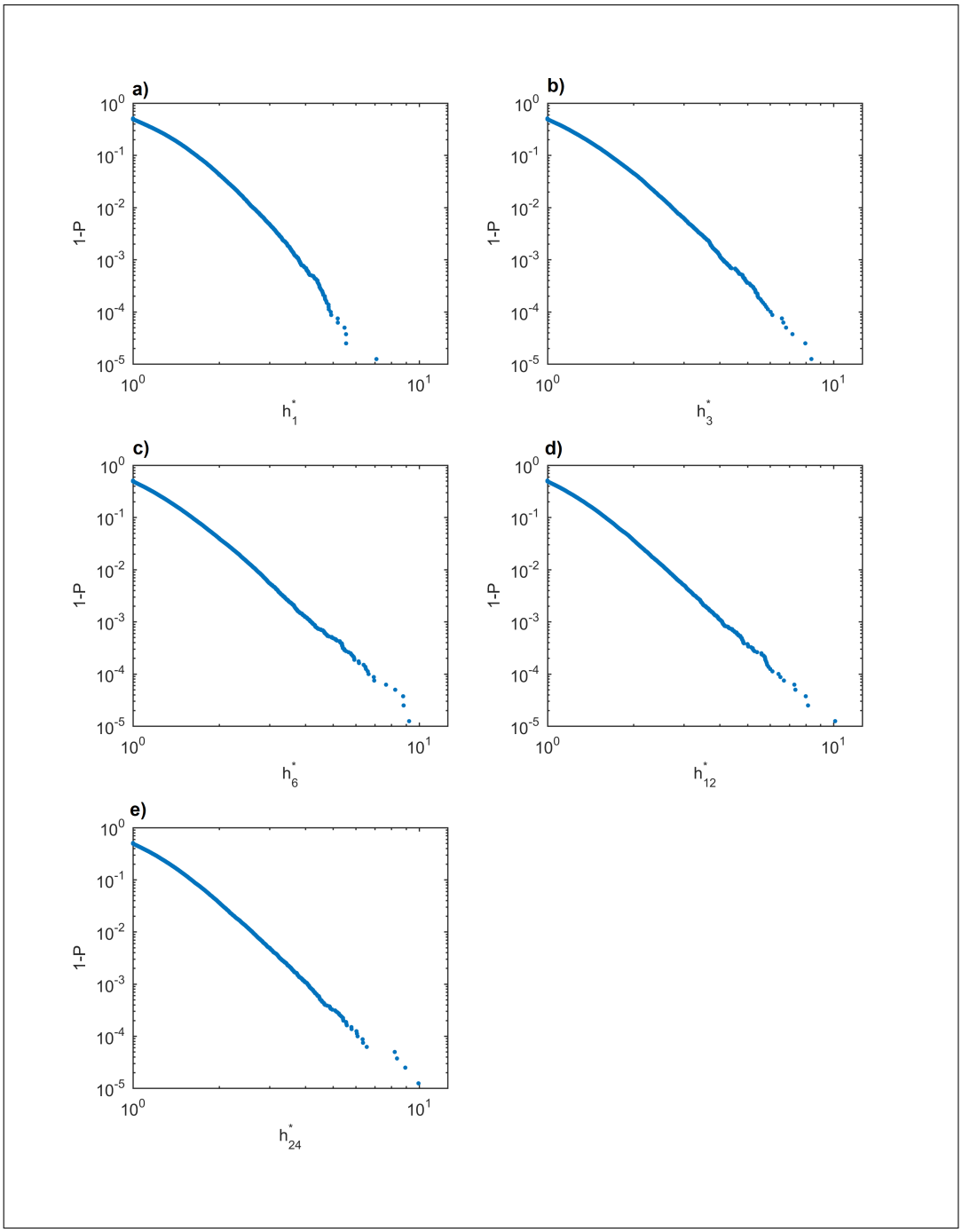


Fig. 5.14 Normalized rainfall depths: Empirical frequency exceedances (1-*P*) of the upper tails for durations from 1 (a) to 24 (e) hours on bi-logarithmic scale.

1. The need for a theoretically-based criterion for the identification of the threshold value.

2. The problems arising from the inhomogeneity of the rainfall datasets.

The first issue has been faced in the literature. [23] proposes a theoretical formula for the assessment of the probability that a certain data belongs to the extraordinary component of the *TCEV* distribution. This approach is applied in [30] to a case study in Italy, in order to identify a separation criterion. The authors concluded that different thresholds values needed to be identified for the different regions analyzed, but with little detail in the adopted criterion.

According to [169], both the "ordinary" and the "extraordinary" components of the *TCEV* distribution follow an $EV1$ (i.e., Gumbel) distribution. The 5 series obtained pulling up together all the data for every duration are therefore plotted on the Gumbel probability paper, to exploit all the information included in the *RED* to assess the existence of any clue the an effective existence of a two different populations. Graphs related to the normalized rainfall are reported in figures 5.15, 5.16, 5.17, for the absolute values graphs are attached in appendix D.3.

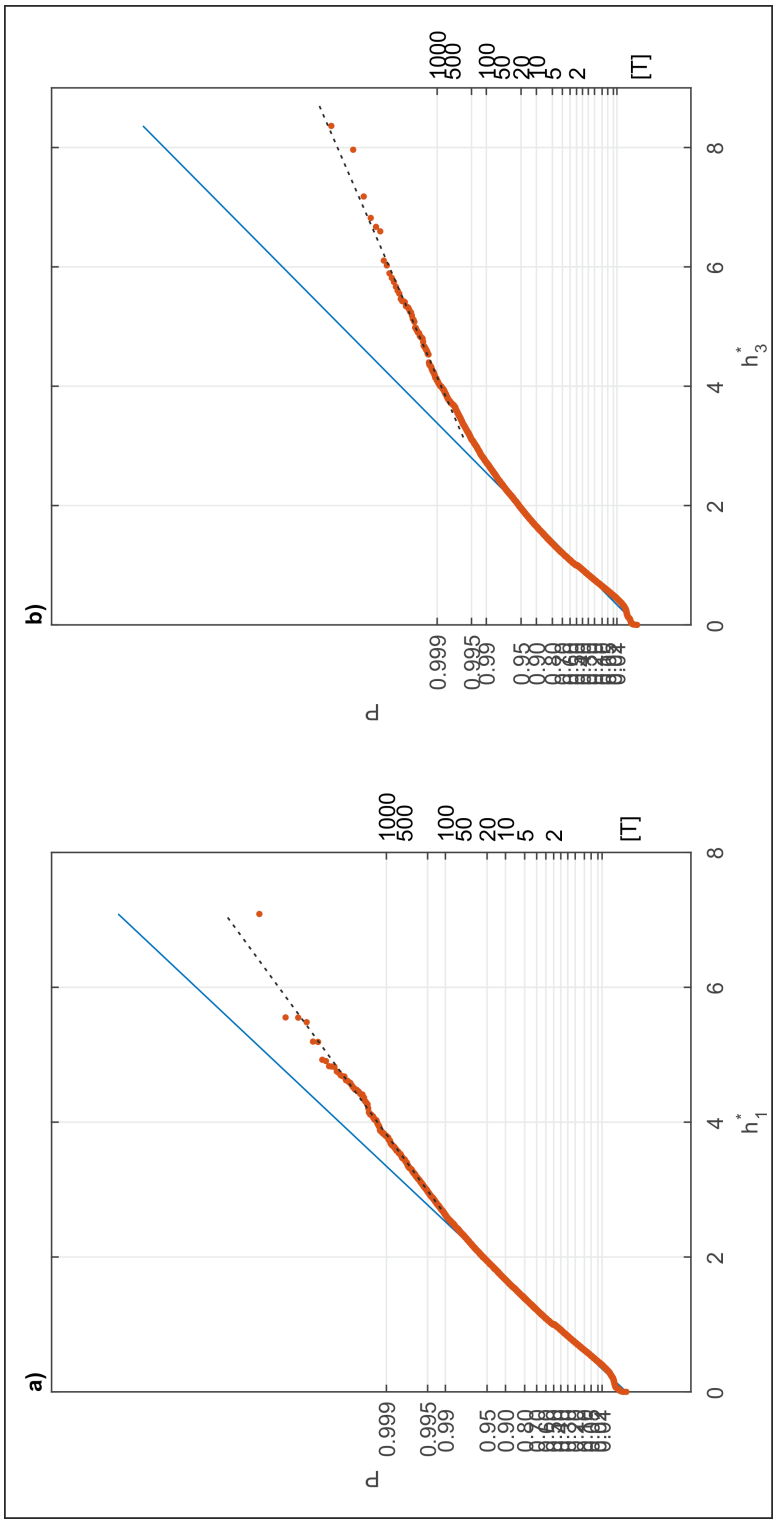


Fig. 5.15 Annual normalized maxima for (a) 1 hour and (b) 3 hours duration on Gumbel probability paper. The blue line represent the Gumbel distribution with parameters estimated by the L-moments method. A qualitative black dotted line traces the right tale of the distributions.

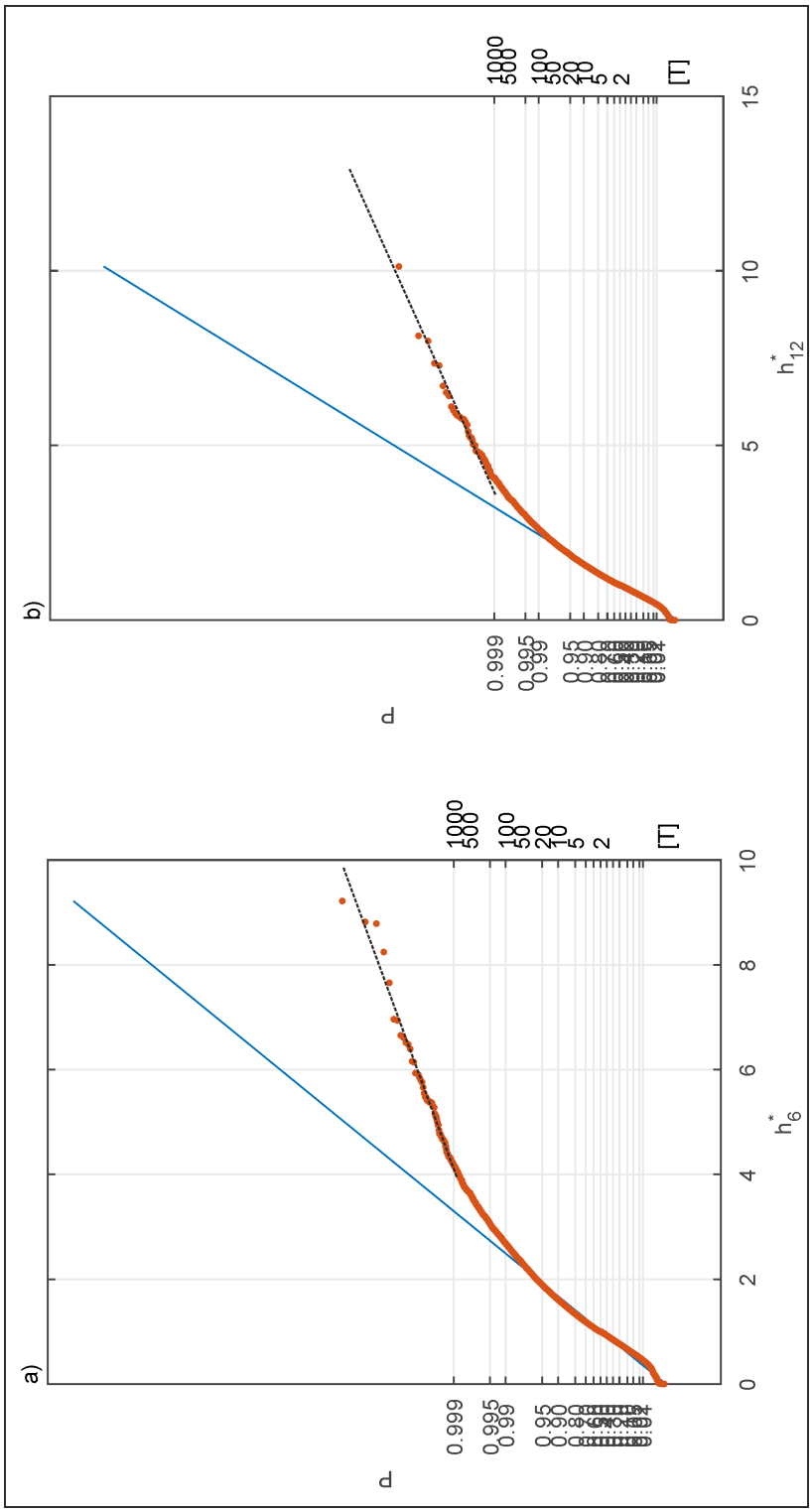


Fig. 5.16 Annual normalized maxima for (a) 6 and (b) 12 hours duration on Gumbel probability paper. The blue line represent the Gumbel distribution with parameters estimated by the L-moments method. A qualitative black dotted line traces the right tale of the distributions.

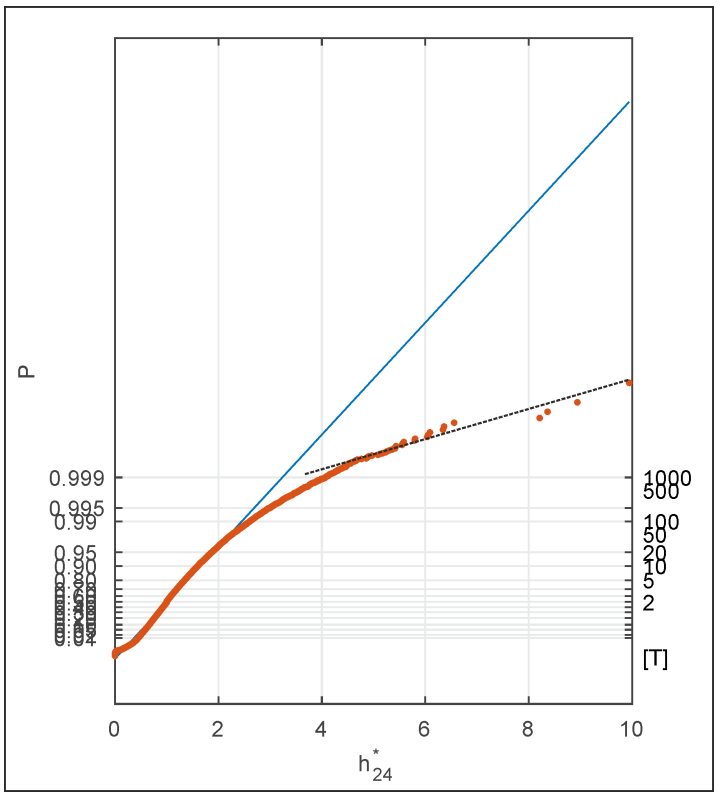


Fig. 5.17 Annual normalized maxima for 24 hours duration on Gumbel probability paper. The blue line represent the Gumbel distribution with parameters estimated by the L-moments method. A qualitative black dotted line traces the right tale of the distributions.

From all the figures a well-defined component significantly deviating from the Gumbel distribution (plotted in blue) emerges. Considering the normalized rainfall, for h^* larger than ~4, the data seems generally lined-up. The only exception is the 1 hour durations, that shows a different behavior. The behavior of this component, overlying a qualitative straight line to the upper tails is then traced. This value coincides roughly with the percentile considered so far. This analysis does not provide the definitive test of the existence of a independent population of "extraordinary" events, as many other non-uniformity factors can lead to values far from an average behavior (e.g., different origin of the events, different areas, etc.). However, this evidence, can be considered an interesting clue for further analysis aimed at assessing the consistency and the spatial distribution of this family of events that, somehow, seems to differ from the rest of the population.

5.4.3 Effects of the uneven spatial distribution of the stations

In this section the spatial distribution of the rain gauge network and its effect on the results of the analysis is explored. Rain gauge spatial distribution turns out to be a significant potential source of non-uniformity of the results at the country scale. The uneven spatial distribution of the network, and its not uniform growth in time must be taken into account in the analysis to test if some spatial bias is affecting the results. E.g., higher densities of stations in some areas could easily lead to the clustering of a higher number of exceedances in the more gauged areas. In this regard, the geographical shape of Italy is a source of potential problems in itself, with the relevant development of coastal stations, not being accompanied (obviously) by any measurements on the sea. This could undoubtedly produce problems with border artifacts.

To assess the entity of these effects, a gridded domain with an initial mesh size of 50 km is introduced. Figure 5.18 shows the number of station-year, i.e., the total number of data per cell. If data consistency changes for the different durations the shortest one is considered. The non uniformity of the network density clearly emerges at first sight, with some cells presenting almost 10 times the number of data of other cells. The most densely gauged cells can be found in the North-West of the country, in particular in Liguria region, in the northern Toscana and, in the North-East. Lazio and Umbria show, in the center of Italy, cells with less than 200 station-years.

The results obtained in the previous sections has been gridded accordingly. Figure 5.19 shows the gridded highest quantiles for the 1 and 24 hours duration. For all the other durations maps are available in appendix D.4.

Panels (a) and (b) of figure 5.19 report the number of over threshold records per square considering the absolute values. A very clustered behavior is apparent, especially for the 1 hour duration, with almost half of the values in the Ligurian area. For the other durations the clusterization is still evident, with almost no exceedances in the central part of the country. On the other hand, when considering the normalized rainfall (panels (c) snd (d)) a more uniform pattern is obtained. High-density squares still emerge, but there are no areas of the peninsula characterized by no exceedances at all. Finally, panel (e) and (f) show the number of exceedances considering the

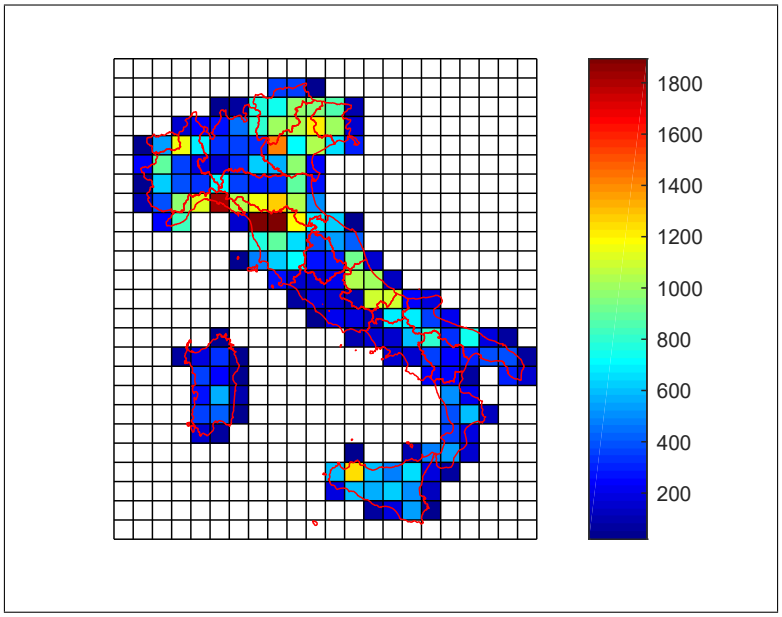


Fig. 5.18 Number of station-year per cell over a 50 km grid.

normalized rainfall, divided by the number of station-year per cell, as reported in figure 5.18.

From panel (e) and (f) an almost homogeneous pattern of super-extremes across the country emerges. For all the durations the differences between the different cells are quite negligible This is confirmed even exploring other spatial resolutions, as shown in appendix D.4.2.

The results above show the large influence of the density of the rain gauge network on the assessment of the risk related to super-extreme rainfall. Ignoring this component could lead to misinterpreting the results of the analysis, focusing the attention on a limited area of the country. From the outcome of this study the spatial susceptibility to these events seems quite uniformly distributed.

The same maps, reported at different resolutions in appendix D.4.2, seems to suggest that the few outliers in terms of occurrence/density ratio can be attributed to the location and size of the selected grid, with some sea-borders artifacts. The normalized spatial distribution of the super-extremes can be therefore considered quite uniform at the country scale.

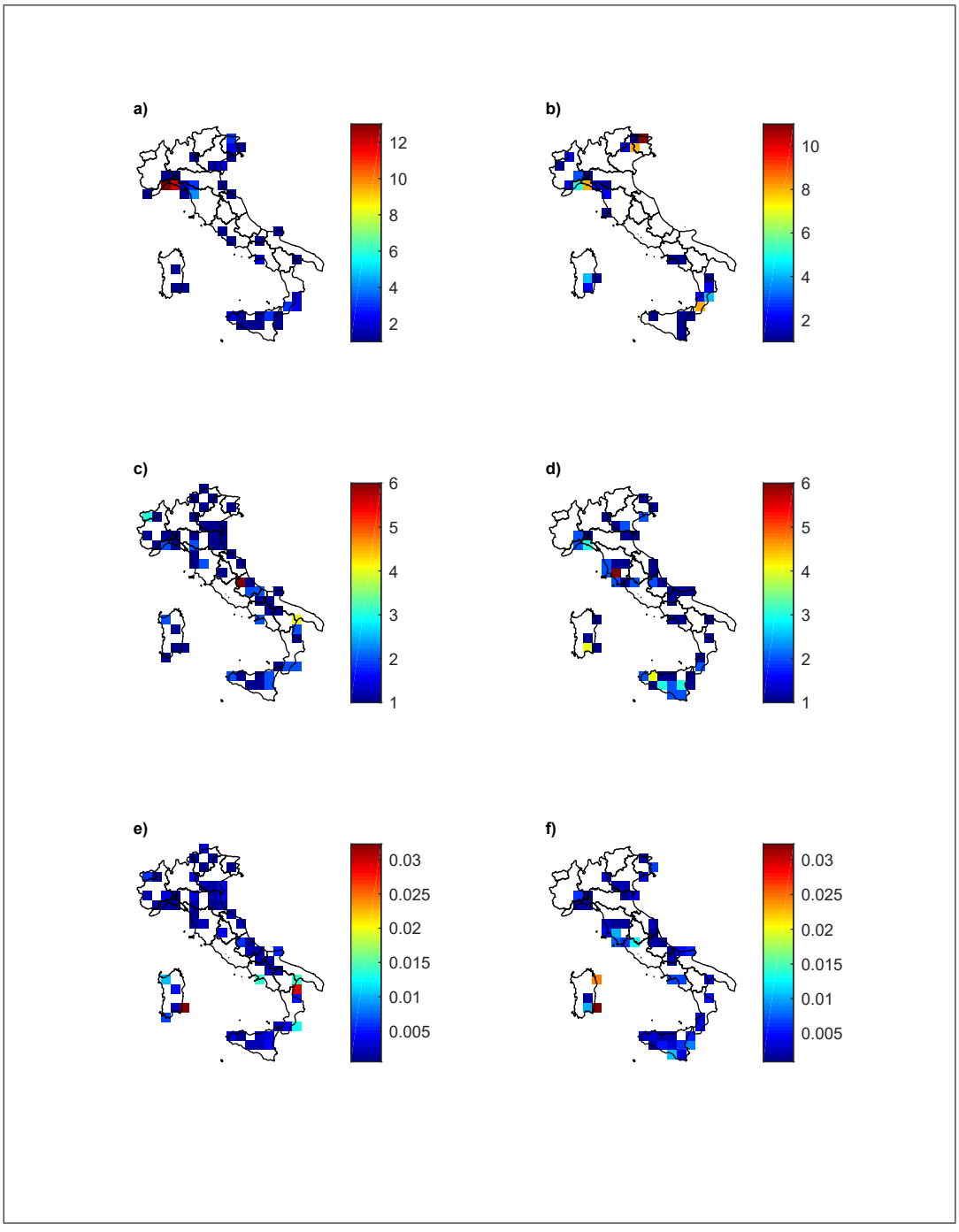


Fig. 5.19 Number of exceedances of the 0.999 cumulative probability distribution on a 50 km gridded domain considering (a) and (b) the absolute rainfall amounts and ((c) and (d)) the normalized ones. Panels (e) and (f) show the same values of (c) and (d) normalized with the numer of station-years per cell. Left panels refer to 1 hour duration, right panels to 24 hours.

5.4.4 Morphological clues of the spatial distribution of the over-threshold extremes

Even considering the dense developed database, the significant issue concerning the spatial problem of the extreme rainfall hazard assessment in areas that have never recorded an over-threshold event is still open. Are there any evident differences between those cells and the others? Did any event affect those cells in the past without being recorded? Are there any chance of future development of intense rainfall systems? For trying to give an answer to those questions with the available data, further analyses are needed.

In this section the possible role of morpho-climatological factors on the development and intensity of the analyzed events is preliminary explored, following the examples of different authors (e.g., [157]). The morphological analysis, combined with the significant contribution of authors who have investigated long time series of synoptic meteorological configurations in search of evidences of causative factors for extraordinary rainfall amounts [110, 77], is proposed as natural future development of the work proposed in this dissertation.

More in the detail, an operative example of analysis that tries to extract useful indexes from morphological structures that can favor the development of extreme events is proposed. This study stems from the mechanisms that are known to be able to generate extreme rainfall amounts, e.g., the orographic lift, and convert this in simple morphological characteristics. The focus is on simple geometries, in order to keep the procedure amenable for an operational and straightforward use on a complex wide territory. Different authors have explored the feasibility of this kind of approach (e.g., [157, 76]), but most of them ended up focusing on smaller domains, due to the accuracy required in the definition of the indexes. The analysis is carried out on the 80 events identified with the relative threshold defined in section 5.4 for the durations 1 and 24 hours. Codes, names and coordinates of the stations are attached in appendix D.4.3. For representing the topography, the ASTER GDEM v2 Digital Terrain Model is adopted. The DTM has 1 arcseg resolution (~30 m) in space and is available at [142]. For further informations refer to [187].

This preliminary approach tries to expand the work of [76] to a wider domain. In [76] the authors present a case study related to the Campania region. They find a set

of topographic indexes related to an amplification of the recorded rainfall amounts. Those indexes are related to a set of morphological objects identified in a preliminary analysis in relation to the typical directions of wet air masses in the region. The first drawback of the approach is that such a detailed morphological analysis is hard to reproduce at the national scale. Furthermore, the direction of wind and streams at the national scale are not so easy to be classified.

Simpler characteristics, able to be built and analyzed on a large continuous spatial domain, will be here considered. The identification of the characteristics stems from the mechanism generally related to the development of extraordinary rainfall. Basically: wet air masses originating from the sea are transported forward to the continent (warm or cold currents); when they collide with a sudden barrier they release water vapor content because of the saturation generated. If the amount of available humidity is large and the barrier shape allows an almost total blocking of the flow an evidence of this mechanism would result in a specific index.

Three main morphological characteristics are considered in this scheme:

1. The distance from the sea
2. The slope of the obstacle
3. The elevation of the obstacle

At first, the slope and elevation characteristics are analyzed in the 8 cardinal directions, in neighborhoods from 1 to 4 km^2 in the surrounding of the stations presenting at least a super-extreme. The 1 hour and 24 hours dataset are analyzed separately and any correlation between the characteristics of the stations is explored. Results seems to suggest quite weak similarities between the different stations. The complex interactions that characterize the phenomena apparently prevent a simple mono-variate regression to give meaningful information.

Knowing that wet air masses come from the sea, a 180° domain in the surrounding of the minimum distance between the station and the sea is explored, with a maximum radius of 2 km. Sections centered in the station coordinate are analyzed using an angular step of 5° (e.g., figure 5.20a).

After analysing some well-known case study, a critical height for the block of the stream of 1000 m is considered.

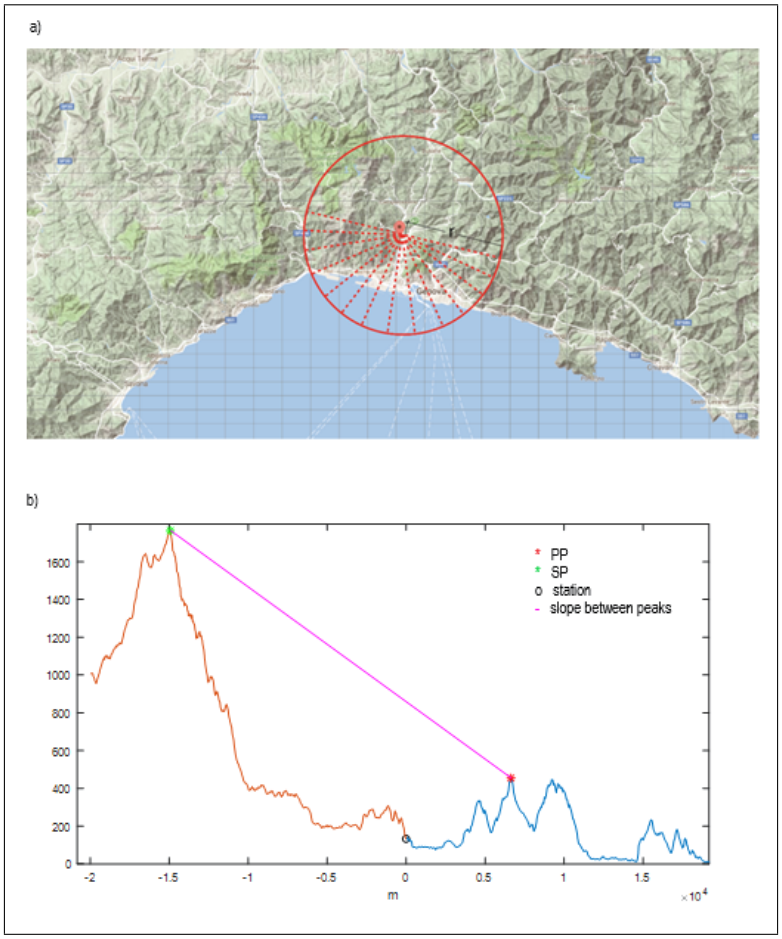


Fig. 5.20 (a) Range of slope calculation profiles. (b) Station 8027 Bugnato, an example of the scheme adopted for the analysis.

From the station towards the sea along the section, the higher peak with altitude below 1000 m is considered, and referred as the primary peak (). If any peak above 1000 m altitude interposes between the station and the seas, the section is considered null, as the wet air mass can not overly the station.

From the station towards the hinterland, the peak generating the larger slope with the *PP* is considered (it will be referred as secondary peak). Considering that the *PP* can facilitate the ascent of the air fronts and lead to the overriding of the *SP*, a difference in the elevation of at lest 200 m is required. If this condition is not respected, the actual *SP* is ignored, and a peak with larger elevation is considered. If no compatible peaks are available, the section is considered null. The slope between the *PP* and the *SP* is then evaluated for each non-null section (as in figure5.20b).

For some stations (mostly the one located in the Pianura Padana) it has not been possible to identify the main directions, as all of the section have been considered null. The flat areas in which this stations are located and the large distance from the sea give attribution to a different nature of the precipitation system. At this preliminary stage these stations have been ignored.

The maximum slope between all the sections and the related direction are then considered for each station. The second and third slopes and the related sections are also taken into account. Results are reported in figure 5.21 for the 1 hour duration and in figure 5.22 for the 24 hours duration.

As expected, results are far from depicting an unique pattern for all the station. The differences in the climatic characteristic and orographic settings across the country imply different directions and different mechanism of the super-extreme genesis. On the other hand, this technique downgrades the problem of the complex orographic setting, reducing the problem to a mono-dimensional one or to a subsequent application of mono-dimensional models in the principal direction identified (to represent the reality in a more complete way, the channeling of the air masses in curved valleys should be taken into account).

For a preliminary assessment of the validity of the obtained directions, a first visual comparison has been carried out with the main tracks of the extra-tropical explosive cyclones. Those phenomena, characterized by exceptionally and unusually large deepening in the mid-latitudes [110], with rate of at least 1 hPa/h for 24 h [24] and characteristics similar to the tropical ones, having warm core systems with a

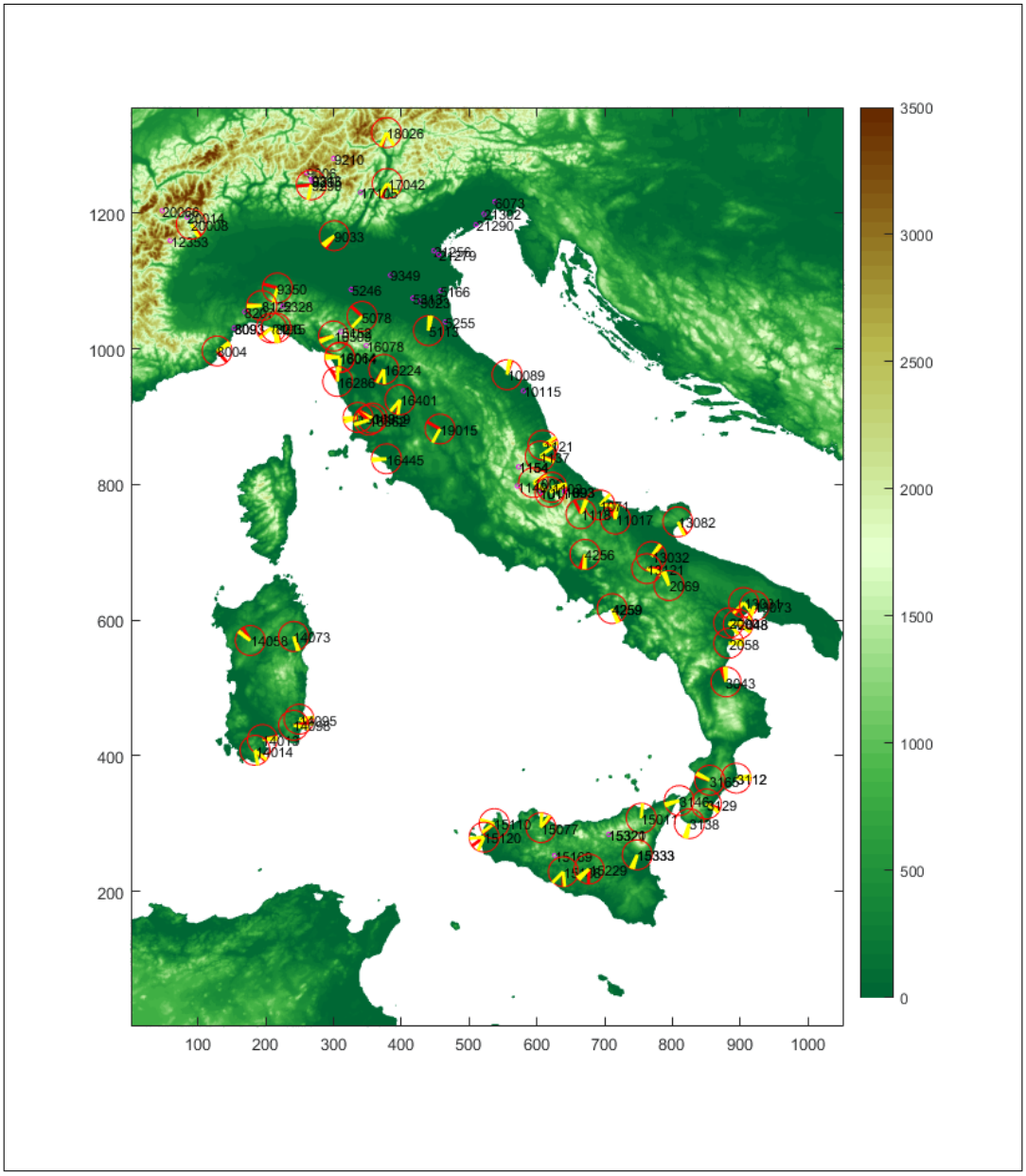


Fig. 5.21 Results of the analysis of the slope between peaks for the 1 hour duration. The red line refer to the maximum slope detected. The yellow ones to the second and third ones.

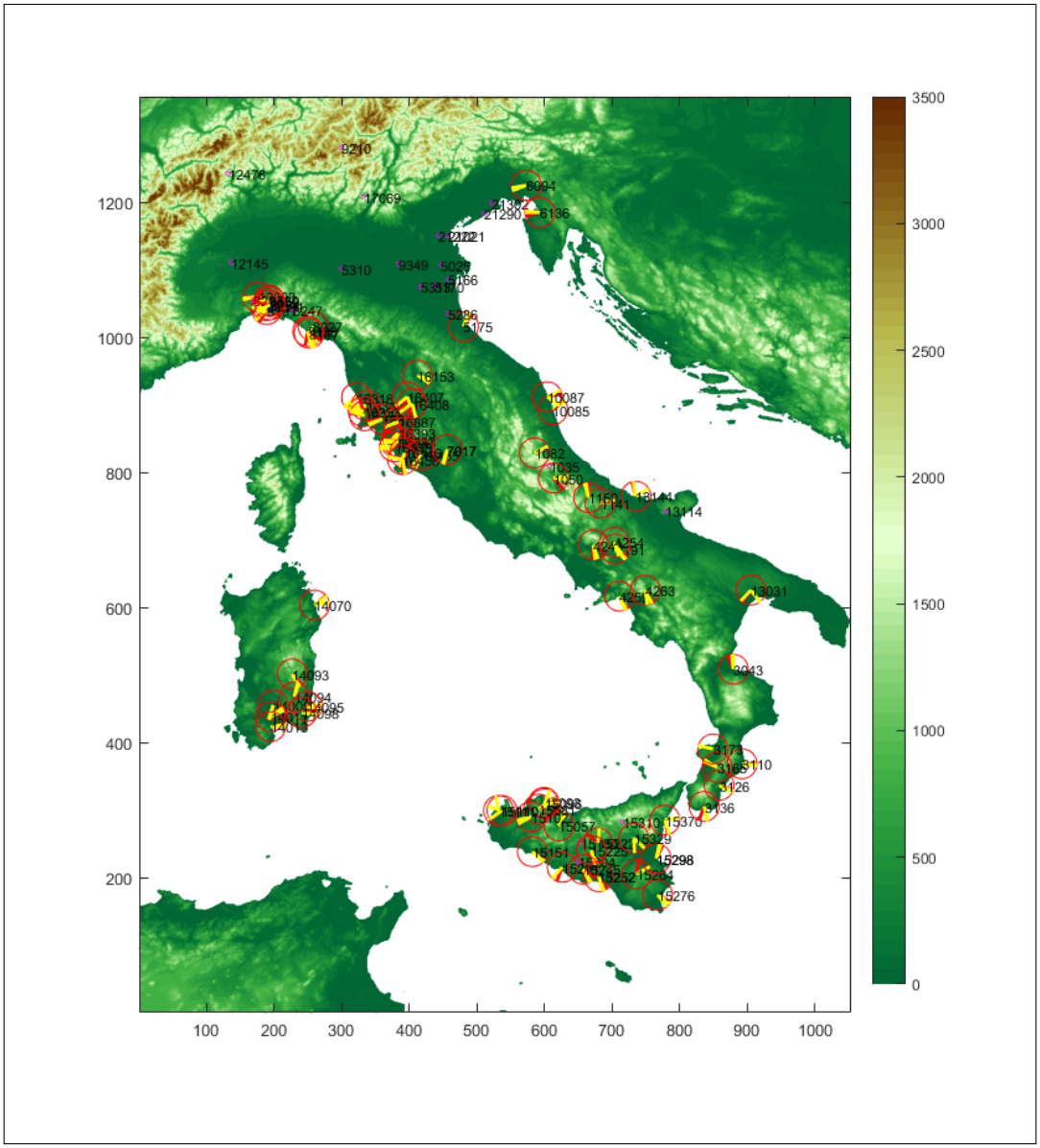


Fig. 5.22 Results of the analysis of the slope between peaks for the 1 hour duration. The red line refer to the maximum slope detected. The yellow ones to the second and third ones.

clear eye, are quite common in the Mediterranean area, are frequently associated uncommon rainfall amounts [202]. In [110] the authors, after studying the path of 40 years of cyclone tracks, conclude that most of the explosive cyclones with cyclogenesis in the Gulf of Genoa, in the Tyrrhenian Sea and the Ionian Sea tend to move southwards and towards the East Mediterranean. The ones whose cyclogenesis is located in the Gulf of Syrte (North Africa, near Libya) tends to move northwards to the Italian peninsula.

Considering main directions, extremely simplified in figure 5.23, some affinties with the previously identified main directions can be identified.

The proposed analysis is still far from being able to provide evidences of the existence of areas more prone to the development of extraordinary rainfall events, as, without the introduction of further climatological information, an objective identification of the its potentiality can not be pursued. However, the proposed framework is simple, consistent with the knowledges on the genesis of those phenomena and, further refined, can be an interesting instrument for the fast-identification of extreme rainfall prone areas. Further efforts are to be addressed in two main directions:

- Increase the detail in the identification of the cyclone tracks, to assess the real interactions with the slopes.
- Convert the 1-dimensional model to a 2-dimensional model, able to represent the wet air flow as a plan, interacting with a non-regular surface. This would allow to take account the phenomena of channeling and diversion of the air masses when colliding with an obstacle.

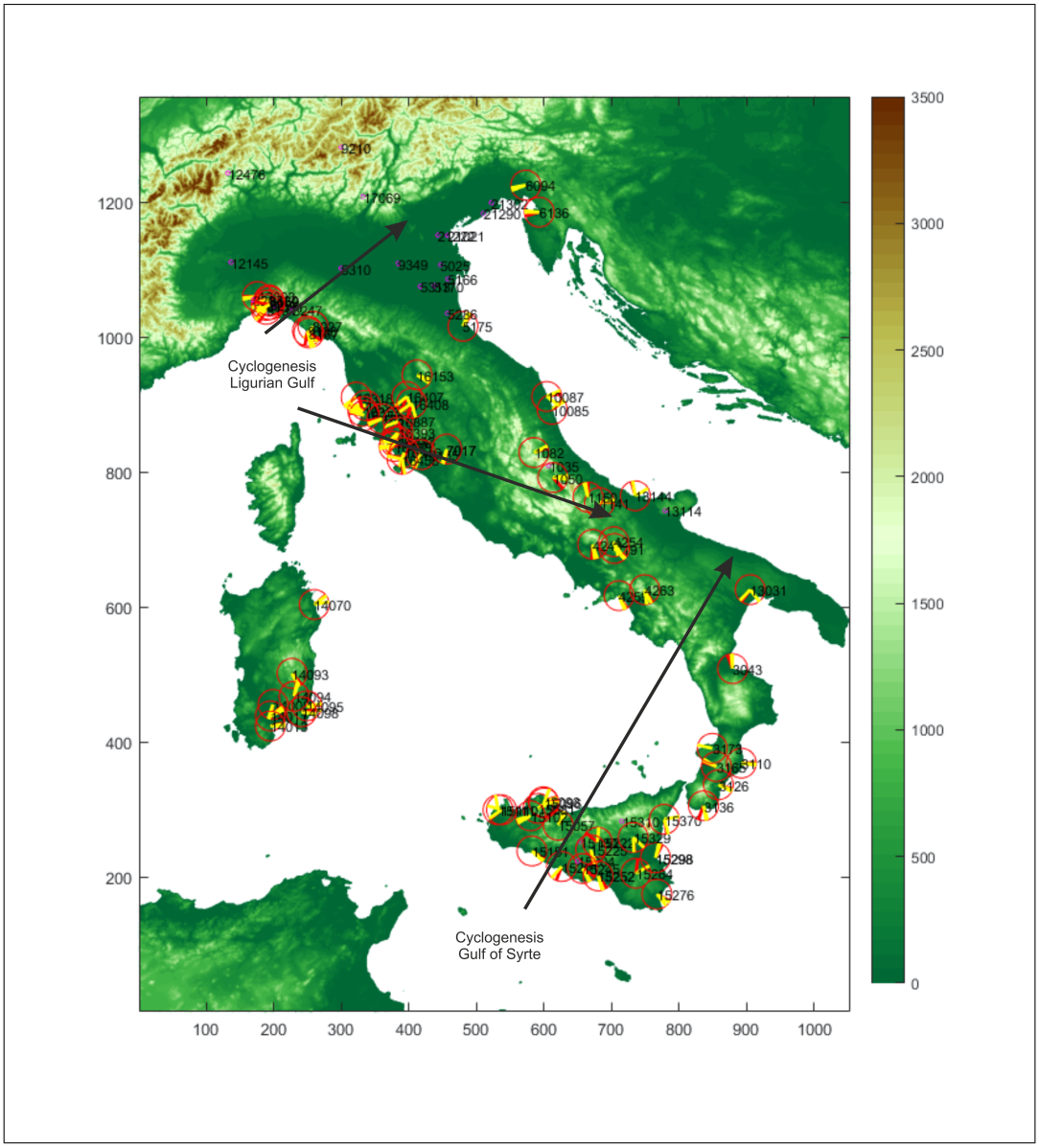


Fig. 5.23 Scheme of the main cyclones directions, overlapped to the map of figure 5.22

Chapter 6

Conclusions

The research work described in this dissertation investigates empirical evidences, systematic data, tools and methodologies for improving the assessment of the hazard related to extreme rainstorms. The dissertation reports the efforts made in different directions, aimed at providing advances for the definition of a comprehensive framework for improving the rainstorm estimation in gauged and ungauged locations, at different spatial scales, according to different technical needs. In this regard, each chapter individually covers a defined technical and spatial domain, as summarized in the following.

Chapter 2 describes a robust statistical framework for dealing with uneven and fragmented rainfall records on a regional non-uniform spatial domain, in order to estimate the frequency of the extremes in ungauged sites. The proposed "patched kriging" technique allows at exploiting all the information available from the recorded series, independently on their length, and to provide annual extreme rainfall estimates in ungauged areas. The methodology involves the sequential application of the ordinary kriging equations, producing a homogeneous dataset of synthetic series with uniform lengths. This provides robustness and low sensitivity to local artifacts during the distribution estimation phase. L-moments expressions weighted on the kriging variances are implemented, to compensate the smoothing effects of kriging. The distributions commonly considered for rainfall analysis in the study area are then evaluated on the patched dataset. Results underline the importance of a data-driven evaluation of the best-fitting distribution at the local scale. Despite the evident improvements in the significance of the resulting maps, the spatialization of the

rainfall information is probably still far from being consistent with the real behavior of the rainfall fields. Interpolation techniques can only reflect the estimation variance determined by the spatial and temporal data resolution. The "patched kriging" allows to significantly increase the amount of information that can be extracted from the rain-gauge records, but the point at-gauge information turns out to be not sufficient for representing complex localized rainfall systems. This limitation is inherent in the nature of rain gauge data and can not be tackled without inserting other source of information in the system, e.g., spatial remote sensed data. In the subsequent chapters, some aspects concerning the potential of remotely sensed information for integrating rain gauge data in the spatial characterization of rainfall fields are then explored.

In chapter 3, weather radar data are analyzed in comparison to to high temporal resolution raingauge data. Data deriving from C-band weather radar are known to underestimate rainfall depths, particularly the in the domain of the high intensities. Using a set of data related to several events recorded in the North-Western Italy, an adaptive procedure aimed at increase the quality of the radar rainfall estimation is developed. The so-defined ATS methodology makes use of confined spatial and temporal domains for pursuing a quasi real-time adaptive calibration of the relation between radar reflectivity and rainfall rate, that is, therefore, spatially varying. The final aim is to allow building systematic high-resolution rainfall maps, regardless of local conditions, characteristics of the radar, sampling methodologies and spatio-temporal distribution of the events under analysis. Results obtained are encouraging, as the ATS technique outperforms the classical estimation methods for most of the analyzed events. The technique still needs to be improved for providing more robust performance dealing with convective events. This is not surprising, as the uneven distribution of the rain gauges minimizes the help that neighbor stations can provide, in each time step of each event, to the identification of a local numerically robust spatial domain. Refinements are under study, and would entail the development of adaptive calibration windows that follow spatially the evolution of the event, considering also other radar variables in the calibration phase.

Chapter 4 investigates the validity of the information content of $TRMM$ rainfall satellite products for assessing the spatio-temporal characteristics of extreme rainfall at the synoptic scale. A preliminary analysis carried out in the Liguria region

confirms a marked underestimation of extreme rainfall depths, but also highlights the ability of *TRMM* data to identify the dates of occurrence of severe rainfall events. Considering the second feature, the work is focused on the ability of both *TRMM* and the recently launched *GPM* mission to help identifying the timing of severe rainfall events on wide spatial domains. Comparing the date of occurrence of the most severe daily rainfall recorded each year by a global rain gauge network with the ones estimated by *TRMM* a match rates approaching 50% is found, indicating significant consistency between the two data sources. This figure rises to 60% for *GPM*, indicating the potential for this new mission to improve the accuracy associated with *TRMM*. Satellite-derived timing information could help in connecting events identified in ungauged areas with the characterization of extreme rainfall in gauged regions, possibly improving the identification of extreme rainstorm-prone areas at the synoptic scale. Moreover, once longer series will be available, this kind of information could effectively be combined with rain gauge data for better defining the synoptic structure of the storm systems and ginving insights in design rainfall estimation for large basins. Further efforts for assessing the quality of results in areas with low gauge densities are required. The actual uneven distribution of the global rain-gauge databases at the different latitudes could affect the results of the analysis, preventing the possibility to unequivocally ascribe the spatial varibility in the quality of the timing-detection to the adopted verification procedure rather than to external factors (e.g., the different types of dominant events at different latitudes, etc.).

Chapter 5 is focused on the analysis of rainfall extremes at the Italian national scale, with particular attention at the detection of elements that can help in defining the notion of rainfall super-extremes. Like other Mediterranean areas, Italy is indeed prone to the development of events with explosive rainfall intensity, lasting for several hours. The main triggering mechanisms of these events is quite well known, but still this knowledge does not help in building rainstorm maps compatible with the actual probability of occurrence of these super-events. Efforts are then addressed at identifying the existence of a meaningful population of records deviating from the ordinary definition of extreme value and assessing their stationarity and spatial distribution across the country. To this aim, several sources of data are analysed, in view of the assembling of the first comprehensive dataset of extreme rainfall of short duration in Italy. The resulting dataset, referred to as *RED*, includes annual maximum rainfall recorded in 1 to 24 consecutive hours from more than 4500 stations

across the country, spanning the period between 1916 and 2014. The analysis of the dataset provides information on the features of extreme rainfall at the national scale. Preliminary normalization with the median value, aimed at removing the influence of the local climatology has been carried out and the stationarity in time of the higher quantiles is checked by pooling up all the data for each duration year by year and computing various quantiles, for removing the effect of the variability in time of the station number. The applied Mann-Kendall tests show that, even considering the autocorrelation of the data, the presence of any significant trend can be neglected. The empirical cumulative distributions of the pooled up normalized data for the different durations on the Gumbel probability paper seems to suggest the coexistence of different populations of extreme values, as suggested in [169, 168]. However, this can be just considered a clue of the existence of a "super-extreme" class of events, requiring further research efforts to overcome the inhomogeneities affecting such a wide database in space and time. The identification an homogeneous category of extraordinary events at the country scale would allow one to further explore the occurrence of these kind of phenomena, apparently negligible if considering the gauging site, but potentially revealing higher frequencies when considered on a wide-area.

For an exploratory analysis of the spatial distribution of the hazard related to those extraordinary events at the country scale, the top 0.001 values of the empirical cumulative frequency distribution for each duration is considered. Given the whole amount of 80000 data for each duration this involves considering in average the highest event per year. The spatial pattern of the locations reporting over-threshold values seems to suggest the presence of areas more prone to the development of this kind of events. However, an artifact could be constituted by the uneven station density in the space. To assess this effect, a gridded domain with a mesh size of 50 km has been used to compute, in each cell, the number of exceedances divided by the total number of available data. The consequent corrected pattern of occurrences normalized by the station density seems more uniform than the previous one for all the durations. This outcome is confirmed also changing the spatial resolution of the grid. This approach, however, cannot tackle the problem of the extraordinary rainfall hazard assessment in areas where an over-threshold event has never occurred (or, possibly, never been recorded). The adoption of a gridded domain, moreover, involves problem of the susceptibility of the outcomes

to the discretization of the domain. An approach continuous in space, capable to take into account different morpho-climatic information, as the ones proposed in [77, 76] could significantly improve the results of the analysis. Far from being able to provide evidences of the existence of areas more prone to the development of extraordinary rainfall events, the preliminary morphological analysis described in the last section of this dissertation moves in this direction. The complexity of the analyzed domain prevents a straightforward identification of the extraordinary rainfall hazard assessment based only on morphological characteristics. However, the proposed framework is simple, consistent with the knowledges on the genesis of those phenomena and, further refined and combined with climatic informations (e.g., cyclone tracks and/or meteorological reanalysis) could be an interesting instrument for the identification of extreme rainfall prone areas, making a remarkable difference in the next steps of research in extreme rainfall analysis and prediction.

References

[1] Acquaotta, F., S. Fratianni, C. Cassardo, and R. Cremonini. "On the continuity and climatic variability of the meteorological stations in Torino, Asti, Vercelli and Oropa". In: *Meteorology and atmospheric physics* 103.1-4 (2009), pp. 279–287.

[2] AghaKouchak, A., E. Habib, and A. Bárdossy. "Modeling radar rainfall estimation uncertainties: Random error model". In: *Journal of Hydrologic Engineering* 15.4 (2009), pp. 265–274.

[3] AghaKouchak, A., N. Nasrollahi, and E. Habib. "Accounting for uncertainties of the TRMM satellite estimates". In: *Remote Sensing* 1.3 (2009), pp. 606–619.

[4] Alavi, N., J. S. Warland, and A. A. Berg. "Filling gaps in evapotranspiration measurements for water budget studies: evaluation of a Kalman filtering approach". In: *Agricultural and Forest Meteorology* 141.1 (2006), pp. 57–66.

[5] Alexander, L. V. and J. M. Arblaster. "Assessing trends in observed and modelled climate extremes over Australia in relation to future projections". In: *International Journal of Climatology* 29.3 (2009), pp. 417–435.

[6] Alfieri, L., P. Claps, and F. Laio. "Time-dependent ZR relationships for estimating rainfall fields from radar measurements". In: *Natural Hazards and Earth System Sciences* 10.1 (2010), pp. 149–158.

[7] Allamano, P., P. Claps, F. Laio, and C. Thea. "A data-based assessment of the dependence of short-duration precipitation on elevation". In: *Physics and Chemistry of the Earth, Parts A/B/C* 34.10 (2009), pp. 635–641.

[8] Altinbilek, D., E. Barret, T. Oweis, E. Salameh, and F. Siccardi. *Rainfall climatology on the Mediterranean, EU-AVI 080 Project ACROSS—analyzed climatology rainfall obtained from satellite and surface data in the Mediterranean basin*. Project reports. Genova, GE: Università degli Studi di Genova, 1997.

[9] Anagnostou, E. N. and W. F. Krajewski. "Real-time radar rainfall estimation. Part I: Algorithm formulation". In: *Journal of Atmospheric and Oceanic Technology* 16.2 (1999), pp. 189–197.

[10] Anagnostou, E. N. and W. F. Krajewski. "Real-time radar rainfall estimation. Part II: Algorithm formulation". In: *Journal of Atmospheric and Oceanic Technology* 16.2 (1999), pp. 198–205.

[11] Andréassian, V., C. Perrin, E. Parent, and A. Bárdossy. "The Court of Miracles of Hydrology: can failure stories contribute to hydrological science?" In: *Hydrological Sciences Journal* 55.6 (2010), pp. 857–871.

[12] ARPA Lombardia. *Progetto Strada*. Accessed: 2016-08-01. URL: http://idro.arpalombardia.it/pmapper-4.0/map.phtml.

[13] ARPA Piemonte. *Le precipitazioni intense in Piemonte - Distribuzione regionale delle piogge e caratterizzazione statistica dei valori estremi*. Final report. Torino, IT: Regione Piemonte, 2013.

[14] ARPA Piemonte. *Banca dati meteorologica*. Accessed: 2016-08-01. URL: http://www.regione.piemonte.it/ambiente/aria/rilev/ariaday/annali/meteorologici.

[15] ARPACAL. *Centro Funzionale Multirischi*. Accessed: 2016-08-01. URL: http://www.cfd.calabria.it/.

[16] ARPAL. *Atlante Climatico della Liguria*. Final report. Genova, IT: Regione Liguria, 2013.

[17] ARPAL. *Consultazione Dati Meteoclimatici*. Accessed: 2016-08-01. URL: http://www.cartografiarl.regione.liguria.it/SiraQualMeteo/script/PubAccessoDatiMeteo.asp.

[18] Asante, K. O., G. A. Arlan, S. Pervez, and J. Rowland. "A linear geospatial streamflow modeling system for data sparse environments". In: *International Journal of River Basin Management* 6.3 (2008), pp. 233–241.

[19] Atlas, D., C. W. Ulbrich, and R. Meneghini. "The multiparameter remote measurement of rainfall". In: *Radio Science* 19.1 (1984), pp. 3–22.

[20] Awadallah, A. G. and N. A. Awadallah. "A Novel Approach for the Joint Use of Rainfall Monthly and Daily Ground Station Data with TRMM Data to Generate IDF Estimates in a Poorly Gauged Arid Region". In: *Open Journal of Modern Hydrology* 03.01 (2013), pp. 1–7.

[21] Barrett, E. C. and D. W. Martin. *The use of satellite data in rainfall monitoring*. Academic press, 1981. ISBN: 9780120796809.

[22] Battan, L. *Radar Observation of the Atmosphere*. University of Chicago Press: Chicago, IL, USA, 1973. ISBN: 9781878907271.

[23] Beran, M., J. R. M. Hosking, and N. Arnell. "Comment on Two-Component Extreme Value Distribution for Flood Frequency Analysis by Fabio Rossi, Mauro Florentino, and Pasquale Versace". In: *Water Resources Research* 22.2 (1986), pp. 263–266.

[24] Bergeron, T. "The problem of tropical hurricanes". In: *Quarterly Journal of the Royal Meteorological Society* 80.344 (1954), pp. 131–164.

[25] Bernard, M. M. "Formulas for rainfall intensities of long duration". In: *Transactions of the American Society of Civil Engineers* 96.1 (1932), pp. 592–606.

[26] Berne, A. and W. F. Krajewski. "Radar for hydrology: Unfulfilled promise or unrecognized potential?" In: *Advances in Water Resources* 51 (2013), pp. 357–366.

[27] Bindlish, R., T. J. Jackson, E. Wood, H. Gao, P. Starks, D. Bosch, and V. Lakshmi. "Soil moisture estimates from TRMM Microwave Imager observations over the Southern United States". In: *Remote Sensing of Environment* 85.4 (2003), pp. 507–515.

[28] Blöschl, G., A. Viglione, and A. Montanari. "Emerging Approaches to Hydrological Risk Management in a Changing World". In: *Climate Vulnerability: Understanding and Addressing Threats to Essential Resources*. Ed. by Pielke Sr, R. A. Oxford: Elsevier Inc., Academic Press, 2013, pp. 3–10. ISBN: 9780123847034.

[29] Blöschl, G. and E. Zehe. "On hydrological predictability". In: *Hydrological processes* 19.19 (2005), pp. 3923–3929.

[30] Boni, G., A. Parodi, and R. Rudari. "Extreme rainfall events: Learning from raingauge time series". In: *Journal of hydrology* 327.3 (2006), pp. 304–314.

[31] Borga, M., E. N. Anagnostou, and E. Frank. "On the use of real-time radar rainfall estimates for flood prediction in mountainous basins". In: *Journal of Geophysical Research: Atmospheres* 105.D2 (2000), pp. 2269–2280.

[32] Brandes, E. A. "Optimizing rainfall estimates with the aid of radar". In: *Journal of Applied Meteorology* 14.7 (1975), pp. 1339–1345.

[33] Bringi, V. and V. Chandrasekar. *Polarimetric Doppler weather radar: principles and applications*. Cambridge University Press, 2001. ISBN: 9780521019552.

[34] Bringi, V., V. Chandrasekar, J. Hubbert, E. Gorgucci, W. Randeu, and M. Schoenhuber. "Raindrop size distribution in different climatic regimes from disdrometer and dual-polarized radar analysis". In: *Journal of the Atmospheric Sciences* 60.2 (2003), pp. 354–365.

[35] Bruen, M. and F. O'Loughlin. "Towards a nonlinear radar-gauge adjustment of radar via a piece-wise method". In: *Meteorological Applications* 21.3 (2014), pp. 675–683.

[36] Brunetti, M., M. Maugeri, and T. Nanni. "Variations of temperature and precipitation in Italy from 1866 to 1995". In: *Theoretical and Applied Climatology* 65.3-4 (2000), pp. 165–174.

[37] Brunetti, M., M. Maugeri, F. Monti, and T. Nanni. "Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series". In: *International journal of climatology* 26.3 (2006), pp. 345–381.

[38] Buffoni, L., M. Maugeri, and T. Nanni. "Precipitation in Italy from 1833 to 1996". In: *Theoretical and Applied Climatology* 63.1-2 (1999), pp. 33–40.

[39] Buishand, T. "Extreme rainfall estimation by combining data from several sites". In: *Hydrological Sciences Journal* 36.4 (1991), pp. 345–365.

[40] Cai, Y., C. Jin, A. Wang, D. Guan, J. Wu, F. Yuan, and L. Xu. "Spatio-Temporal Analysis of the Accuracy of Tropical Multisatellite Precipitation Analysis 3B42 Precipitation Data in Mid-High Latitudes of China". In: *PLOS ONE* 10.4 (2015), pp. 1–22.

[41] Caracciolo, C., F. Porcu, and F. Prodi. "Precipitation classification at mid-latitudes in terms of drop size distribution parameters". In: *Advances in Geosciences* 16 (2008), pp. 11–17.

[42] Cassola, F., F. Ferrari, and A. Mazzino. "Numerical simulations of Mediterranean heavy precipitation events with the WRF model: A verification exercise using different approaches". In: *Atmospheric Research* 164 (2015), pp. 210–225.

[43] Centro Funzionale di Protezione Civile Prov.Aut.Trento. *Meteotrentino*. Accessed: 2016-08-01. URL: http://www.meteotrentino.it/.

[44] Chapon, B., G. Delrieu, M. Gosset, and B. Boudevillain. "Variability of rain drop size distribution and its effect on the Z-R relationship: A case study for intense Mediterranean rainfall". In: *Atmospheric Research* 87.1 (2008), pp. 52–6.

[45] Chumchean, S., A. Sharma, and A. Seed. "An integrated approach to error correction for real-time radar-rainfall estimation". In: *Journal of Atmospheric and Oceanic Technology* 23.1 (2006), pp. 67–79.

[46] Ciach, G. J. "Local random errors in tipping-bucket rain gauge measurements". In: *Journal of Atmospheric and Oceanic Technology* 20.5 (2003), pp. 752–759.

[47] Ciach, G. J. and W. F. Krajewski. "Radar-rain gauge comparisons under observational uncertainties". In: *Journal of Applied Meteorology* 38.10 (1999), pp. 1519–1525.

[48] Ciullo, A., A. Viglione, A. Castellarin, M. Crisci, and G. Di Baldassarre. "Socio-hydrological modelling of flood-risk dynamics: comparing the resilience of green and technological systems". In: *Hydrological Sciences Journal* (2017), pp. 1–12.

[49] Claps, P., C. Barberis, M. D. Agostino, E. Gallo, G. Laguardia, F. Laio, F. Miotto, F. Plebani, G. Vezzù, A. Viglione, and M. Zanetta. "Development of an Information System of the Italian basins for the CUBIST project". In: *EGU General Assembly 2008*. 2008.

[50] Claps, P., E. Caporali, V. Chiarello, D. R., D. De Luca, L. Giuzio, A. Libertino, F. Lo Conti, S. Manfreda, V. Noto, and P. Versace. "Stima operativa delle piogge estreme sul territorio nazionale: nuovi metodi e possibili sinergie". In: *Atti del XXXV Convegno Nazionale di Idraulica e Costruzioni Idrauliche*. 2016, pp. 557–560.

[51] Clarke, R. T., R. D. de Paiva, and C. B. Uvo. "Comparison of methods for analysis of extremes when records are fragmented: A case study using Amazon basin rainfall data". In: *Journal of Hydrology* 368.1 (2009), pp. 26–29.

[52] Cleveland, W. S. and S. J. Devlin. "Locally weighted regression: an approach to regression analysis by local fitting". In: *Journal of the American statistical association* 83.403 (1988), pp. 596–610.

[53] Coe, J. A., J. A. Michael, R. A. Crovelli, W. Z. Savage, W. T. Laprade, and W. D. Nashem. "Probabilistic assessment of precipitation-triggered landslides using historical records of landslide occurrence, Seattle, Washington". In: *Environmental & Engineering Geoscience* 10.2 (2004), pp. 103–122.

[54] Coleman, T. F. and Y. Li. "An interior trust region approach for nonlinear minimization subject to bounds". In: *SIAM Journal on optimization* 6.2 (1996), pp. 418–445.

[55] Coles, S. *An introduction to statistical modeling of extreme values*. Springer Series in Statistics. London: Springer-Verlag, 2001. ISBN: 9781852334598.

[56] Cremonini, R. and R. Bechini. "Heavy rainfall monitoring by polarimetric C-Band weather radars". In: *Water* 2.4 (2010), pp. 838–848.

[57] CUBIST Team, The. *CUBIST*. Accessed: 2015-08-29. URL: http://www.cubist.polito.it.

[58] Cunnane, C. "Factors affecting choice of distribution for flood series". In: *Hydrological Sciences Journal* 30.1 (1985), pp. 25–36.

[59] Dalrymple, T. *Flood-frequency analyses, manual of hydrology: Part 3*. Technical notes. USGPO, 1960.

[60] Davini, P., R. Bechini, R. Cremonini, and C. Cassardo. "Radar-based analysis of convective storms over Northwestern Italy". In: *Atmosphere* 3.1 (2011), pp. 33–58.

[61] Dayan, U., K. Nissen, and U. Ulbrich. "precipitation over the eastern and western Mediterranean". In: *Natural Hazards and Earth System Sciences* 15 (2015), pp. 2525–2544.

[62] De Luca, C., P. Furcolo, F. Rossi, P. Villani, and C. Vitolo. "Extreme rainfall in the Mediterranean". In: *Proceedings of the International Workshop on Advances in Statistical Hydrology*. Citeseer. 2010, pp. 23–25.

[63] DeGaetano, A. T. and D. S. Wilks. "Radar-guided interpolation of climatological precipitation data". In: *International Journal of Climatology* 29.2 (2009), pp. 185–196.

[64] Deidda, R. and M. Puliga. "Sensitivity of goodness-of-fit statistics to rainfall data rounding off". In: *Physics and Chemistry of the Earth, Parts A/B/C* 31.18 (2006), pp. 1240–1251.

[65] Dinku, T., S. J. Connor, and P. Ceccato. "Comparison of CMORPH and TRMM-3B42 over Mountainous Regions of Africa and South America". In: *Satellite Rainfall Applications for Surface Hydrology*. Ed. by Gebremichael, M. and F. Hossain. Springer Netherlands, 2010, pp. 193–204.

[66] Dobesch, H., P. Dumolard, and I. Dyras. *Spatial interpolation for climate data: the use of GIS in climatology and meteorology*. John Wiley & Sons, 2013. ISBN: 9781280847837.

[67] Doviak, R. J. and D. S. Zrnic. *Doppler Radar and Weather Observations*. 2nd ed. Dover Publication: Dover, Uk, 2006. ISBN: 9780486450605.

[68] Easterling, D. R., G. A. Meehl, C. Parmesan, S. A. Changnon, T. R. Karl, and L. O. Mearns. "Climate extremes: observations, modeling, and impacts". In: *Science* 289.5487 (2000), pp. 2068–2074.

[69] Elachi, C. and J. J. Van Zyl. *Introduction to the physics and techniques of remote sensing*. Vol. 28. John Wiley & Sons, 2006. ISBN: 9780471848103.

[70] Elshorbagy, A. A., U. Panu, and S. Simonovic. "Group-based estimation of missing hydrological data: I. Approach and general methodology". In: *Hydrological Sciences Journal* 45.6 (2000), pp. 849–866.

[71] Elshorbagy, A., S. Simonovic, and U. Panu. "Estimation of missing streamflow data using principles of chaos theory". In: *Journal of Hydrology* 255.1 (2002), pp. 123–133.

[72] Ferrari, E. "Regional rainfall and flood frequency analysis in Italy". In: *International Conference Development in Hydrology of Mountainous Areas, Stará Lesná, Slovakia*. 1994, pp. 12–16.

[73] Fiori, E., A. Comellas, L. Molini, N. Rebora, F. Siccardi, D. Gochis, S. Tanelli, and A. Parodi. "Analysis and hindcast simulations of an extreme rainfall event in the Mediterranean area: The Genoa 2011 case". In: *Atmospheric Research* 138 (2014), pp. 13–29.

[74] Fleming, K., J. Awange, M. Kuhn, and W. Featherstone. "Evaluating the TRMM 3B43 monthly precipitation product using gridded raingauge data over Australia". In: *Australian Meteorological and Oceanographic Journal* 61.3 (2011), pp. 171–184.

[75] Flocas, H. A., I. Simmonds, J. Kouroutzoglou, K. Keay, M. Hatzaki, V. Bricolas, and D. Asimakopoulos. "On cyclonic tracks over the eastern Mediterranean". In: *Journal of Climate* 23.19 (2010), pp. 5243–5257.

[76] Furcolo, P., A. Pelosi, and F. Rossi. "Statistical identification of orographic effects in the regional analysis of extreme rainfall". In: *Hydrological Processes* (2015), pp. 1342–1353.

[77] Gabriele, S. and F. Chiaravalloti. "Searching regional rainfall homogeneity using atmospheric fields". In: *Advances in Water Resources* 53 (2013), pp. 163–174.

[78] Gelfand, A. E., P. Diggle, P. Guttorp, and M. Fuentes. *Handbook of spatial statistics*. CRC press, 2010. ISBN: 9781420072877.

[79] Gilbert, R. O. *Statistical methods for environmental pollution monitoring*. John Wiley & Sons, 1987. ISBN: 9780470310045.

[80] Goudenhoofdt, E. and L. Delobbe. "Evaluation of radar-gauge merging methods for quantitative precipitation estimates". In: *Hydrology and Earth System Sciences* 13.2 (2009), pp. 195–203.

[81] Grimaldi, S., S. Kao, A. Castellarin, S. Papalexiou, A. Viglione, F. Laio, H. Aksoy, and A. Gedikli. "Statistical Hydrology". In: *Treatise on Water Science*. Ed. by Wilderer, P. Oxford: Elsevier, 2011, pp. 479–511. ISBN: 9780444531995.

[82] Grubbs, F. E. "Procedures for detecting outlying observations in samples". In: *Technometrics* 11.1 (1969), pp. 1–21.

[83] Habib, E., W. F. Krajewski, and A. Kruger. "Sampling errors of tipping-bucket rain gauge measurements". In: *Journal of Hydrologic Engineering* 6.2 (2001), pp. 159–166.

[84] Hamed, K. H. and A. R. Rao. "A modified Mann-Kendall trend test for autocorrelated data". In: *Journal of Hydrology* 204.1-4 (1998), pp. 182–196.

[85] Heinrich, U. *Zur Methodik der räumlichen Interpolation mit geostatistischen Verfahren*. Springer, 1992. ISBN: 9783824420278.

[86] Higgins, R., J. E. Schemm, W. Shi, and A. Leetmaa. "Extreme precipitation events in the western United States related to tropical forcing". In: *Journal of climate* 13.4 (2000), pp. 793–820.

[87] Hosking, J. R. M. *Some theoretical results concerning L-moments*. Research report. New York, US: Watson IBM Research Center, 1989.

[88] Hosking, J. R. M. "L-Moments: analysis and estimation of distributions using linear combinations of order statistics". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 52.1 (1990), pp. 105–124.

[89] Hosking, J. R. M. and J. F. Wallis. "Parameter and Quantile Estimation for the Generalized Pareto Distribution". In: *Technometrics* 29.3 (1987), pp. 339–349.

[90] Hosking, J. R. M. and J. R. Wallis. *Regional frequency analysis: an approach based on L-moments*. Cambridge University Press, 1997. ISBN: 9780511881640.

[91] Hou, A. Y., R. K. Kakar, S. Neeck, A. A. Azarbarzin, C. D. Kummerow, M. Kojima, R. Oki, K. Nakamura, and T. Iguchi. "The Global Precipitation Measurement Mission". In: *Bulletin of the American Meteorological Society* 95.5 (2014), pp. 701–722.

[92] Houze Jr, R. A. "Stratiform precipitation in regions of convection: A meteorological paradox?" In: *Bulletin of the American Meteorological Society* 78.10 (1997), pp. 2179–2196.

[93] Houze, R. A. and P. V. Hobbs. "Organization and structure of precipitating cloud systems". In: *Advances in Geophysics* 24 (1982), pp. 225–315.

[94] Hov, Ø., U. Cubasch, E. Fischer, P. Höppe, T. Iversen, N. Gunnar Kvamstø, W. Kundzewicz, D. Rezacova, D. Rios, F. Duarte Santos, et al. *Extreme weather events in Europe: Preparing for climate change adaptation*. Final report. Oslo, NO: Norwegian Meteorological Institute, 2013.

[95] Huffman, G. J., D. T. Bolvin, and E. J. Nelkin. *Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation*. Technical Documentation. Washington, DC 20546: NASA, 2014.

[96] Huffman, G. J., D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, and E. F. Stocker. "The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales". In: *Journal of Hydrometeorology* 8.1 (2007), pp. 38–55.

[97] Iguchi, T., T. Kozu, J. Kwiatkowski, R. Meneghini, J. Awaka, and K. Okamoto. "Uncertainties in the Rain Profiling Algorithm for the TRMM Precipitation Radar". In: *Journal of the Meteorological Society of Japan* 87A (2009), pp. 1–30.

[98] Isaaks, E. H. and M. R. Srivastava. *An Introduction to Applied Geostatistics*. Oxford University Press, USA, 1990. ISBN: 9780195050134.

[99] ISPRA. *Hydrological Yearbooks*. Accessed: 2017-01-22. URL: http://www.isprambiente.gov.it/en/projects/inland-waters-and-marine-waters/hydrological-yearbooks?set_language=en.

[100] ISPRA. *Inquadramento storico del monitoraggio idro-meteografico e delle relative competenze*. Accessed: 2017-01-22. URL: http://www.isprambiente.gov.it/it/progetti/acque-interne-e-marino-costiere-1/progetto-annali/inquadramento-storico-del-monitoraggio-idro-meteografico-e-delle-relative-competenze.

[101] Jenkinson, A. F. "The frequency distribution of the annual maximum (or minimum) values of meteorological elements". In: *Quarterly Journal of the Royal Meteorological Society* 81.348 (1955), pp. 158–171.

[102] Joss, J. and A. Waldwogel. "A method to improve the accuracy of radar-measured amounts of precipitation." In: *In Proceedings of 14th Conference of Radar Meteorology, Tucson, AZ, USA*. 1970, pp. 237–238.

[103] Journel, A. G. and C. J. Huijbregts. *Mining geostatistics*. Academic press, 1978. ISBN: 9780123910509.

[104] Kendall, M. *Rank correlation methods*. Oxford University Press, 1990. ISBN: 9780195208375.

[105] Khan, S. I., P. Adhikari, Y. Hong, H. Vergara, R. F. Adler, F. Policelli, D. Irwin, T. Korme, and L. Okello. "Hydroclimatology of Lake Victoria region using hydrologic model and satellite remote sensing data". In: *Hydrology and Earth System Sciences* 15.1 (2011), pp. 107–117.

[106] Kidder, S. Q. and T. H. V. Haar. *Satellite meteorology: an introduction*. Gulf Professional Publishing, 1995. ISBN: 9780124064300.

[107] Klüppelberg, C. "Subexponential distributions and integrated tails". In: *Journal of Applied Probability* 25.01 (1988), pp. 132–141.

[108] Klüppelberg, C. "Subexponential distributions and characterizations of related classes". In: *Probability Theory and Related Fields* 82.2 (1989), pp. 259–269.

[109] Kolarova, M. and P. Hamouz. "Approaches to the analysis of weed distribution". In: *Atlas of Weed Mapping*. John Wiley and Sons, Ltd, 2016, pp. 438–453. ISBN: 9781118720691.

[110] Kouroutzoglou, J., H. Flocas, K. Keay, I. Simmonds, and M. Hatzaki. "Climatological aspects of explosive cyclones in the Mediterranean". In: *International Journal of Climatology* 31.12 (2011), pp. 1785–1802.

[111] Koutsoyiannis, D. and A. Langousis. "Precipitation". In: *Treatise on Water Science*. Ed. by Wilderer, P. Oxford: Elsevier, 2011, pp. 27–77. ISBN: 9780444531995.

[112] Koutsoyiannis, D. "Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation". In: *Hydrological Sciences Journal* 49.4 (2004), pp. 575–590.

[113] Koutsoyiannis, D. "A critical review of probability of extreme rainfall: principles and models". In: *Advances in Urban Flood Management* (2007), pp. 139–166.

[114] Koutsoyiannis, D., D. Kozonis, and A. Manetas. "A mathematical framework for studying rainfall intensity-duration-frequency relationships". In: *Journal of Hydrology* 206.1 (1998), pp. 118–135.

[115] Koutsoyiannis, D., A. Montanari, H. F. Lins, and T. A. Cohn. "Climate, hydrology and freshwater: towards an interactive incorporation of hydrological experience into climate research". In: *Hydrological Sciences Journal* 54.2 (2009), pp. 394–405.

[116] Krajewski, W. and J. Smith. "Radar hydrology: rainfall estimation". In: *Advances in Water Resources* 25.8 (2002), pp. 1387–1394.

[117] Krajewski, W. F., G. Villarini, and J. A. Smith. "RADAR-rainfall uncertainties where are we after thirty years of effort?" In: *Bulletin of the American Meteorological Society* 91 (2010), pp. 87–94.

[118] Kummerow, C., J. Simpson, O. Thiele, W. Barnes, A. Chang, E. Stocker, R. Adler, A. Hou, R. Kakar, F. Wentz, et al. "The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit". In: *Journal of Applied Meteorology* 39.12 (2000), pp. 1965–1982.

[119] Lee, G. W. and I. Zawadzki. "Variability of drop size distributions: Time-scale dependence of the variability and its effects on rain estimation". In: *Journal of Applied Meteorology* 44.2 (2005), pp. 241–255.

[120] Legates, D. R. "Real-time calibration of radar precipitation estimates". In: *The Professional Geographerogr.* 52.2 (2000), pp. 235–246.

[121] Li, H., J. Sheffield, and E. F. Wood. "Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching". In: *Journal of Geophysical Research: Atmospheres* 115.D10 (2010), pp. 1–20.

[122] Libertino, A., S. Macchia, and P. Claps. "Nubifragi eccezionali in Italia: analisi preliminare di rilevanza sugli eventi 1920-2000". In: *Atti del XXXV Convegno Nazionale di Idraulica e Costruzioni Idrauliche*. 2016, pp. 549–552.

[123] Libertino, A., P. Allamano, P. Claps, R. Cremonini, and F. Laio. "Radar estimation of intense rainfall rates through adaptive calibration of the ZR relation". In: *Atmosphere* 6.10 (2015), pp. 1559–1577.

[124] Libertino, A., A. Sharma, V. Lakshmi, and P. Claps. "A global assessment of the timing of extreme rainfall from TRMM and GPM for improving hydrologic design". In: *Environmental Research Letters* 11.5 (2016), pp. 1–9.

[125] Lin, Y.-L. *Mesoscale dynamics*. Cambridge University Press, 2007. ISBN: 9780521808750.

[126] Liu, Z. "Comparison of versions 6 and 7 3-hourly TRMM multi-satellite precipitation analysis (TMPA) research products". In: *Atmospheric Research* 163 (2015), pp. 91–101.

[127] Llasat-Botija, M., M. Llasat, and L. López. "Natural hazards and the press in the western Mediterranean region". In: *Advances in Geosciences* 12 (2007), pp. 81–85.

[128] Lo Conti, F., K.-L. Hsu, L. Noto, and S. Sorooshian. "Evaluation and comparison of satellite precipitation estimates with reference to a local area in the Mediterranean Sea". In: *Atmospheric Research* 138 (2014), pp. 189–204.

[129] Ly, S., C. Charles, and A. Degré. "Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale. A review". In: *Biotechnologie, Agronomie, Société et Environnement* 17.2 (2013), pp. 392–406.

[130] Manfreda, S., A. Sole, and G. De Costanzo. *Le precipitazioni estreme in Basilicata*. Editrice Universo Sud, 2015. ISBN: 9788899432034.

[131] Mann, H. "Nonparametric Tests against Trend". In: *Econmetrica* 13.3 (1945), pp. 245–259.

[132] Marshall, J. S. and W. M. K. Palmer. "The distribution of raindrops with size". In: *Journal of Meteorology* 5.4 (1948), pp. 165–166.

[133] Mathevet, T. and R. Garçon. "Tall tales from the hydrological crypt: are models monsters?" In: *Hydrological Sciences Journal–Journal des Sciences Hydrologiques* 55.6 (2010), pp. 857–871.

[134] Menne, M. J., I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston. "An Overview of the Global Historical Climatology Network-Daily Database". In: *Journal of Atmospheric and Oceanic Technology* 29.7 (2012), pp. 897–910.

[135] Menne, M., I. Durre, B. Korzeniewski, S. McNeal, K. Thomas, X. Yin, S. Anthony, R. Ray, R. Vose, B. Gleason, et al. *Global Historical Climatology Network-Daily (GHCN-Daily), Version 3. 22*. NOAA National Climatic Data Center. Accessed: 2015-08-01. 2012.

[136] Messeri, A., M. Morabito, G. Messeri, G. Brandani, M. Petralli, F. Natali, D. Grifoni, A. Crisci, G. Gensini, and S. Orlandini. "Weather-related flood and landslide damage: a risk index for Italian regions". In: *PloS one* 10.12 (2015), pp. 1–17.

[137] Ministero dell'Ambiente, della Tutela del Territorio e del Mare. *Il rischio idrogeologico in Italia*. Institutional report. Roma, IT, 2008.

[138] Mugnai, A., S. Michele, E. Smith, F. Baordo, P. Bauer, B. Bizzarri, P. Joe, C. Kidd, F. Marzano, A. Tassa, J. Testud, and G. Tripoli. "Snowfall Measurements by Proposed European GPM Mission". In: *Measuring Precipitation From Space*. Ed. by Levizzani, V., P. Bauer, and F. Turk. Vol. 28. Advances In Global Change Research. Springer Netherlands, 2007, pp. 655–674. ISBN: 9781402058349.

[139] Munich RE. *NatCatSERVICE*. Accessed: 2017-04-01. URL: http://natcatservice.munichre.com/.

[140] Myers, D. E. "Spatial interpolation: an overview". In: *Geoderma* 62.1 (1994), pp. 17–28.

[141] Nanding, N., M. A. Rico-Ramirez, and D. Han. "Comparison of different radar-raingauge rainfall merging techniques". In: *Journal of Hydroinformatics* 17.3 (2015), pp. 422–445.

[142] NASA LP DAAC. *ASTER GDEM V2. ASTER GDEM is a product of NASA and METI. NASA EOSDIS Land Processes DAAC, USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota (https://lpdaac.usgs.gov), 2011. Accessed: 2016-10-23*. URL: http://gdex.cr.usgs.gov/gdex/.

[143] Ochoa, A., L. Pineda, P. Willems, and P. Crespo. "Evaluation of TRMM 3B42 (TMPA) precipitation estimates and WRF retrospective precipitation simulation over the Pacific-Andean basin into Ecuador and Peru". In: *Hydrology and Earth System Sciences Discussions* 11.1 (2014), pp. 411–449.

[144] Olea, R. A. "Geostatistics for engineers and earth scientists". In: *Technometrics* 42.4 (2000), pp. 444–445.

[145] Oliver, M. A. and R. Webster. "Kriging: a method of interpolation for geographical information systems". In: *International Journal of Geographical Information System* 4.3 (1990), pp. 313–332.

[146] Papalexiou, S. M., D. Koutsoyiannis, and C. Makropoulos. "How extreme is extreme? An assessment of daily rainfall distribution tails". In: *Hydrology and Earth System Sciences* 17.2 (2013), pp. 851–862.

[147] Papalexiou, S. M. and D. Koutsoyiannis. "Battle of extreme value distributions: A global survey on extreme daily rainfall". In: *Water Resources Research* 49.1 (2013), pp. 187–201.

[148] Pappas, C., S. M. Papalexiou, and D. Koutsoyiannis. "A quick gap filling of missing hydrometeorological data". In: *Journal of Geophysical Research: Atmospheres* 119.15 (2014), pp. 9290–9300.

[149] Pathak, C. and R. Teegavarapu. "Utility of optimal reflectivity-rain rate (Z-R) relationships for improved precipitation estimates." In: *In Proceedings of the World Environmental and Water Resources Congress*, pp. 4681–4691.

[150] Peel, M. C. and G. Blöschl. "Hydrological modelling in a changing world". In: *Progress in Physical Geography* 35.2 (2011), pp. 249–261.

[151] Pereira, F. and J. Augusto. “Improving WSR-88D hourly rainfall estimates”. In: *Weather Forecast.* 13.4 (1998), pp. 1016–1028.

[152] Pettersen, S. *Weather analysis and forecasting: Volume II: Weather and weather systems*. 1956. ISBN: 9780070496866.

[153] Phillips, D. L., J. Dolph, and D. Marks. “A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain”. In: *Agricultural and Forest Meteorology* 58.1 (1992), pp. 119–141.

[154] Pilgrim, D. *Australian Rainfall and Runoff: A Guide to Flood Estimation*. Institution of Engineers Australia, 1987. ISBN: 9780858254344.

[155] Protezione Civile Puglia. *Annali Idrologici - Parte I*. Accessed: 2016-08-01. URL: http://www.protezionecivile.puglia.it/centro-funzionale/analisielaborazione-dati.

[156] Protezione Civile Regione Marche. *Annali Idrologici Regione Marche*. Accessed: 2016-08-01. URL: http://console.protezionecivile.marche.it/viewdoc.asp?co_id=511.

[157] Prudhomme, C. and D. W. Reed. “Mapping extreme rainfall in a mountainous region using geostatistical techniques: a case study in Scotland”. In: *International Journal of Climatology* 19.12 (1999), pp. 1337–1356.

[158] Raghavan, S. *Radar Metorology*. Kluwer Academic Publisher: Dordrecht, The Netherlands, 2003. ISBN: 1402016042.

[159] Rasmussen, K. L., S. L. Choi, M. D. Zuluaga, and R. A. Houze. “TRMM precipitation bias in extreme storms in South America”. In: *Geophysical Research Letters* 40.13 (2013), pp. 3457–3461.

[160] Reale, O. and R. Atlas. “Tropical cyclone–like vortices in the extratropics: Observational evidence and synoptic analysis”. In: *Weather and forecasting* 16.1 (2001), pp. 7–34.

[161] Rebora, N., L. Molini, E. Casella, A. Comellas, E. Fiori, F. Pignone, F. Siccardi, F. Silvestro, S. Tanelli, and A. Parodi. “Extreme rainfall in the mediterranean: what can we learn from observations?” In: *Journal of Hydrometeorology* 14.3 (2013), pp. 906–922.

[162] Reed, D. *Flood Estimation Handbook: Overview*. Vol. 1. Institute of Hydrology Wallingford, 1999. ISBN: 9781906698003.

[163] Reiss, R.-D. and M. Thomas. *Statistical analysis of extreme values, from insurance, finance, hydrology and other field*. Basel, Boston, Berlin: Birkhaüser Verlag, 2001. ISBN: 9783764357689.

[164] Rendon, S., B. Vieux, and C. Pathak. “Estimation of regionally specific Z-R relationships for radar-based hydrologic prediction.” In: *In Proceedings of the World Environmental and Water Resources Congress*. 2010, pp. 4668–4680.

[165] Rendon, S., B. Vieux, and C. Pathak. “Continuous forecasting and evaluation of derived Z-R relationships in a sparse rain gauge network using NEXRAD”. In: *Journal of Hydrologic Engineering* 18.2 (2013), pp. 175–182.

[166] Ripley, B. D. *Spatial statistics*. Vol. 575. John Wiley & Sons, 2005. ISBN: 9780471083672.

[167] Romilly, T. G. and M. Gebremichael. "Evaluation of satellite rainfall estimates over Ethiopian river basins". In: *Hydrology and Earth System Sciences* 15.5 (2011), pp. 1505–1514.

[168] Rossi, F., G. Scannapieco, and P. Villani. "Una proposta operativa per la rivalutazione del rischio idrologico di alluvione in Italia". In: *Atti del XXXV Convegno Nazionale di Idraulica e Costruzioni Idrauliche*. 2016, pp. 529–532.

[169] Rossi, F., M. Fiorentino, and P. Versace. "Two-component extreme value distribution for flood frequency analysis". In: *Water Resources Research* 20.7 (1984), pp. 847–856.

[170] Rudari, R., D. Entekhabi, and G. Roth. "Large-scale atmospheric patterns associated with mesoscale features leading to extreme precipitation events in Northwestern Italy". In: *Advances in Water Resources* 28.6 (2005), pp. 601–614.

[171] Rutter, C. M. "An Introduction to Applied Geostatistics". In: *Journal of the American Statistical Association* 86.414 (1991), pp. 548–550.

[172] Sanders, F. and J. R. Gyakum. "Synoptic-dynamic climatology of the "bomb"". In: *Monthly Weather Review* 108.10 (1980), pp. 1589–1606.

[173] Seo, D., J. Breidenbach, and E. Johnson. "Real-time estimation of mean field bias in radar rainfall data". In: *Journal of Hydrology* 223.3 (1999), pp. 131–147.

[174] Seo, D.-J. and J. Breidenbach. "Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements". In: *Journal of Hydrometeorology* 3.2 (2002), pp. 93–111.

[175] Siccardi, F. "Rainstorm hazards and related disasters in the North-West Mediterranean region". In: *Remote Sensing Reviews* 14.1-3 (1996), pp. 5–21.

[176] Singh, R., T. Wagener, K. van Werkhoven, M. E. Mann, and R. Crane. "A trading-space-for-time approach to probabilistic continuous streamflow predictions in a changing climate accounting for changing watershed behavior". In: *Hydrology and Earth System Sciences* 15.11 (2011), pp. 3591–3603.

[177] Singh, R., T. Wagener, K. van Werkhoven, M. E. Mann, and R. Crane. "A trading-space-for-time approach to probabilistic continuous streamflow predictions in a changing climate – accounting for changing watershed behavior". In: *Hydrology and Earth System Sciences* 15.11 (2011), pp. 3591–3603.

[178] SIR Toscana. *Settore Idrologico Regionale*. Accessed: 2016-08-01. URL: http://www.sir.toscana.it/.

[179] Smith, J. A. and W. F. Krajewski. "A modeling study of rainfall rate-reflectivity relationships". In: *Water Resour. Res.* 29.8 (1993), pp. 2505–2514.

[180] Smith, R. L. "Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone". In: *Statistical Science* (1989), pp. 367–377.

[181] Soliman, M. M. *Engineering hydrology of arid and semi-arid regions*. CRC Press, 2010. ISBN: 9781439815557.

[182] Stalmann, V., R. Draschoff, T. Günther, A. Pfister, D. Prellberg, H. Verworn, and G. Malitz. "Das Niederschlagsregelwerk für die deutsche Wasserwirtschaft". In: *Wasserwirtschaft* 94.10 (2004), pp. 8–27.

[183] Stephens, G. L. *Remote sensing of the lower atmosphere*. Oxford University Press New York, 1994. ISBN: 9780195081886.

[184] Su, F., Y. Hong, and D. P. Lettenmaier. "Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and Its Utility in Hydrologic Prediction in the La Plata Basin". In: *Journal of Hydrometeorology* 9.4 (2008), pp. 622–640.

[185] Svensson, C. and D. A. Jones. "Review of rainfall frequency estimation methods". In: *Journal of Flood Risk Management* 3.4 (2010), pp. 296–313.

[186] Szolgay, J., J. Parajka, S. Kohnová, and K. Hlavčová. "Comparison of mapping approaches of design annual maximum daily precipitation". In: *Atmospheric Research* 92.3 (2009), pp. 289–307.

[187] Tachikawa, T., M. Hato, M. Kaku, and A. Iwasaki. "Characteristics of ASTER GDEM version 2". In: *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*. IEEE. 2011, pp. 3657–3660.

[188] Taleb, N. N. *The black swan: The impact of the highly improbable*. Random house, 2007. ISBN: 9780812973815.

[189] Teegavarapu, R. S. *Floods in a changing climate: extreme precipitation*. Cambridge University Press, 2012. ISBN: 9781107018761.

[190] Teegavarapu, R. S. "Spatial interpolation using nonlinear mathematical programming models for estimation of missing precipitation records". In: *Hydrological Sciences Journal* 57.3 (2012), pp. 383–406.

[191] Teegavarapu, R. S. "Statistical corrections of spatially interpolated missing precipitation data estimates". In: *Hydrological Processes* 28.11 (2014), pp. 3789–3808.

[192] Testik, F. Y. and A. P. Barros. "Toward elucidating the microstructure of warm rainfall: A survey". In: *Reviews of Geophysics* 45.2 (2007), pp. 1–21.

[193] Teugels, J. L. "The class of subexponential distributions". In: *The Annals of Probability* (1975), pp. 1000–1011.

[194] Tian, Y. and C. D. Peters-Lidard. "A global map of uncertainties in satellite-based precipitation measurements". In: *Geophysical Research Letters* 37.24 (2010), pp. 1–6.

[195] Todini, E. "A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements". In: *Hydrology and Earth System Sciences* 5.2 (2001), pp. 187–199.

[196] Tokay, A. and D. A. Short. "Evidence from tropical raindrop spectra of the origin of rain from stratiform versus convective clouds". In: *Journal of Applied Meteorology* 35.3 (1996), pp. 355–371.

[197] Tous, M. and R. Romero. "Medicanes: cataloguing criteria and exploration of meteorological environments". In: *Tethys* 8 (2011), pp. 53–61.

[198] Tous, M. and R. Romero. "Meteorological environments associated with medicane development". In: *International Journal of Climatology* 33.1 (2013), pp. 1–14.

[199] Trigo, I. F., T. D. Davies, and G. R. Bigg. "Objective climatology of cyclones in the Mediterranean region". In: *Journal of Climate* 12.6 (1999), pp. 1685–1696.

[200] Uboldi, F., A. Sulis, C. Lussana, M. Cislaghi, and M. Russo. "A spatial bootstrap technique for parameter estimation of rainfall annual maxima distribution". In: *Hydrology and Earth System Sciences* 18.3 (2014), pp. 981–995.

[201] Uijlenhoet, R. *Parameterization of rainfall microstructure for radar meteorology and hydrology*. Wageningen Univ., 1999. ISBN: 9789058081568.

[202] Ulbrich, U., G. Leckebusch, and J. G. Pinto. "Extra-tropical cyclones in the present and future climate: a review". In: *Theoretical and Applied Climatology* 96.1-2 (2009), pp. 117–131.

[203] Velasco-Forero, C. A., D. Sempere-Torres, E. F. Cassiraga, and J. Jaime Gómez-Hernández. "A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data". In: *Advances in Water Resources* 32.7 (2009), pp. 986–1002.

[204] Vieux, B. E. and S. H. Rendon. *Derivation and evaluation of seasonally specific Z-R relationships.* Final report. West Palm Beach, FL: South Florida Water Management District, 2009.

[205] Villarini, G., P. V. Mandapaka, W. F. Krajewski, and R. J. Moore. "Rainfall and sampling uncertainties: A rain gauge perspective". In: *Journal of Geophysical Research: Atmospheres* 113.D11 (2008), pp. 1–12.

[206] Vrieling, A., J. C. Hoedjes, and M. van der Velde. "Towards large-scale monitoring of soil erosion in Africa: Accounting for the dynamics of rainfall erosivity". In: *Global and Planetary Change* 115 (2014), pp. 33–43.

[207] Wang, G., L. Liu, and Y. Ding. "Improvement of radar quantitative precipitation estimation based on real-time adjustments to Z-R relationships and inverse distance weighting correction schemes". In: *Advances in Atmospheric Sciences* 29.3 (2012), pp. 575–584.

[208] Ward, E., W. Buytaert, L. Peaver, and H. Wheater. "Evaluation of precipitation products over complex mountainous terrain: A water resources perspective". In: *Advances in Water Resources* 34.10 (2011), pp. 1222–1231.

[209] Waymire, E., V. K. Gupta, and I. Rodriguez-Iturbe. "A spectral theory of rainfall intensity at the meso-β scale". In: *Water Resources Research* 20.10 (1984), pp. 1453–1465.

[210] Westra, S., H. Fowler, J. Evans, L. Alexander, P. Berg, F. Johnson, E. Kendon, G. Lenderink, and N. Roberts. "Future changes to the intensity and frequency of short-duration extreme rainfall". In: *Reviews of Geophysics* 52.3 (2014), pp. 522–555.

[211] Westra, S., L. V. Alexander, and F. W. Zwiers. "Global increasing trends in annual maximum daily precipitation". In: *Journal of Climate* 26.11 (2013), pp. 3904–3918.

[212] Wilks, D. S. *Statistical methods in the atmospheric sciences*. Vol. 100. Academic press, 2011. ISBN: 9780127519661.

[213] Wood, S., D. Jones, R. Moore, et al. "Static and dynamic calibration of radar data for hydrological use". In: *Hydrology and Earth System Sciences Discussions* 4.4 (2000), pp. 545–554.

[214] Yufa, W., W. Cuihong, and J. Hongxiang. "Real-time synchronous integration of radar and raingauge measurements based on the quasi same-rain-volume sampling". In: *Acta Meteorologica Sinica* 24.3 (2010), pp. 340–353.

[215] Zulkafli, Z., W. Buytaert, C. Onof, B. Manz, E. Tarnavsky, W. Lavado, and J.-L. Guyot. "A Comparative Performance Analysis of TRMM 3B42 (TMPA) Versions 6 and 7 for Hydrological Applications over Andean–Amazon River Basins". In: *Journal of Hydrometeorology* 15.2 (2014), pp. 581–592.

Appendix A

Additional results of the “patched kriging” methodology

A.1 Variograms

This section reports the sample and theoretical variograms obtained for the individual years for the 1-3-6-12 and 24 hours extremes in Piemonte.

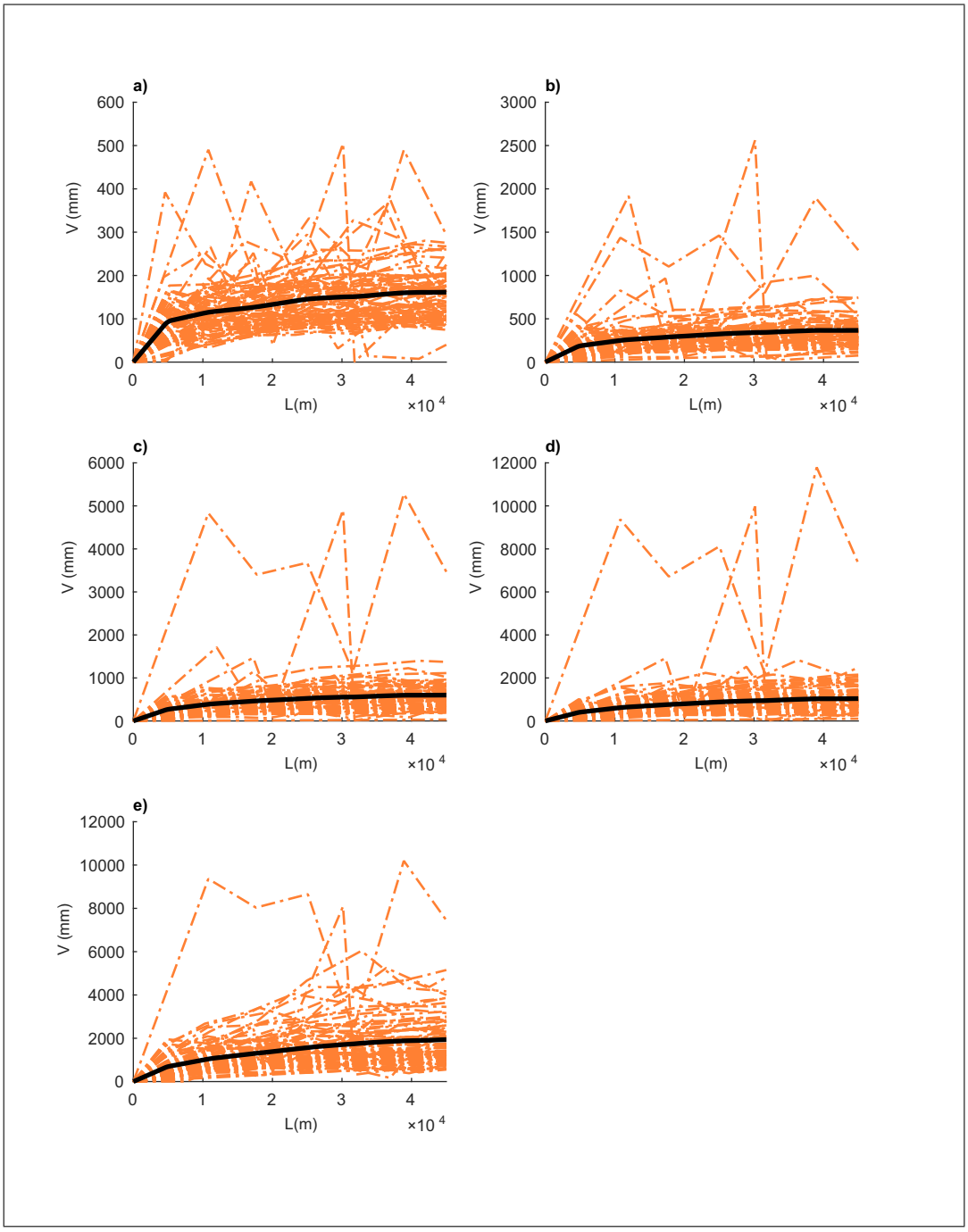


Fig. A.1 Sample variograms for the 1 (a) to 24 (e) hours durations. The orange dashed lines refer to the annual sample variograms, the black curve is the average sample variogram for the considered duration.

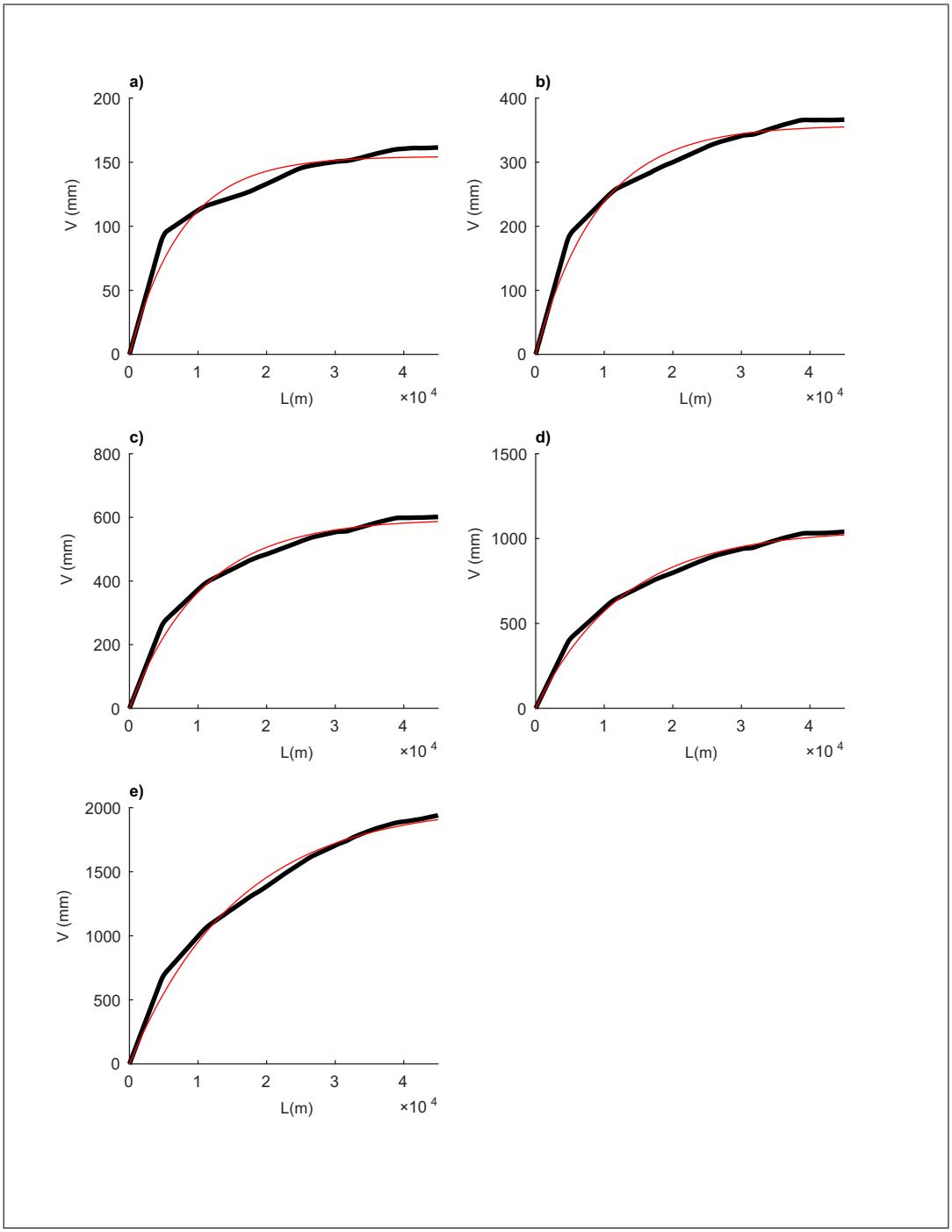


Fig. A.2 Theoretical variograms for the 1 (a) to 24 (e) hours durations. The black curve is the average sample variogram for the duration, the red one the fitted exponential variogram.

A.2 L-moments

This section reports the weighted L-moments maps for the 1-3-6-12 and 24 hours rainfall extremes.

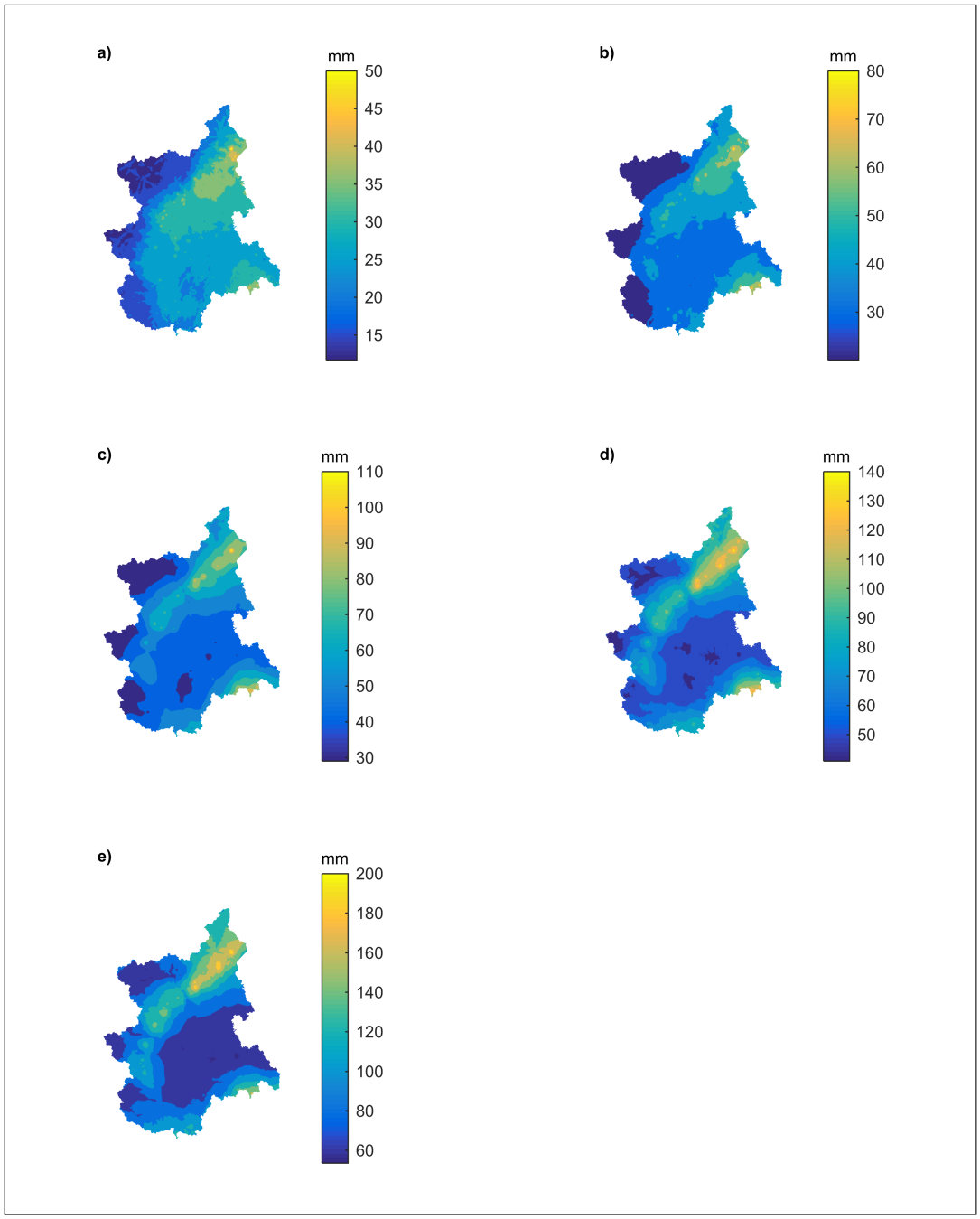


Fig. A.3 Maps of l_1 for the 1 (a) to 24 (e) hours durations.

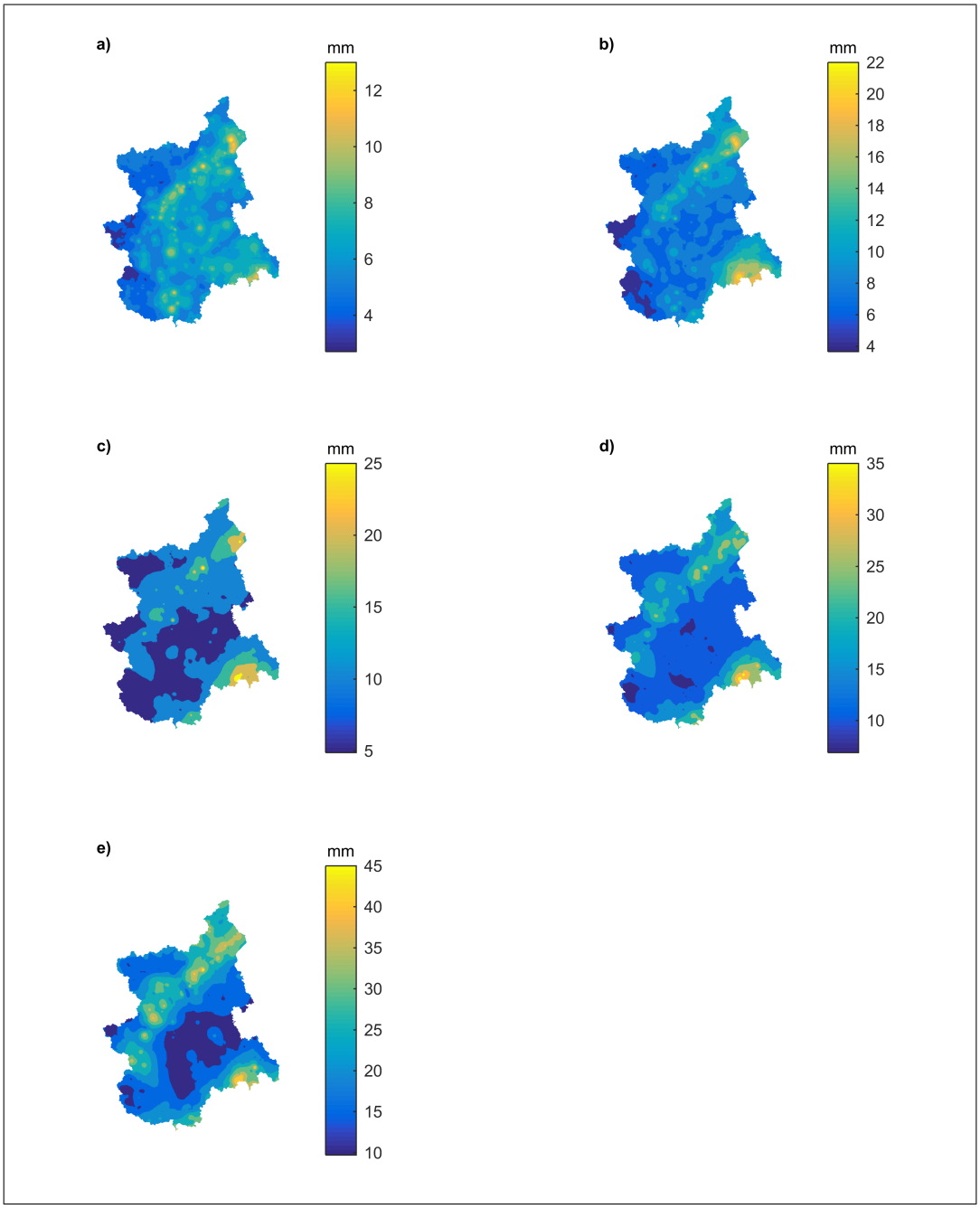


Fig. A.4 Maps of l_2 for the 1 (a) to 24 (e) hours durations.

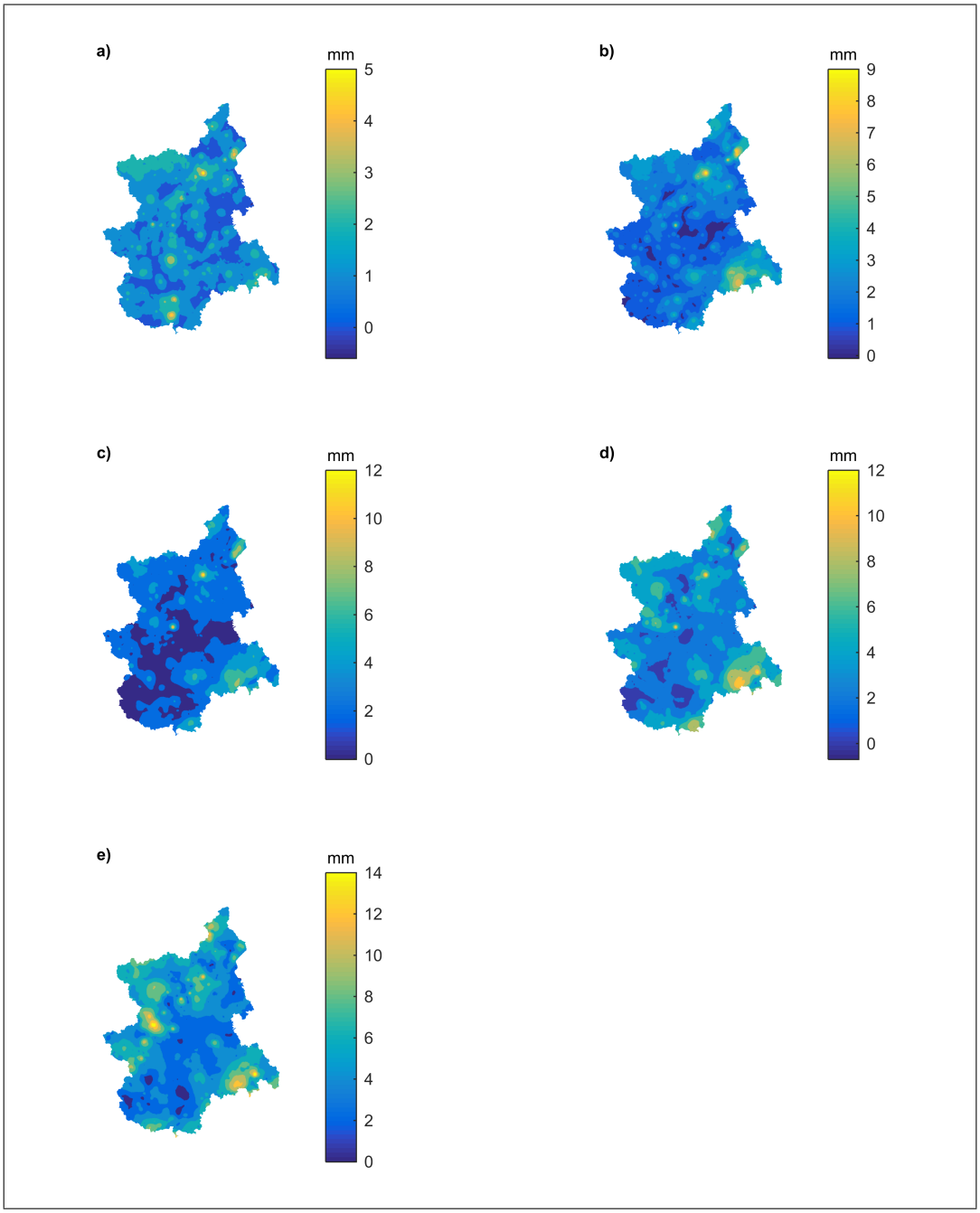


Fig. A.5 Maps of l_3 for the 1 (a) to 24 (e) hours durations.

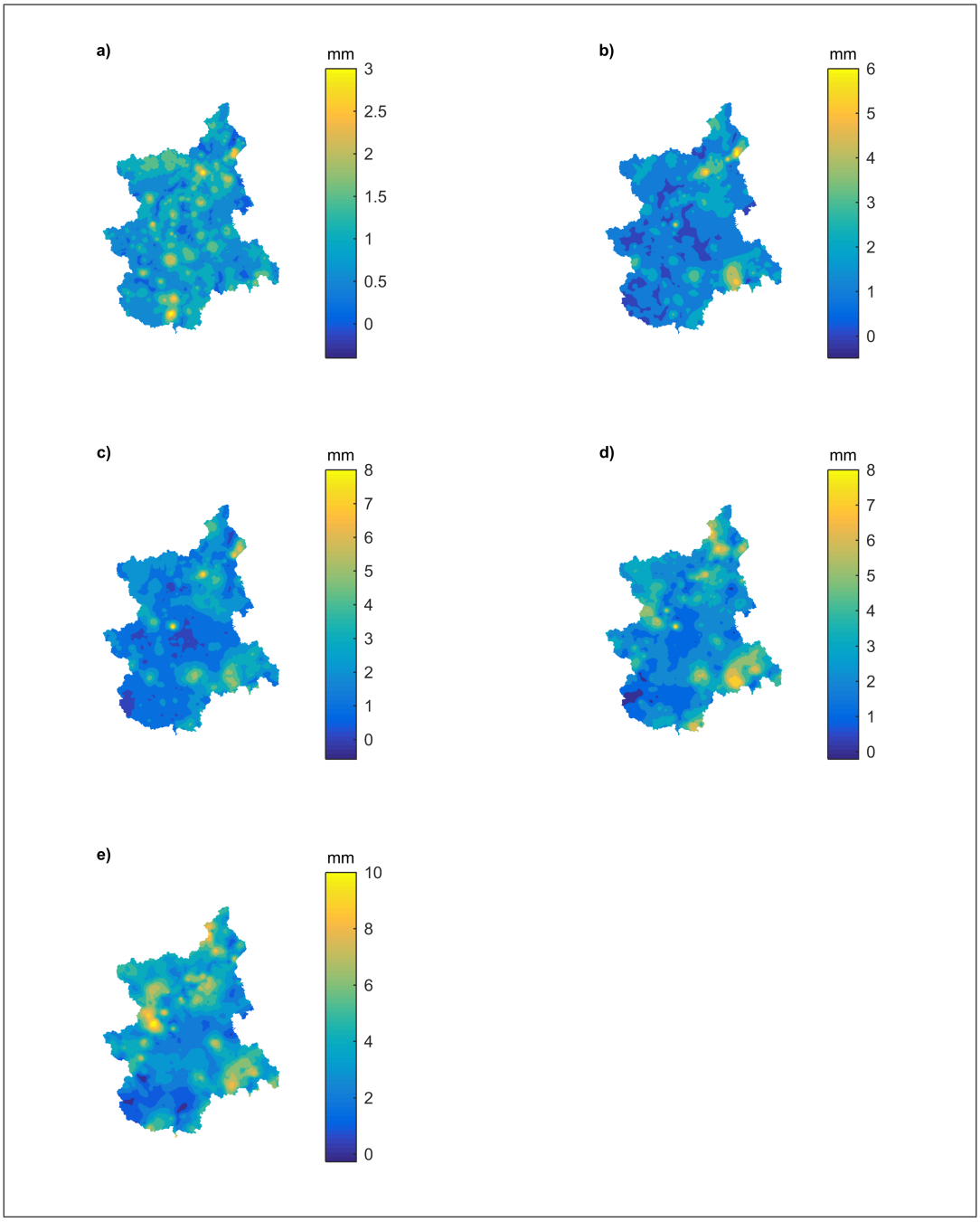


Fig. A.6 Maps of l_4 for the 1 (a) to 24 (e) hours durations.

Appendix B

Description of the rainfall datasets composing the Rainfall Extremes Database

B.1 The National Hydrographic Service dataset

The National Hydrographic Service (*SIN*) and the National Hydrographic and Mareographic Service (*SIMN*) collected annual maxima values for 1-3-6-12 and 24 hours duration in the Hydrological Yearbooks from 1917 to about 2002 (before some local structures were already operating, e.g., in Venezia and Parma). The management of the network and the publication of the Yearbooks was delegated from the Central Administration to local Departmental Offices. Each Departmental Office was supervising a certain number of basins at the national scale. The Italian territory was therefore divided in Compartments. All the Compartments should have provided their data to the Central Service in Roma.

The borders of the different compartments changed frequently during the life of the *SIN* and the *SIMN*. At the beginning, 10 compartments were defined (Venezia, Parma, Pisa, Roma, Napoli, Catanzaro, Chieti, Bologna, Palermo, Cagliari). After many changes, with suppressions and unifications resulting also from the administrative changes in the national configuration, the Decree 85/1991 stated the definitive structure of the *SIMN*. The Central Administration was located in Roma, the 10 Compartmental Offices were located in Venezia, Parma, Bologna, Pescara, Bari, Catanzaro, Napoli, Roma, Pisa e Genova. Bolzano, Trento, Cagliari and Palermo hosted independent Hydrographic Services, coordinated by the *SIMN*. The definitive configuration is shown in figure B.1. For more information on the evolution of the service, refer to [100].

The *ISPRA*'s "Yearbook project" involved the digitization of all data published in the Hydrological Yearbooks since 1921 in order to create a national database. This project is still ongoing and just a PDF archive has been published, and available in [99]. However, for the aim of this work a digital archive is needed.

The "*CUBIST* project", established in 2006, was funded by the Italian Ministry of Education and Research (PRIN 2005) and belongs to the "Prediction in Ungauged Basins" research area. The project was supported by four university partners (Politecnico di Torino, Univesità degli Studi di Bologna, Università degli Studi della Basilicata, Politecnico di Bari) and CNR-ISAFoM (a public research organization). Under the *CUBIST* project a comprehensive nationwide hydrological information system has been set up. The database includes about 6000 pluviografic and plu-

Fig. B.1 The structure of the *SIMN* under the Decree 85/1991 (source: [100]).

viometric stations, 700 temperature stations and about 400 river basins, completely defined as regards their geomorphological characteristics and climatic and hydrologic parameters [49].

The annual maxima for different durations, digitized from the Hydrological Yearbooks were included in the *CUBIST* database [57]. In the detail, the database includes data from 1900 to 2001. Data from 1900 to 1916 are from pre-*SIN* local datasets. The different Local Compartments provided data over different periods, as the shift from the Central Hydrological Systems to the Regional ones was not simultaneous all over the nation. Therefore the length of the different databases are not uniform.

The number of data per year is not constant across the analysed period as shown in Figure B.2, that shows the minimum number of data per year across the 5 available duration. The number of data increases with time, as more station has been installed in the recent years. The data availability deceases in the period of the Second World War, as many records have been missed in the period. After 1980, with the progressive dismissal of the *SIMN*, and the development of the local hydrograpic authorities the data availability decreases rapidly until 2001, when the rain gauges still under the ex-*SIMN* were taken over by the Operational Centres.

The available series are classified according to their length. Results are shown in Figure B.3. Although a large amount of series is shorter than 10 years and therefore unsuitable for a robust statistical analysis, the great abundance of available stations can grant a significant number of series with more than 20-30 data, suitable for our analysis. The spatial distribution of the stations is shown in figure B.4. The color scale refers to the number of the available data per each series. The minimum number across the 5 duration is considered.

B.2 The Local Operational Centers datasets

After the late '80s the local Environmental Agencies started to support the *SIMN* in its work. Gradually, the 21 regional hydrological services took over the tasks and the networks of the national one. Each hydrological service adopted its own rules for the management and dissemination of the collected data and, even if the

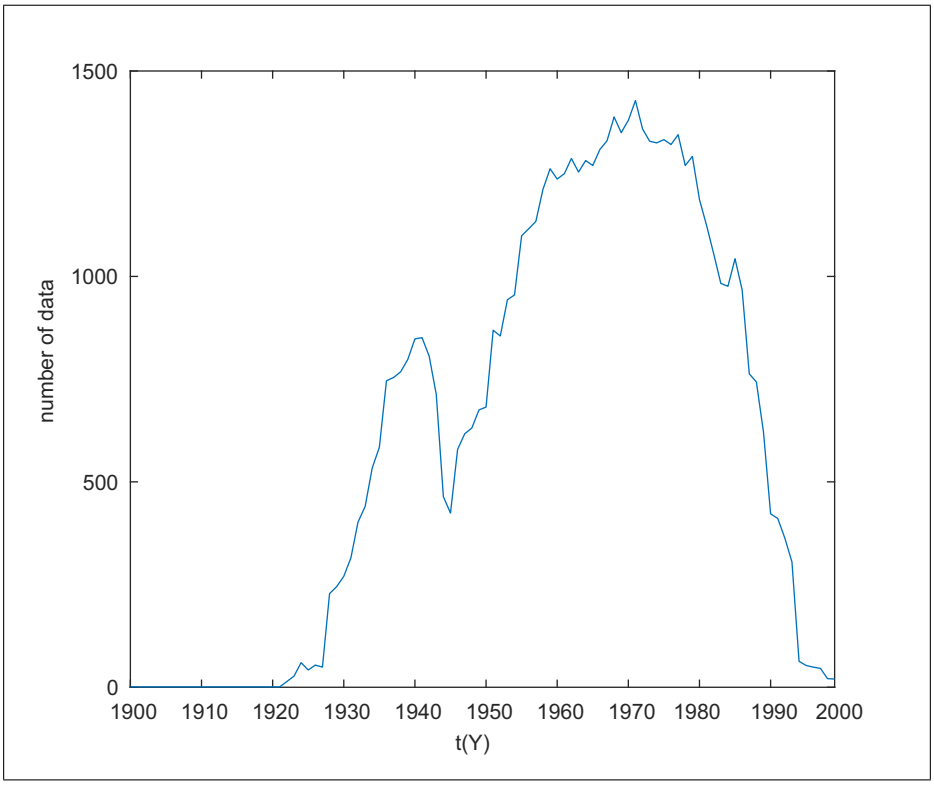


Fig. B.2 Data availability in the *CUBIST* database (the smallest value across the 5 considered durations is considered per each year).

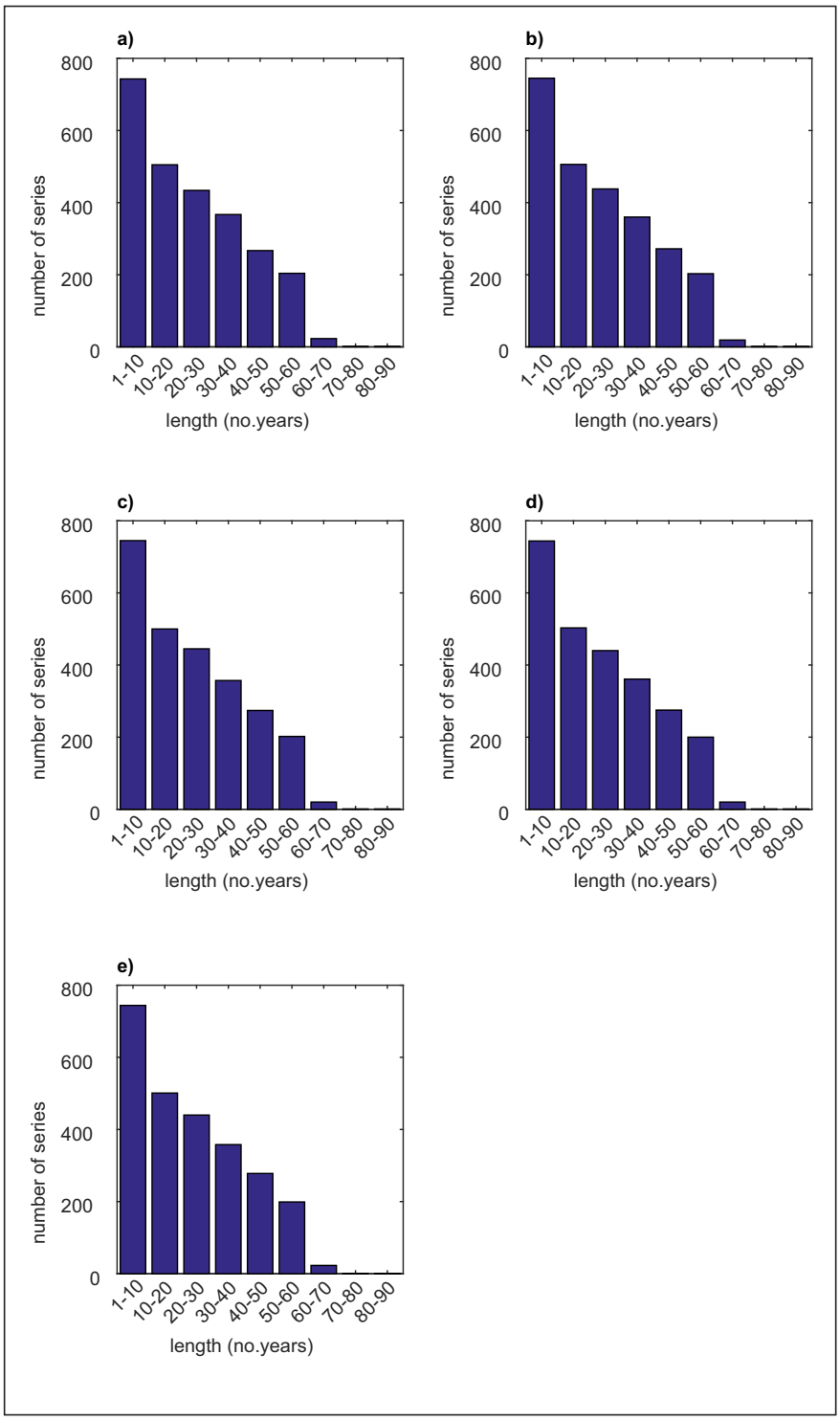


Fig. B.3 Number of station per record length class in the *CUBIST* database for durations from (a) 1 to (e) 24 hours.

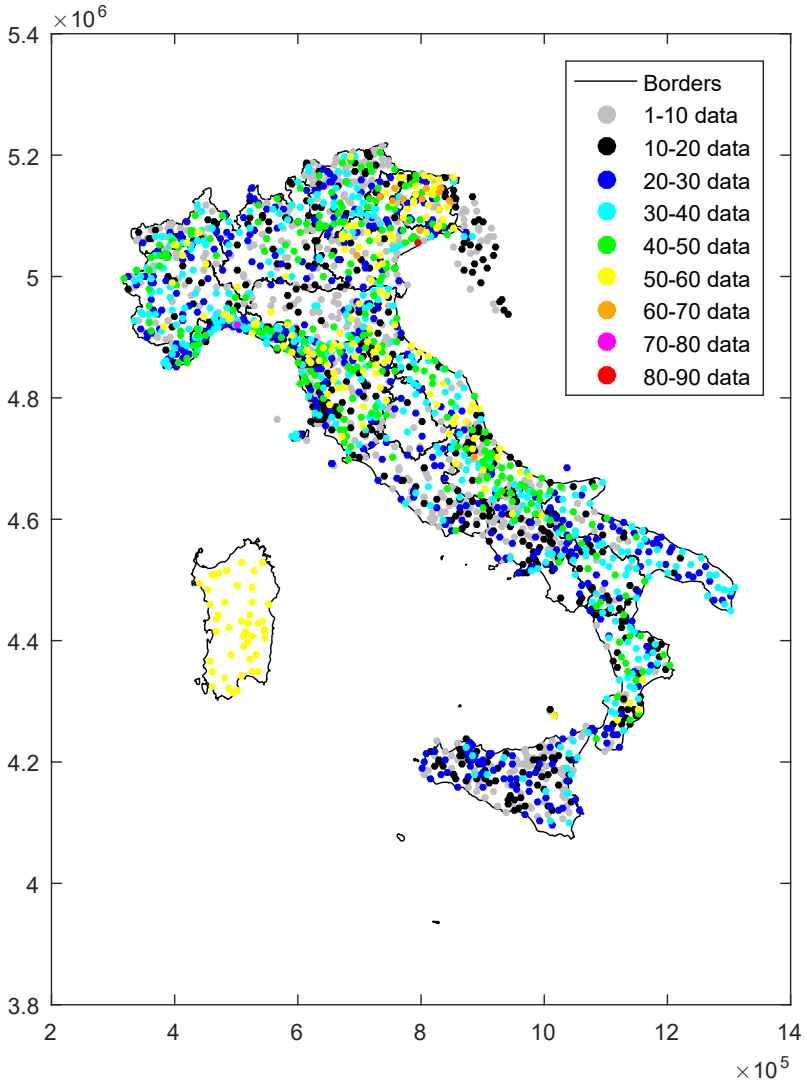


Fig. B.4 Length of the series in the *CUBIST* database represented in space. The color refers to the minimum length among the 5 available durations. If more stations overlap due to the resolution of the picture, the one with the longer series appears on top.

Italian law adopted an Open Source policy for the public data, nowadays, an updated database of the annual rainfall maxima for sub-daily duration at the national scale is still lacking. The different agencies have been therefore contacted, requesting their annual maxima dataset for sub-daily duration. For simplicity, the regions of Italy are listed in table B.1 with the name of the local authorities and the codes, aimed at identifying them in the following. A map showing the location of the regions is shown in figure B.5.

Merging and homogenizing the different dataset has been a quite long and difficult operation. The different Operational Centres provided different type of dataset, with different temporal coverages and spatial reference systems. Duplicate stations are often present in the databases of neighboring region. Some of the datasets also include a complete or partial digitized version of the data from the *SIMN* network (but often the name and code of the *SIMN* stations has not been preserved). A complete analysis for a comparison between the considered databases in their original form is not feasible due to their different characteristics.

B.3 Cleaning and merging the datasets

In this section the steps followed for the definition of the *RED* are described. The first steps of the work have been carried out at the regional scale. For each region all the available data are considered. In the detail for each region 3 types of data could be available:

1. Data from the *CUBIST* database for the 1900-2001 period
2. Data provided from the regional authority
3. Data provided from the regional authorities of the neighboring regions, falling out of their regional borders

Considering that most of the provided dataset have been validated from the related authorities, they are considered reliable and, at first, included directly in the *RED*. In the presence of inconsistencies between the type 2 and type 3 data, preliminary manual merging is carried out. If the same station falls in both the database of the considered region and in the one of the neighboring one/ones, a first

Table B.1 Regions of Italy with the assigned code and the related local Operational Center with references to the availability of digitized data. * the autonomous provinces of Trento and Bolzano/Bolzen, together, constitute the region Trentino Alto Adige (CD: 22).

CD	Region	Operational Center	Digitized data availability
01	Abruzzo	Ufficio Idrografico e Mareografico Regione Abruzzo	Under request
02	Basilicata	Dipartimento Protezione Civile Regione Basilicata	Available in [130]
03	Calabria	Centro Funzionale Multirischi - ARPACAL	Available at [15]
04	Campania	Centro Funzionale Regione Campania	Under request
05	Emilia-Romagna	ARPA Emilia-Romagna	Under request
06	Friuli Venezia Giulia	Ufficio Idrografico Regione Autonoma Friuli Venezia Giulia	Under request
07	Lazio	Centro Funzionale Regione Lazio	Under request
08	Liguria	ARPAL-CFMI-PC	Partially available at [17] and in [16]
09	Lombardia	ARPA Lombardia	Available at [12]
10	Marche	Dipartimento di Protezione Civile Regione Marche	Available at [156]
11	Molise	Centro Funzionale Regione Molise	Under request
12	Piemonte	ARPA Piemonte	Partially available at [14]
13	Puglia	Dipartimento di Protezione Civile Regione Puglia	Available at [155]
14	Sardegna	ARPAS	Under request
15	Sicilia	Osservatorio delle Acque Regione Siciliana	Under request
16	Toscana	Servizio Irdrografico Regionale Toscana	Available at [178]
17	Trento *	Centro Funzionale Provincia Autonoma di Trento	Available at [43]
18	Bolzano/Bolzen *	Ufficio Idrografico Provincia Autonoma di Bolzano	Under request
19	Umbria	Regione Umbria	Under request
20	Valle d'Aosta	Centro Funzionale Regione Autonoma Valle d'Aosta	Under request
21	Veneto	ARPAV	Under request

attempt of merging the series together is carried out, by analysing the data recorded year by year. If the merging is not feasible, higher priority is given to the data provided by the authority of the considered region (that is usually also the owner of the network) and the other data are discarded. This allow to avoid the presence of duplicate series in the *RED*.

Once merged, for each region, type 2 and type 3 datasets, the resulting dataset has to be merged with the type 1 dataset. This operation has been quite complex, as the overlapping period between the different dataset is different for each region and because most of the authorities did not tracked the change in the name/code of the stations. Different operations have to be made according to the type of the dataset that the region has provided. The different procedures performed according to the type of data provided by the regions (as summarized in figure B.5) are summarized in the following paragraphs.

CASE1: Regions providing a complete and merged dataset This is the case of the regions: Abruzzo, Basilicata, Calabria, Friuli Venezia Giulia, Lombardia, Marche, Molise, Puglia, Sardegna, Sicilia, Trentino Alto Adige. These regions digitized the whole *SIMN* database for their regional domains and provide a complete merged *SIMN*-post*SIMN* database. The provided data are inserted in the *RED* without editing and without considering the *CUBIST* series. Only for the Abruzzo and Molise regions some preliminary work was needed. The two regions were divided in 1963, and the databases of the two regions partially overlap. The stations are then divided according to the actual regional boundaries and the duplicate series cleaned.

CASE2: Regions providing an incomplete and merged dataset Regions Liguria, Umbria, Emilia-Romagna and Lazio provided datasets including data from their actual regional network partially merged with subsets of digitized data from the *SIMN* Hydrological Yearbooks. Not all the *SIMN* datasets have been digitized from the local authorities, therefore the dataset lacked a part of the stations included in the *CUBIST* database. To maximize the available information, data from the regional databases and the *CUBIST* one are manually analyzed and merged, in order to avoid duplicate values. Stations are merged together in the presence of a sufficient number of overlapping year. If it is not possible to unravel any doubt, the stations

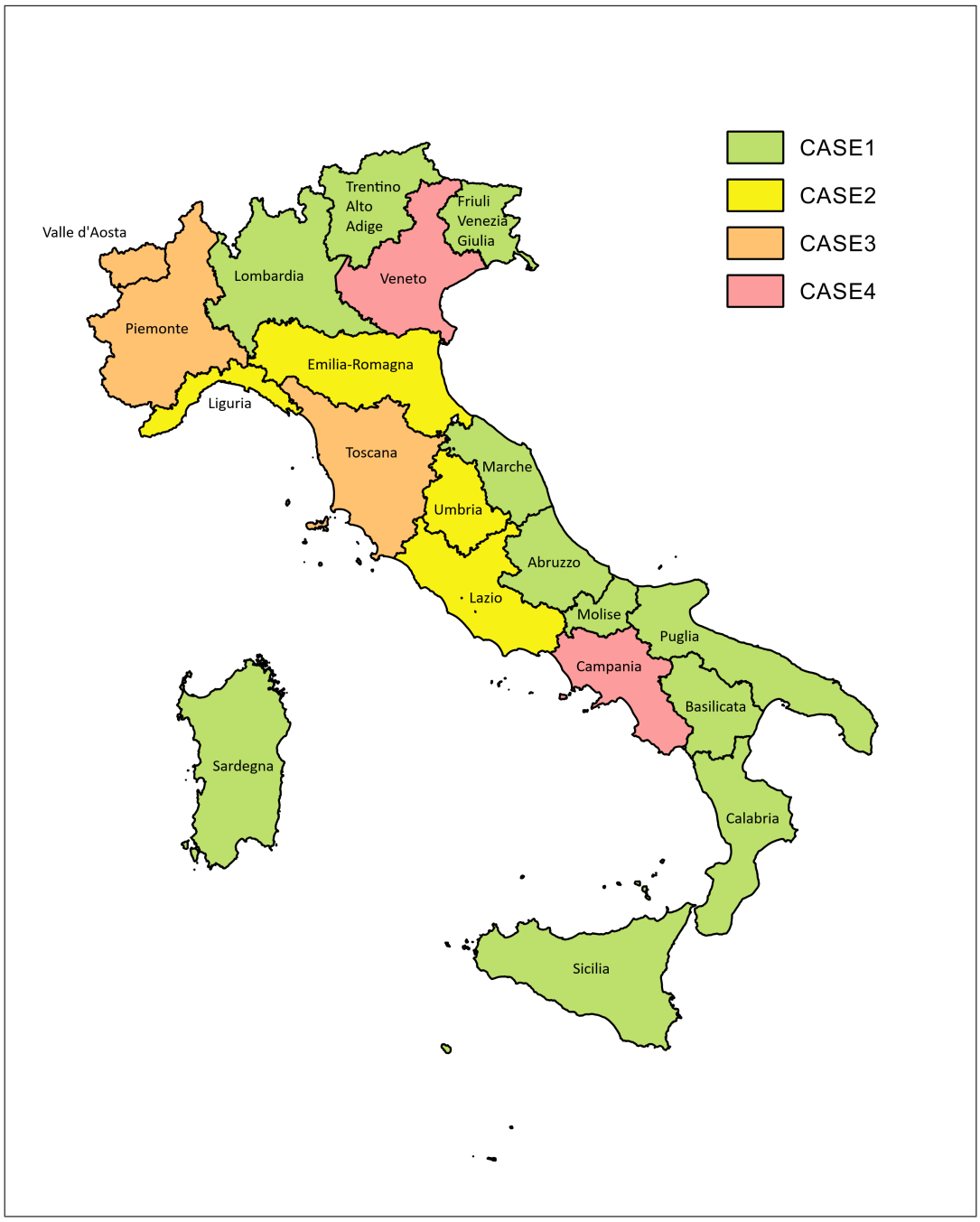


Fig. B.5 Names of the Italian regions and type of datasets provided by the regional authorities.

are considered as separate entities. For the Liguria region, the information in [16] is used to to overcome the lack of information on the continuity of the series.

CASE3: Regions providing two complete and separated datasets The Piemonte, Valle d'Aosta and Toscana regions provided two different dataset: one containing the whole digitized data from the *SIMN* stations and another one containing the digitized data from their actual networks. The data of the two database are merged together to avoid overlapping and the *CUBIST* database ignored. The operation has been feasible thanks of our collaboration with *ARPA* Piemonte, for Piemonte and Valle d'Aosta, and with the Università degli Studi di Firenze, for Toscana.

CASE4: Regions providing just the recent dataset Regions Veneto and Campania provided just the post*SIMN* datasets, including just the data recorded from the network they actually manage. All the information concerning the *SIMN* stations is lacking. The provided dataset is therefore merged with the whole *CUBIST* database for the considered region. Duplicate values are excluded analyzing manually the overlapping period, if present.

20 complete regional dataset are then obtained. The regional datasets are merged together to generate the *RED*.

Before going on with the analysis, some last reliability checks are made on the larger values of the dataset, to identify the presence of anomalous values. For each duration, the larger recorded values are considered. If any value exceeds the overall absolute maximum for the considered duration, obtained from the literature and reported in table B.2, the series is automatically excluded from the dataset. The year of occurrence of the other outliers, for each duration, is further compared, when referring to recent years, with the data from event reports or newspapers. If the data refers to a *SIMN* station, the Hydrological Yearbooks are consulted. If no evidence can be found, in-depth analysis on the quality of the series from which the anomalous value is extracted are then made. In case of further doubts, the related authority is contacted. If it it not possible to unravel any doubt, the value is discarded.

Table B.2 Overall annual maximum rainfall for different durations in Italy.

d (hours)	Station	Year	h_d (mm)
1	Vicomorasso	2011	181.0
3	Giffone	1959	360.1
6	Brugnato	2011	472.0
12	Genova Bolzaneto	1970	717.8
24	Genova Bolzaneto	1970	948.4

Appendix C

The Italian Extreme Rainfalls Database

C.1 Data consistency

In this section the detailed consistency per year per duration is reported.

Table C.1 Number of data per year for each duration in the *RED*.

Year	1h	3h	6h	12h	24h
1916	0	0	0	0	2
1917	0	0	0	0	1
1918	0	0	0	0	3
1919	0	0	0	0	0
1920	1	1	1	1	3
1921	2	0	0	0	4
1922	1	1	1	1	1
1923	22	15	4	5	46
1924	52	48	52	62	23
1925	32	8	15	17	38
1926	49	32	42	51	82
1927	49	22	38	46	57
1928	245	240	240	239	243
1929	256	254	254	254	271
1930	276	270	269	269	284
1931	348	347	348	349	361
1932	399	394	396	386	388
1933	454	441	440	446	454
1934	551	541	539	540	556
1935	525	522	524	520	526
1936	658	653	655	657	651
1937	682	676	675	672	681
1938	698	688	689	695	691
1939	719	725	719	723	724
1940	762	762	763	766	763
1941	767	769	770	770	776
1942	725	718	723	724	724
1943	642	640	643	644	643

1944	428	424	422	425	423
1945	378	373	371	373	382
1946	553	553	554	558	562
1947	594	590	591	593	594
1948	616	609	609	609	614
1949	649	638	638	638	643
1950	666	655	655	658	659
1951	821	823	821	821	825
1952	828	820	816	821	825
1953	920	919	921	919	918
1954	936	934	938	936	941
1955	1059	1059	1060	1057	1062
1956	1074	1070	1072	1068	1074
1957	1106	1106	1106	1104	1114
1958	1188	1184	1182	1181	1187
1959	1229	1225	1224	1221	1230
1960	1213	1208	1208	1208	1212
1961	1236	1226	1230	1232	1237
1962	1275	1269	1272	1271	1272
1963	1255	1256	1256	1255	1257
1964	1288	1286	1285	1279	1287
1965	1283	1284	1284	1278	1282
1966	1331	1330	1325	1315	1328
1967	1345	1341	1340	1335	1338
1968	1410	1411	1411	1407	1413
1969	1368	1368	1367	1363	1363
1970	1385	1384	1383	1380	1384
1971	1436	1436	1436	1433	1437
1972	1449	1451	1450	1445	1452
1973	1355	1352	1353	1351	1350
1974	1322	1335	1338	1335	1335
1975	1363	1365	1366	1370	1370
1976	1316	1321	1321	1316	1319
1977	1344	1342	1344	1342	1339

1978	1278	1275	1277	1275	1275
1979	1300	1302	1304	1295	1302
1980	1321	1320	1323	1319	1320
1981	1269	1266	1266	1261	1267
1982	1236	1241	1245	1240	1245
1983	1237	1238	1241	1238	1236
1984	1249	1252	1250	1246	1239
1985	1357	1357	1361	1361	1361
1986	1328	1330	1328	1324	1331
1987	1334	1337	1339	1340	1343
1988	1329	1337	1343	1337	1341
1989	1334	1334	1333	1330	1333
1990	1365	1366	1370	1368	1372
1991	1480	1483	1484	1479	1479
1992	1629	1629	1637	1632	1631
1993	1588	1590	1596	1598	1602
1994	1669	1674	1678	1681	1686
1995	1676	1676	1680	1681	1683
1996	1836	1836	1835	1831	1834
1997	1729	1732	1733	1739	1744
1998	1653	1654	1656	1654	1654
1999	1650	1648	1651	1652	1654
2000	1764	1768	1770	1772	1774
2001	1927	1931	1930	1931	1931
2002	2125	2125	2125	2125	2124
2003	2198	2200	2198	2196	2195
2004	2277	2277	2277	2277	2280
2005	2138	2138	2138	2140	2147
2006	2050	2052	2053	2053	2053
2007	2086	2084	2084	2084	2082
2008	2122	2125	2127	2126	2125
2009	1977	1975	1975	1971	1969
2010	1926	1929	1931	1931	1929
2011	1521	1521	1524	1521	1522

2012	1165	1164	1165	1164	1164
2013	751	752	752	752	752
2014	531	531	531	531	531

C.2 L-moments of the series

The sample L-moments per each duration for each series with more than 20 years of records are reported in the following figures.

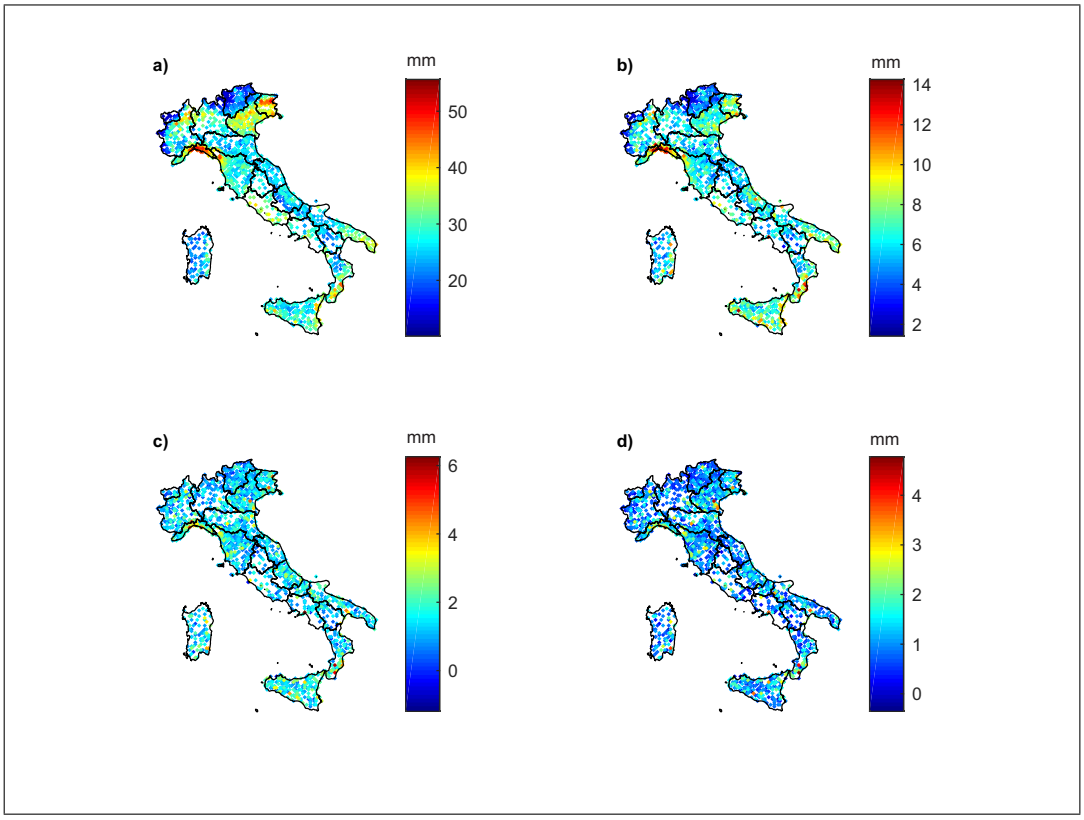


Fig. C.1 l_1 (a), l_2 (b), l_3 (c) and l_4 (d) for the 1 hour duration series.

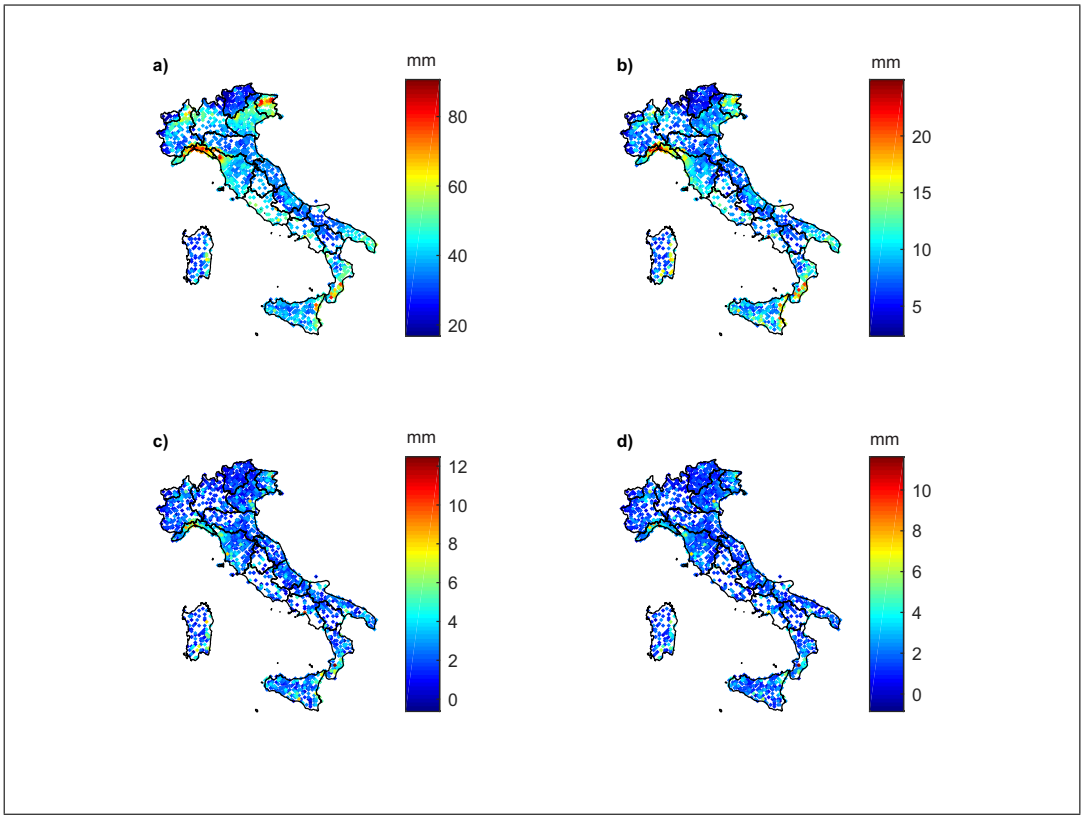


Fig. C.2 l_1 (a), l_2 (b), l_3 (c) and l_4 (d) for the 3 hours duration series.

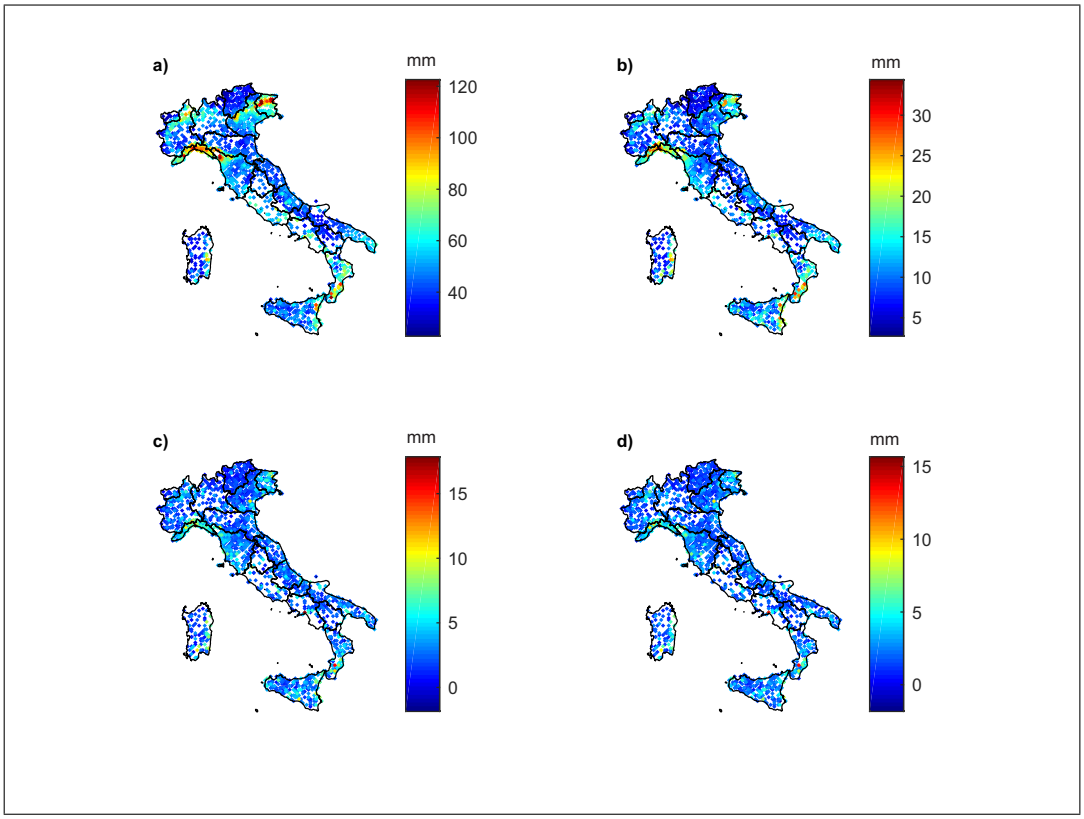


Fig. C.3 l_1 (a), l_2 (b), l_3 (c) and l_4 (d) for the 6 hours duration series.

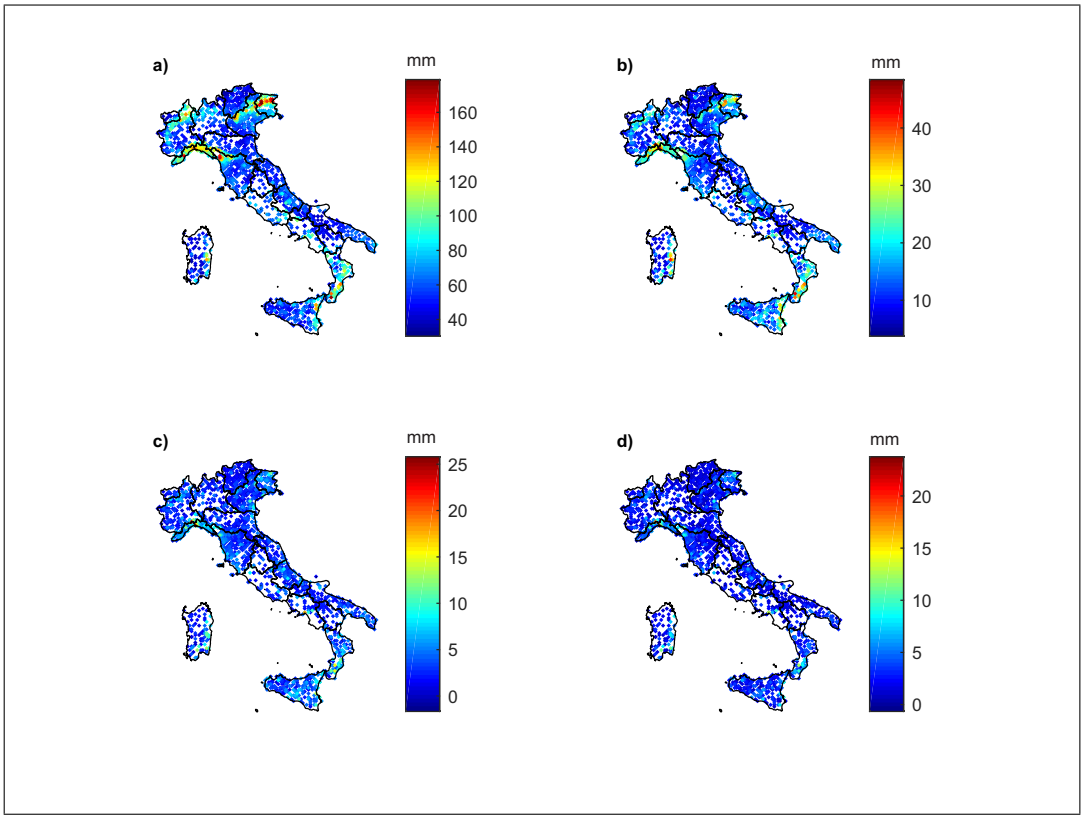


Fig. C.4 l_1 (a), l_2 (b), l_3 (c) and l_4 (d) for the 12 hours duration series.

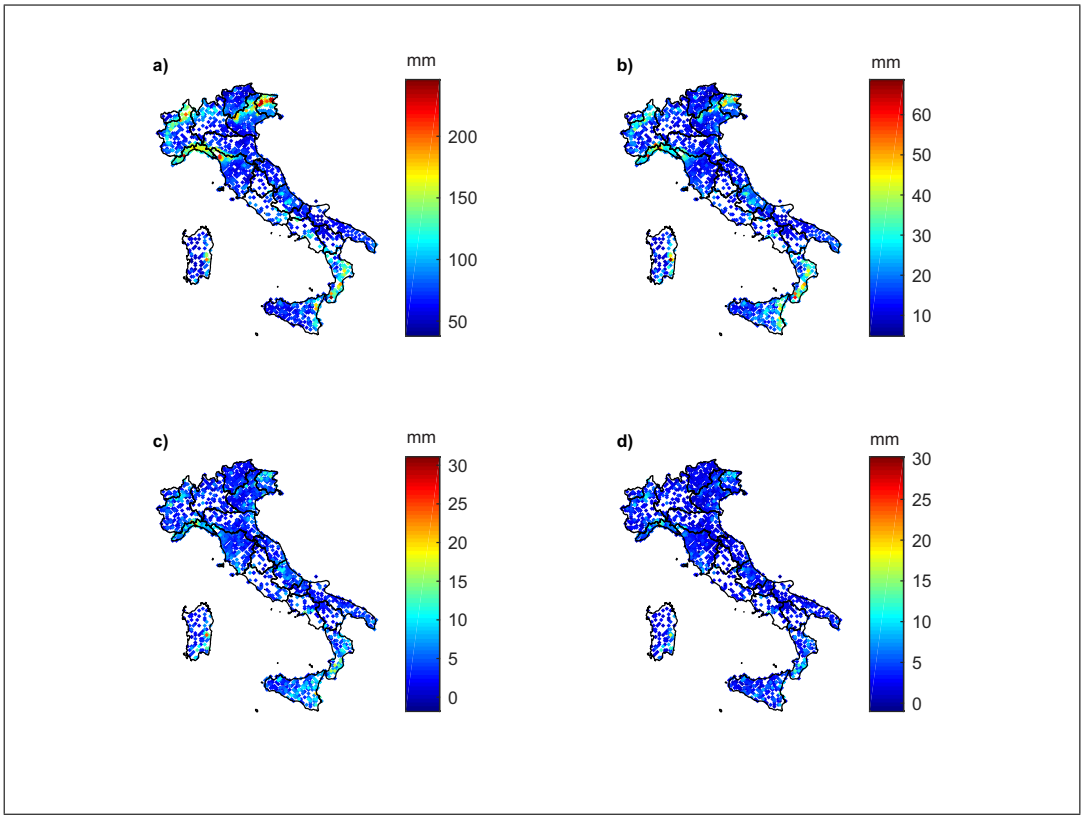


Fig. C.5 l_1 (a), l_2 (b), l_3 (c) and l_4 (d) for the 24 hours duration series.

Appendix D

Extraordinary rainfalls in Italy

D.1 Spatial distribution of the over-threshold extremes

This section reports the number of exceedances of the 0.999 empirical quantiles for the absolute and relative rainfall time series related to the 5 considered durations.

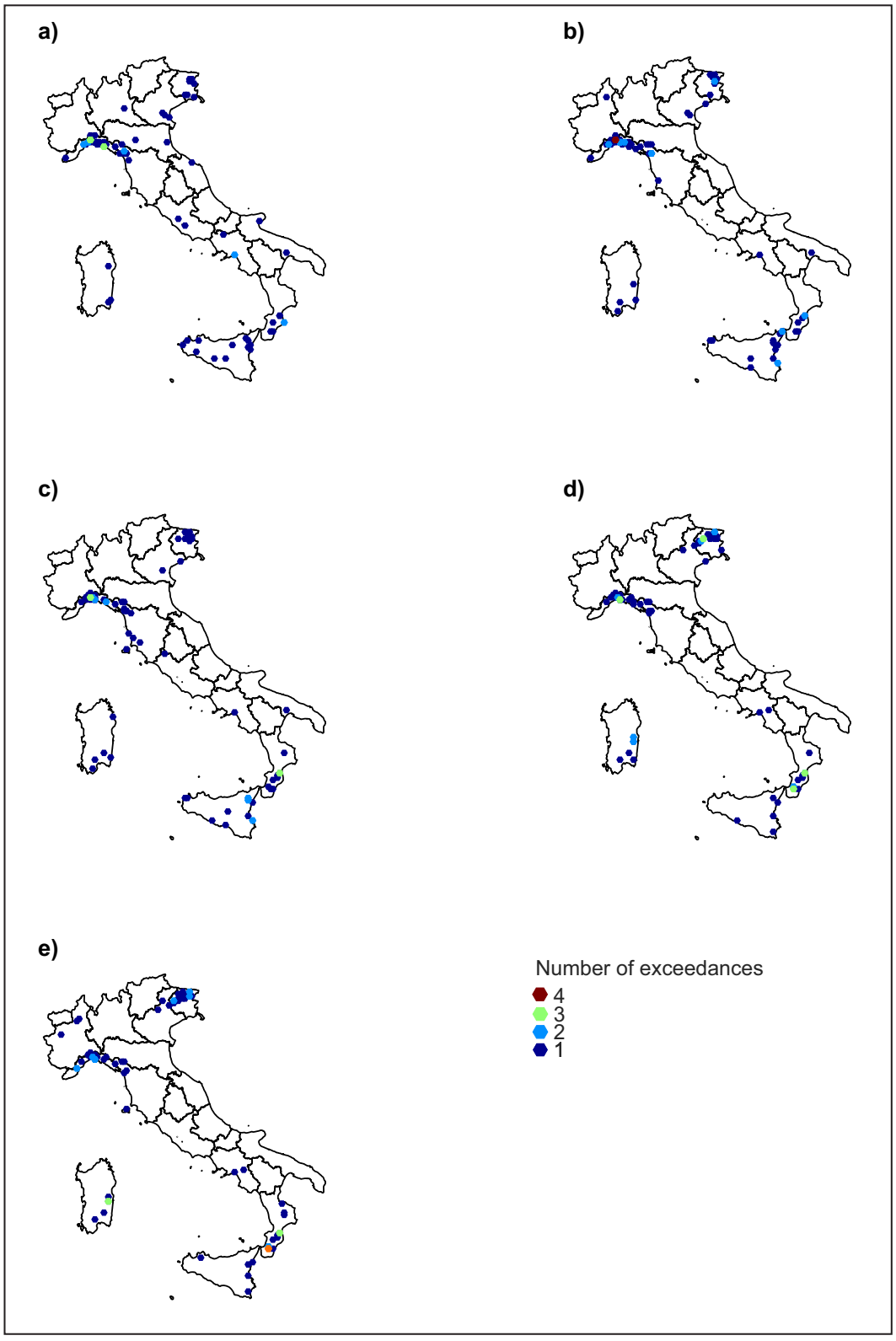


Fig. D.1 Number of exceedances of the 0.999 empirical quantile, considering the absolute rainfall amounts, for the 1 (a) to 24 (e) hours durations.

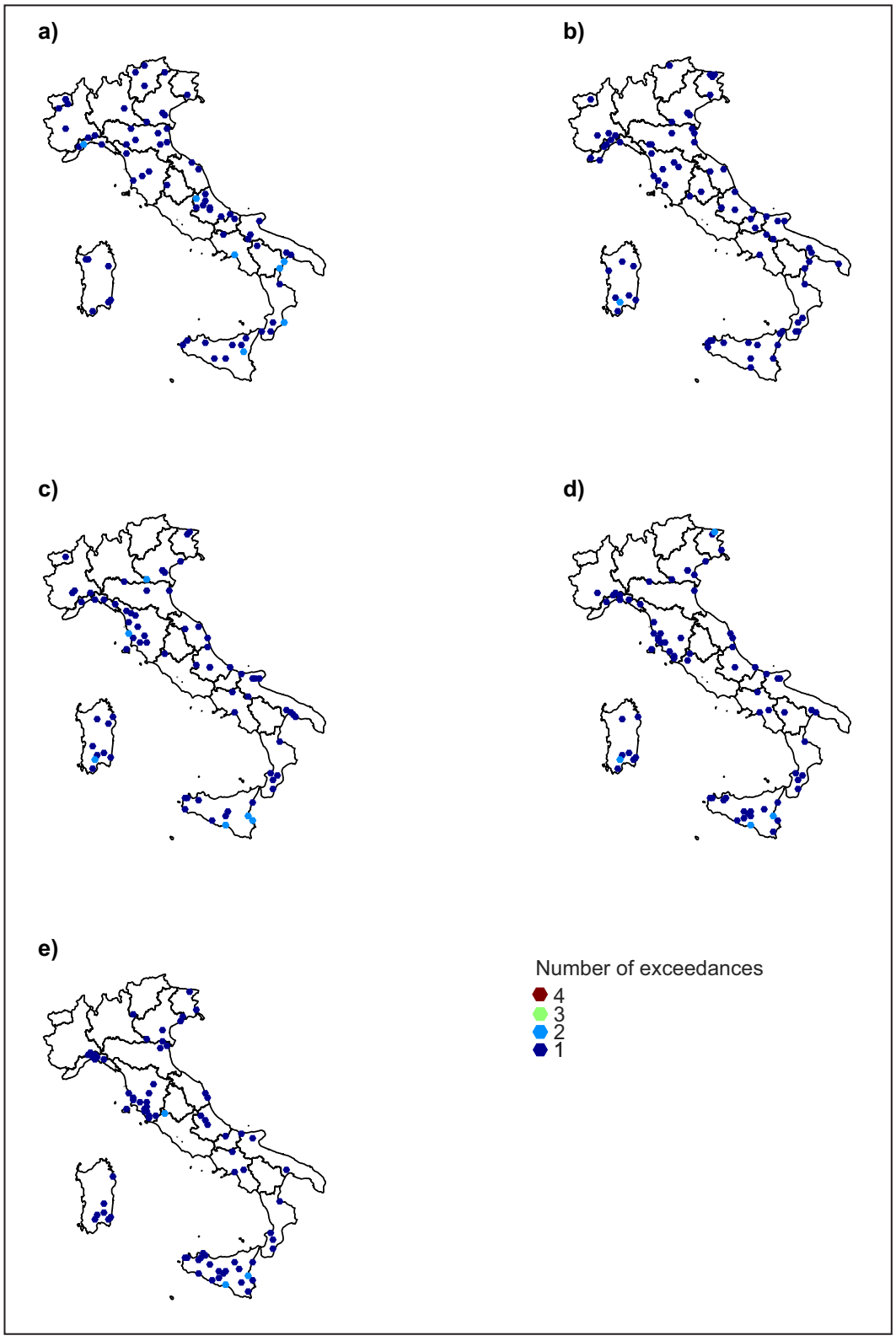


Fig. D.2 Number of exceedances of the 0.999 empirical quantile, considering the normalized rainfall, for the 1 (a) to 24 (e) hours durations.

D.2 Spatial distribution of the over-threshold intensities

This section reports the maps of the amount of the maximum values exceeding the 0.999 empirical quantiles per each series for the 5 considered durations.

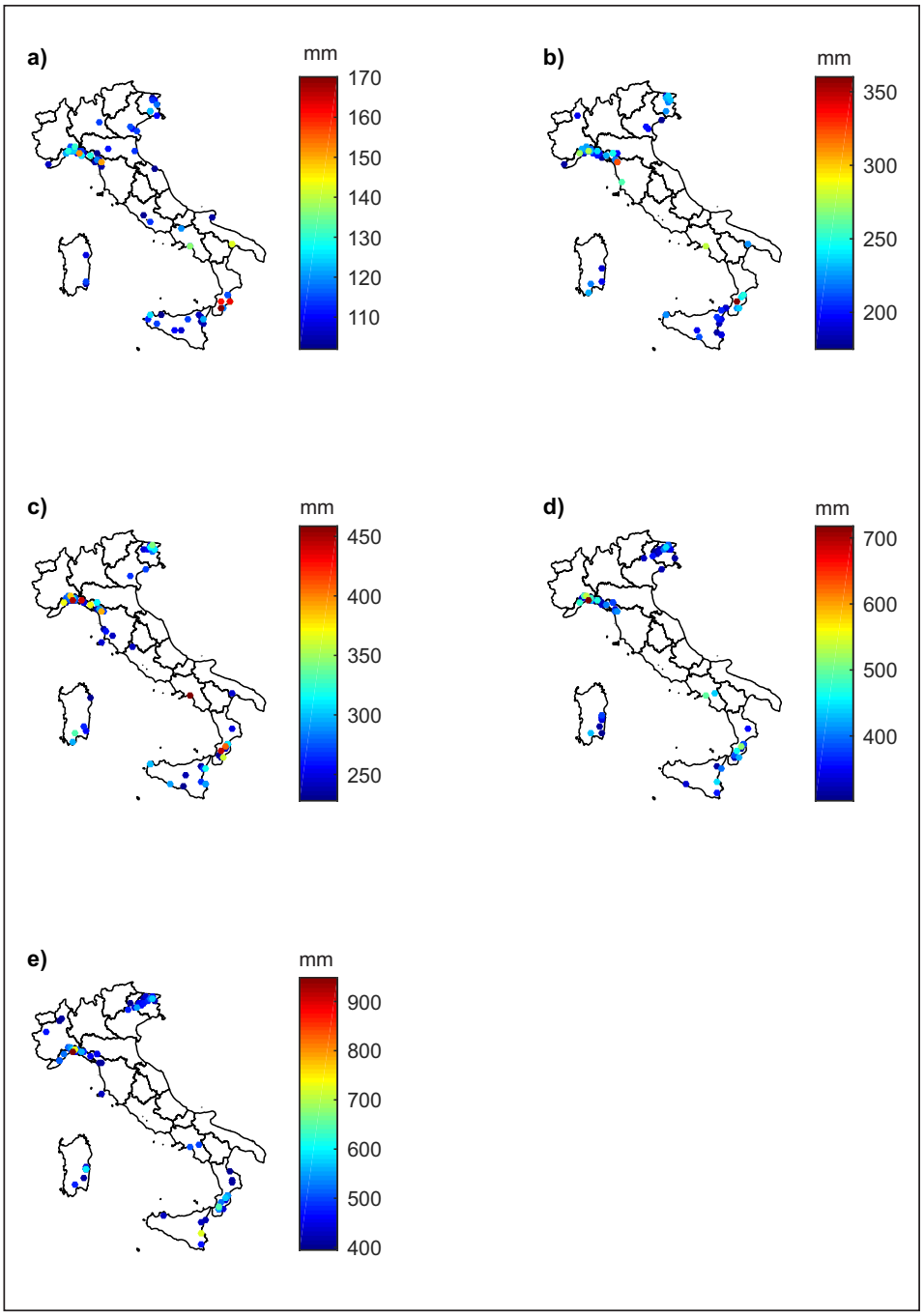


Fig. D.3 Maximum absolute rainfall (h_d, mm) recorded at the stations with at least one value over-threshold for the 1 (a) to 24 (e) hours durations.

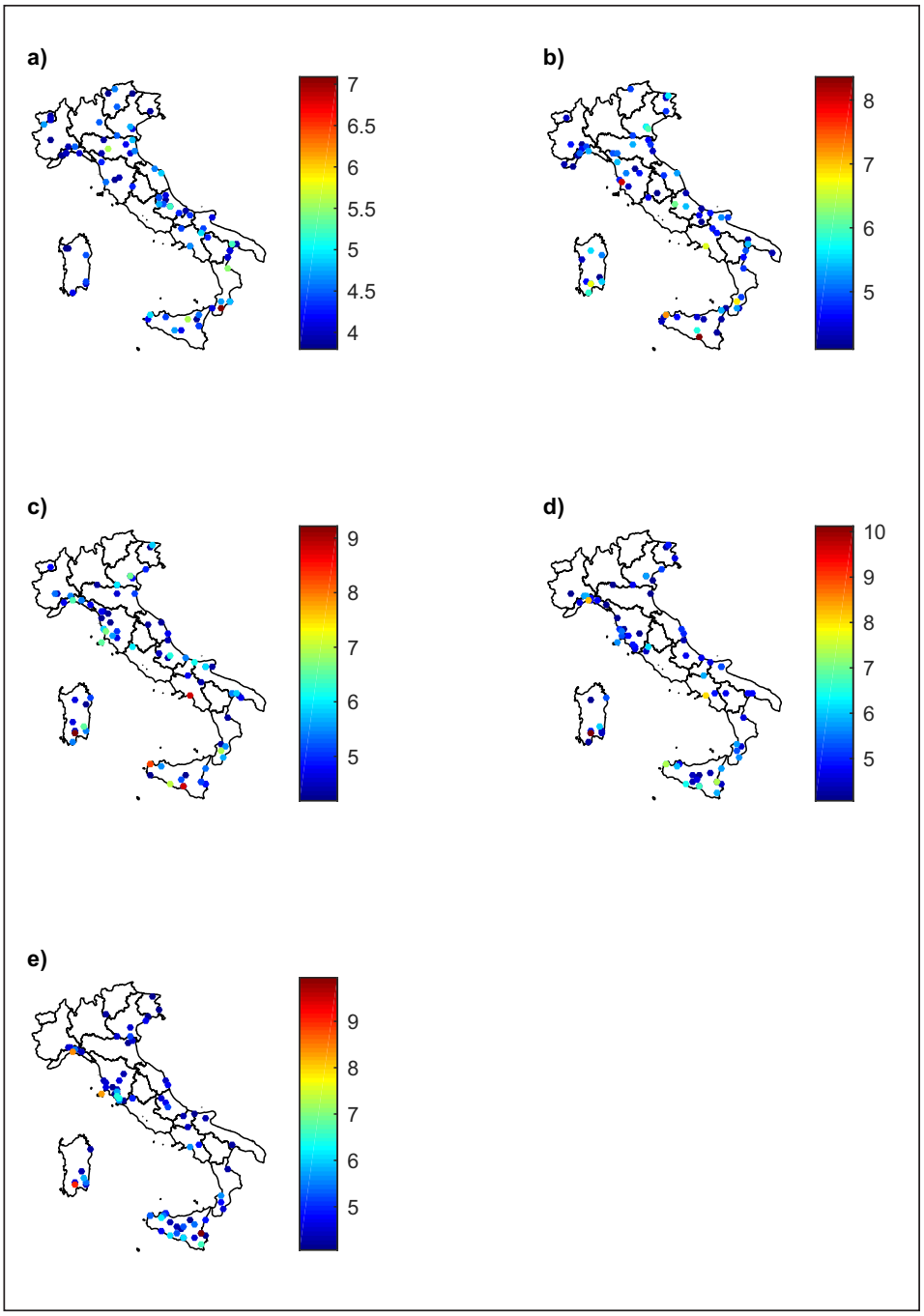


Fig. D.4 Maximum normalized rainfall (h_d^*) recorded at the stations with at least one value over-threshold for the 1 (a) to 24 (e) hours durations.

D.3 Absolute rainfall depths on Gumbel probability paper

This section reports empirical frequency curves of the Italian series of annual maxima for each duration, considering the absolute values, plotted on the Gumbel probability paper.

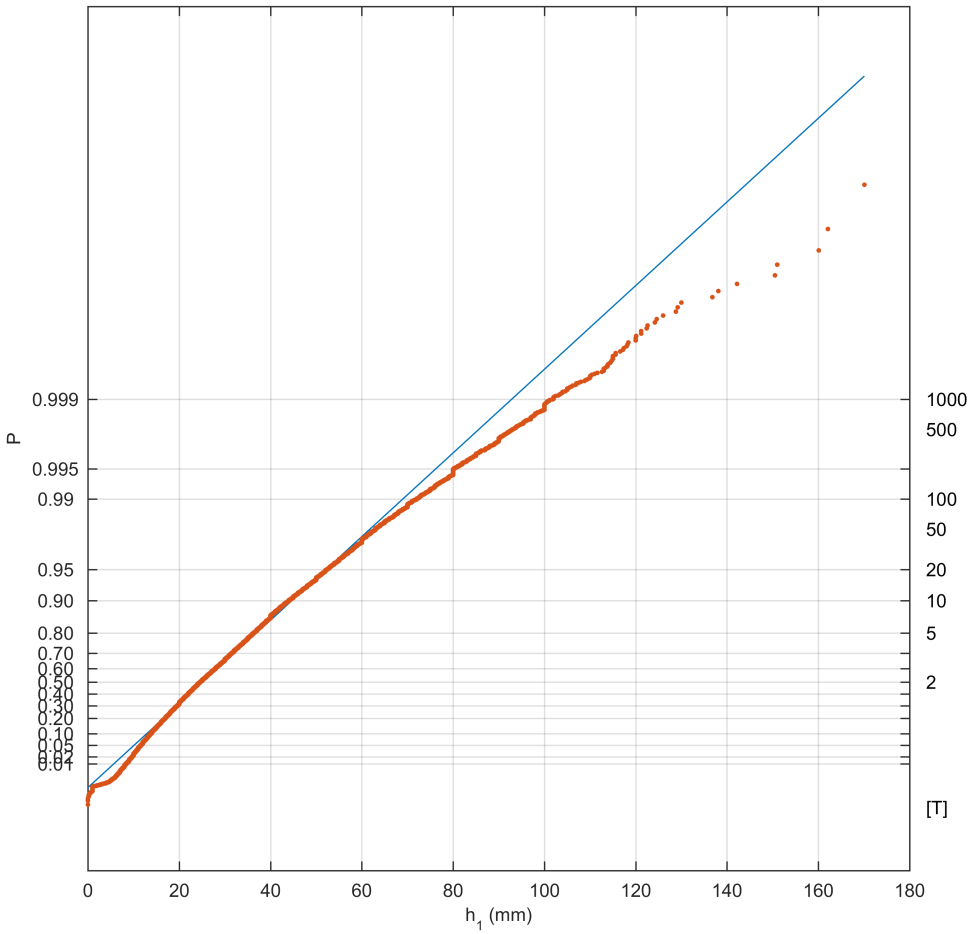


Fig. D.5 Annual maxima (mm) for 1 hour duration on Gumbel probability paper. The blue line represent the Gumbel distribution with parameters estimated by the L-moments method.

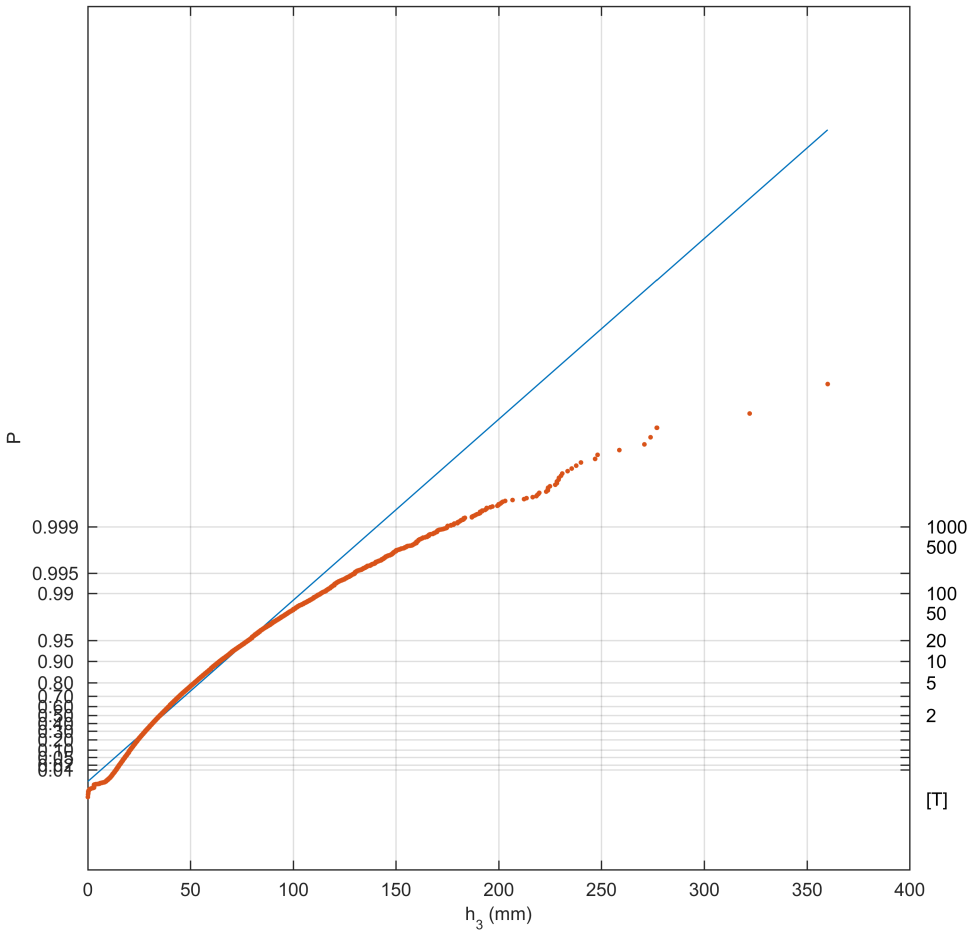


Fig. D.6 Annual maxima (mm) for 3 hours duration on Gumbel probability paper. The blue line represent the Gumbel distribution with parameters estimated by the L-moments method.

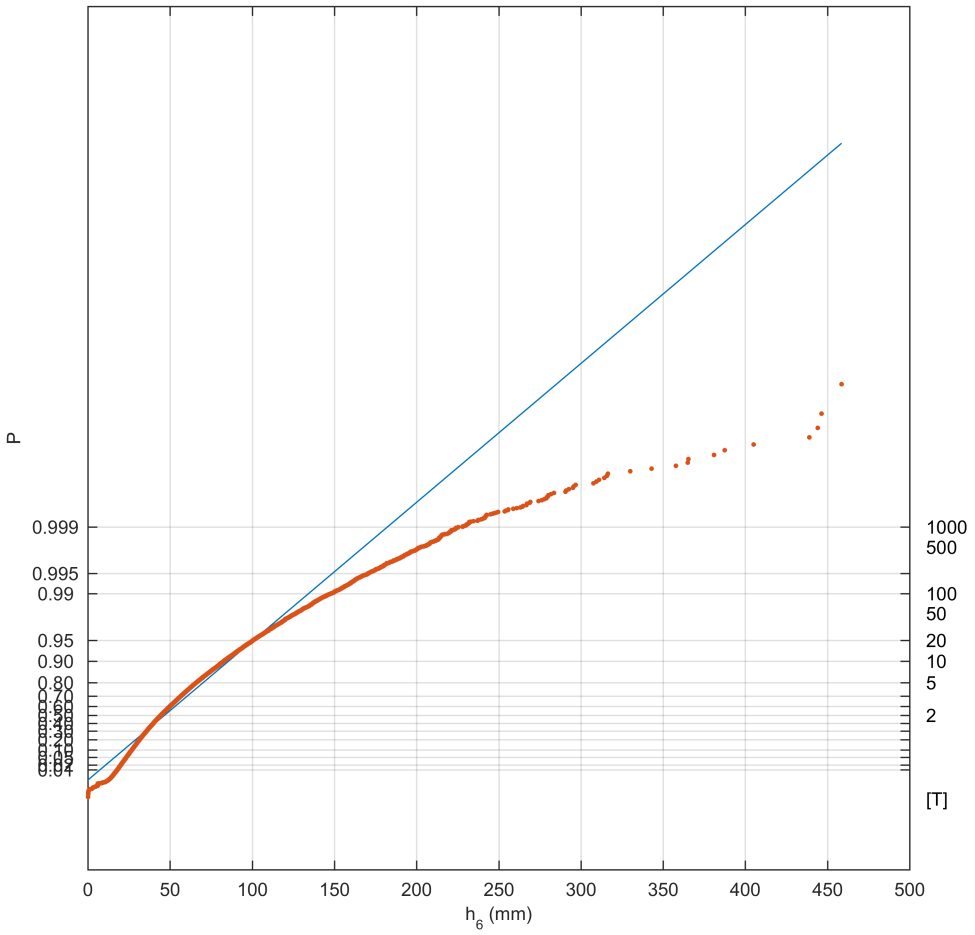


Fig. D.7 Annual maxima (mm) for 6 hours duration on Gumbel probability paper. The blue line represent the Gumbel distribution with parameters estimated by the L-moments method.

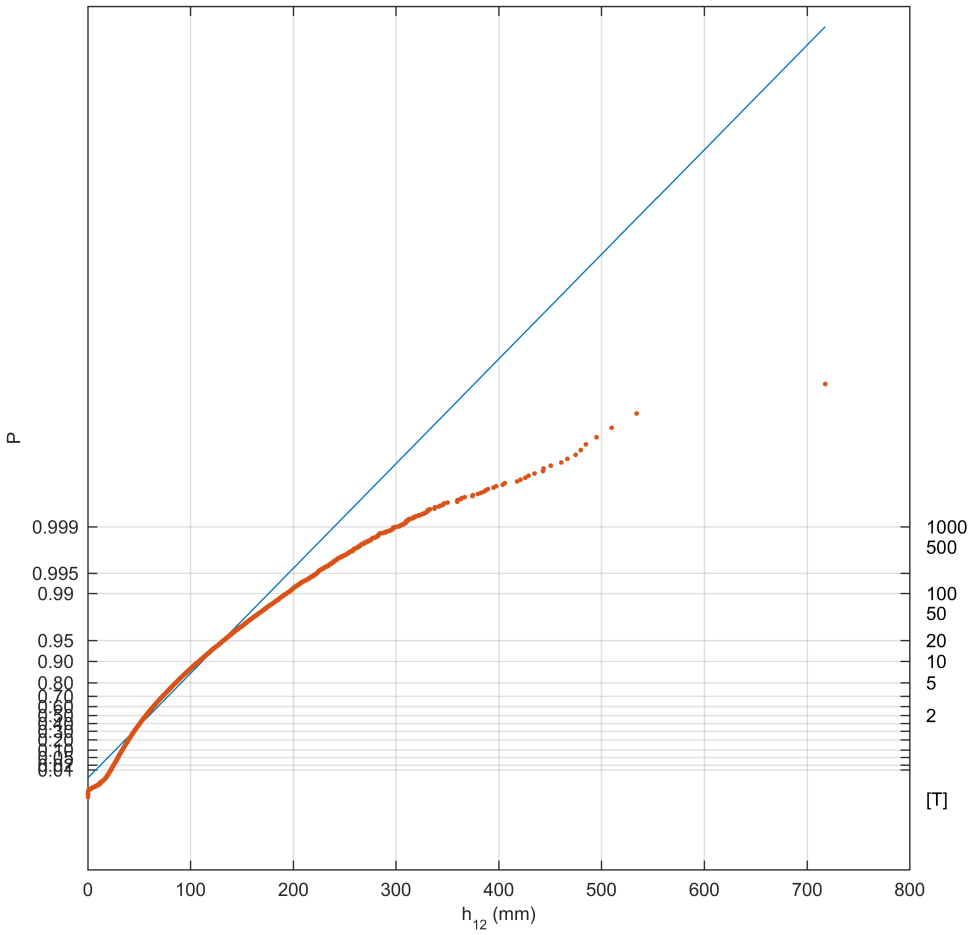


Fig. D.8 Annual maxima (mm) for 12 hours duration on Gumbel probability paper. The blue line represent the Gumbel distribution with parameters estimated by the L-moments method.

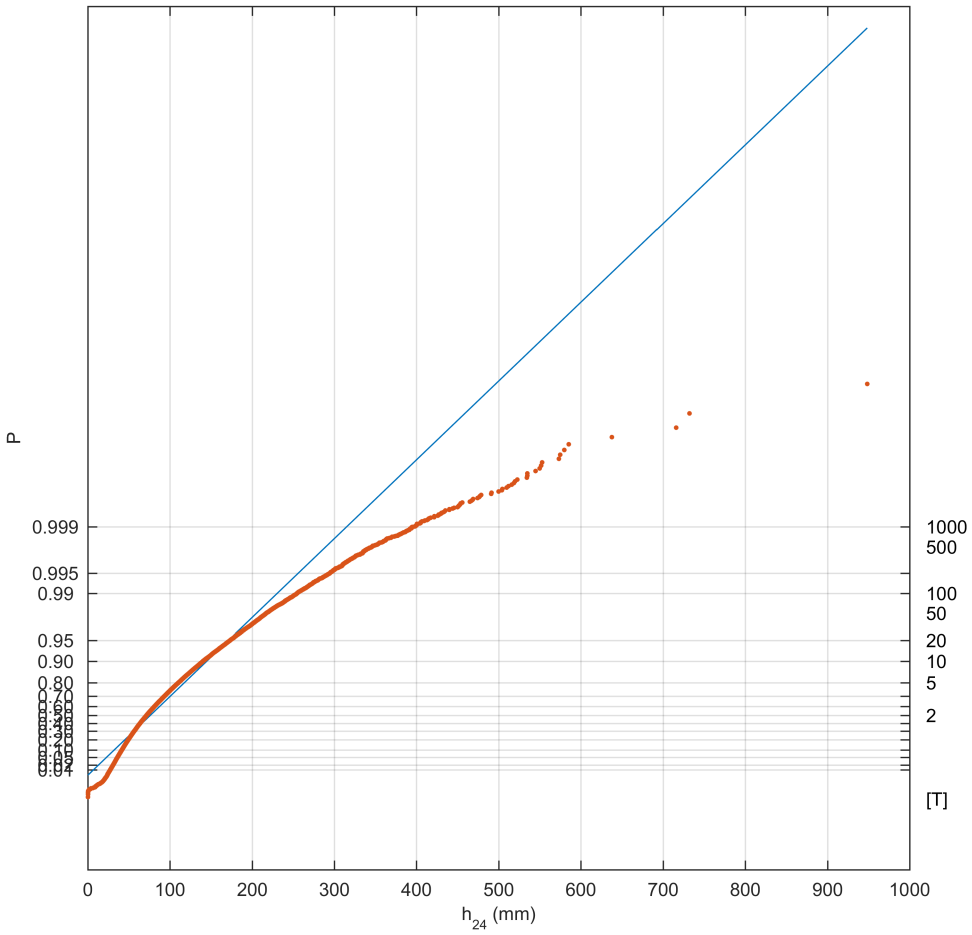


Fig. D.9 Annual maxima (mm) for 24 hours duration on Gumbel probability paper. The blue line represent the Gumbel distribution with parameters estimated by the L-moments method.

D.4 Representation of the spatial distribution of the stations

D.4.1 Comparison on a 50 km gridded domain

This section reports, over a 50 km grid, the number of exceedances of the 0.999 empirical quantile, evaluated considering the absolute rainfall amounts and the normalized ones. The exceedances are compared with the number of exceedances normalized with the number of station-years per cell.

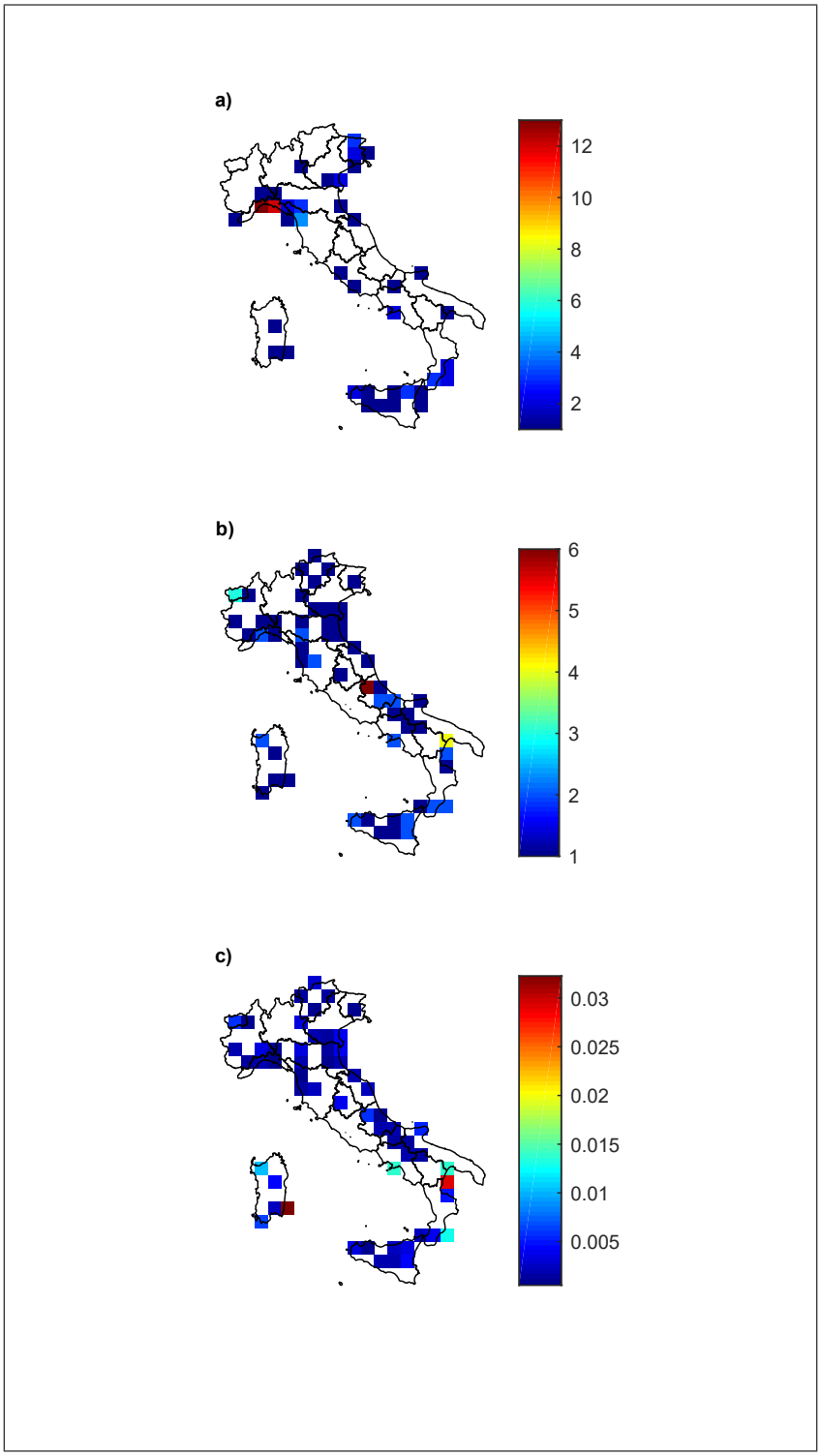


Fig. D.10 (a) Number of absolute rainfall heights (b) and number of normalized rainfall amounts exceeding the 0.999 empirical quantile for the 1 hour duration. (c) Same as (b) but normalized with the number of rain gauges per cell.

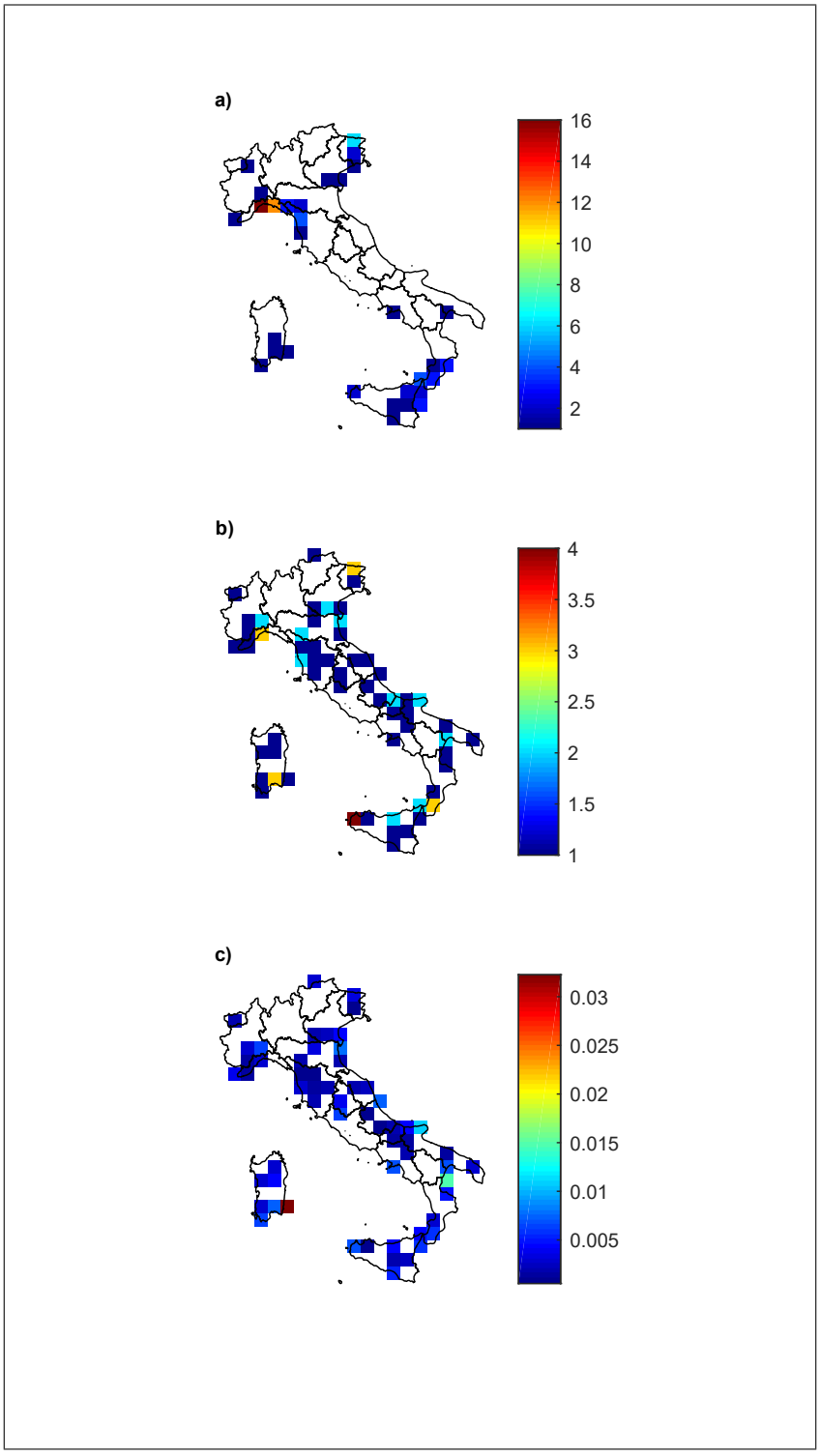


Fig. D.11 (a) Number of absolute rainfall heights (b) and number of normalized rainfall amounts exceeding the 0.999 empirical quantile for the 3 hours duration. (c) Same as (b) but normalized with the number of rain gauges per cell.

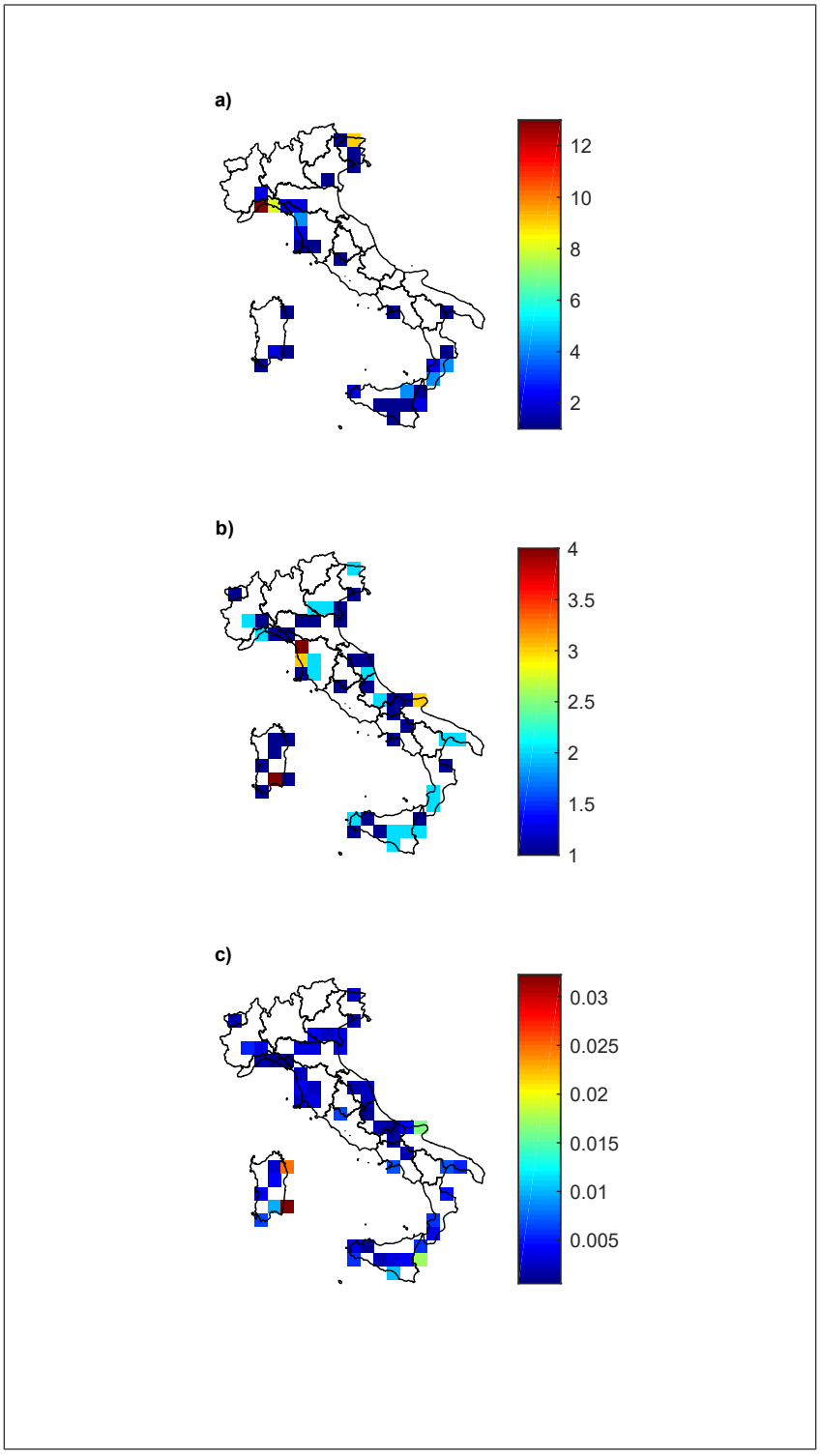


Fig. D.12 (a) Number of absolute rainfall heights (b) and number of normalized rainfall amounts exceeding the 0.999 empirical quantile for the 6 hours duration. (c) Same as (b) but normalized with the number of rain gauges per cell.

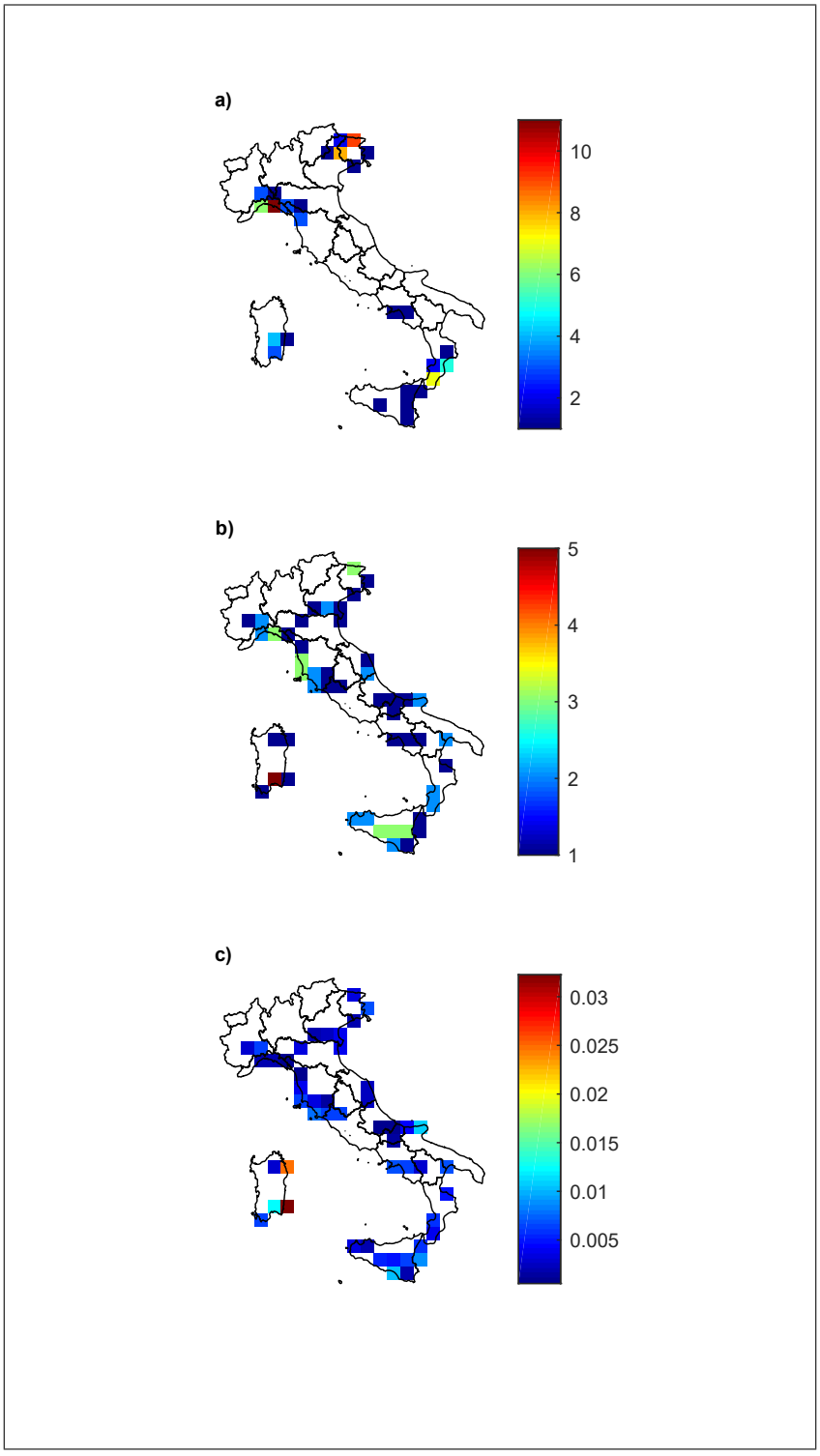


Fig. D.13 (a) Number of absolute rainfall heights (b) and number of normalized rainfall amounts exceeding the 0.999 empirical quantile for the 12 hours duration. (c) Same as (b) but normalized with the number of rain gauges per cell.

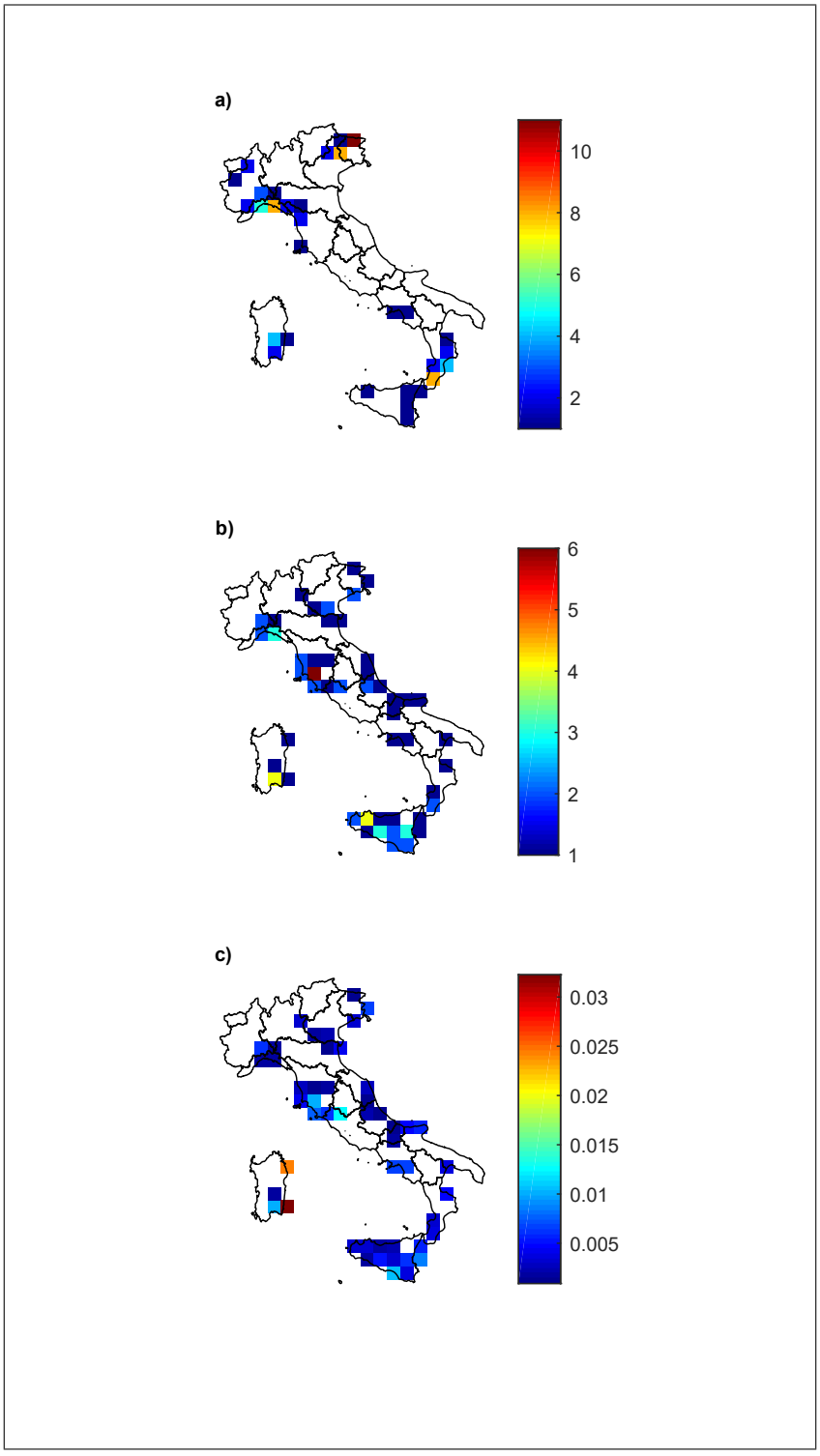


Fig. D.14 (a) Number of absolute rainfall heights (b) and number of normalized rainfall amounts exceeding the 0.999 empirical quantile for the 24 hours duration. (c) Same as (b) but normalized with the number of rain gauges per cell.

D.4.2 Results at different resolutions

This section reports the number of normalized rainfall amounts exceeding the 0.999 empirical quantile, normalized with the number of station-years, considering grids with different cell sizes for the 1-3-6-12 and 24 hours durations.

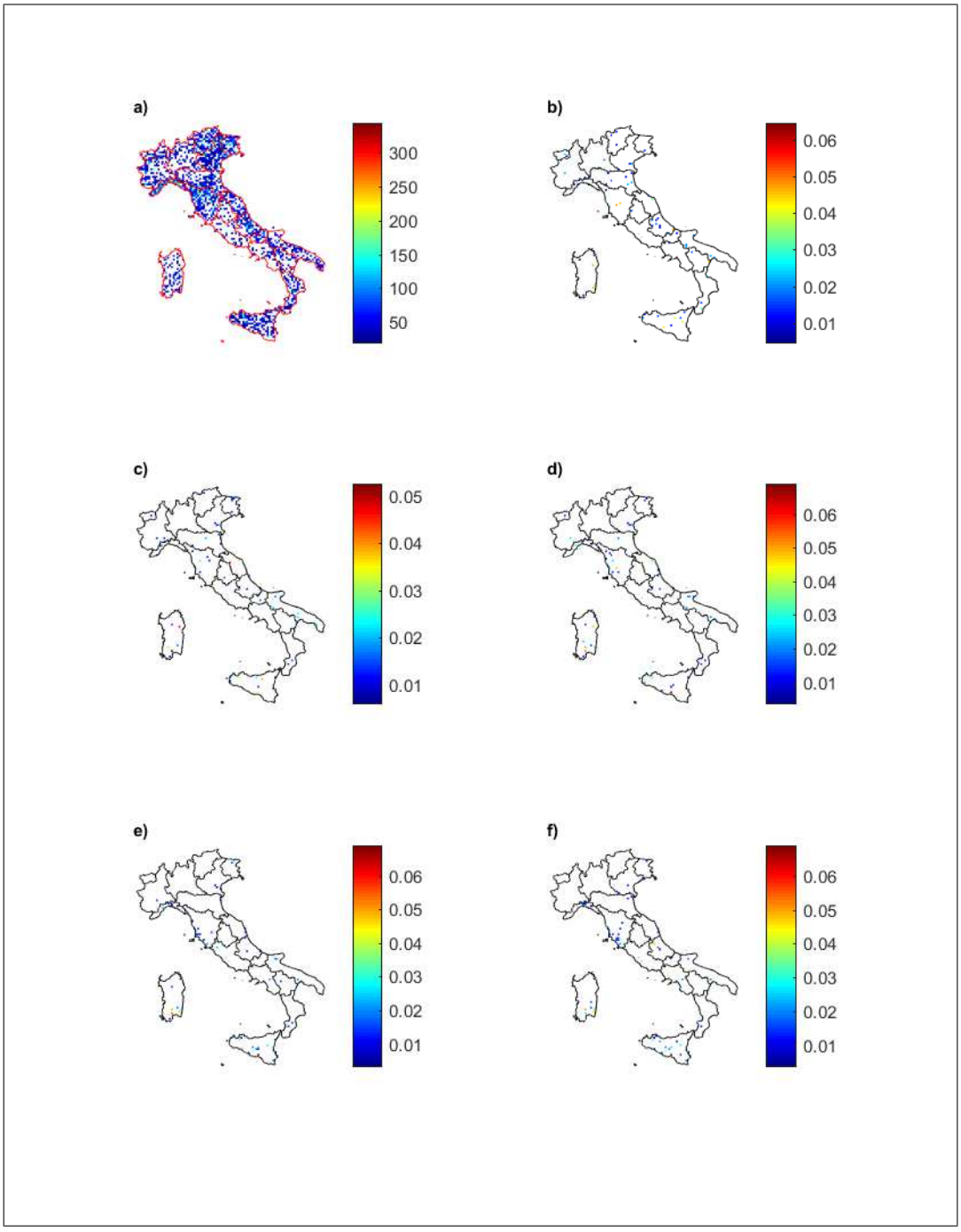


Fig. D.15 (a) Number of station-years per cell on a 10 km gridded domain. Number of normalized rainfall amounts exceeding the 0.999 empirical quantile normalized with the number of station-years for the duration from 1 (b) to 24 (f) hours.

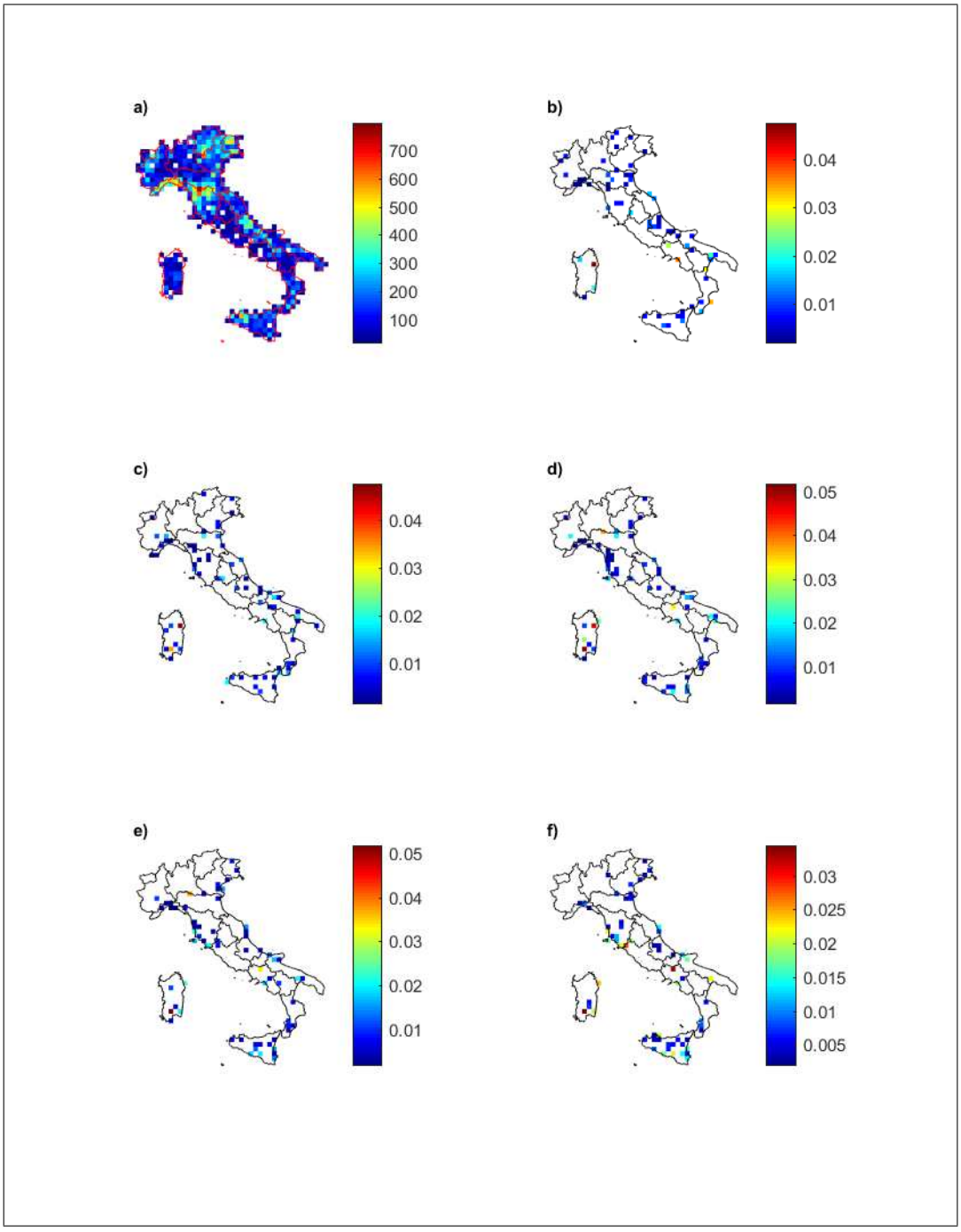


Fig. D.16 (a) Number of station-years per cell on a 25 km gridded domain. Number of normalized rainfall amounts exceeding the 0.999 empirical quantile normalized with the number of station-years for the duration from 1 (b) to 24 (f) hours

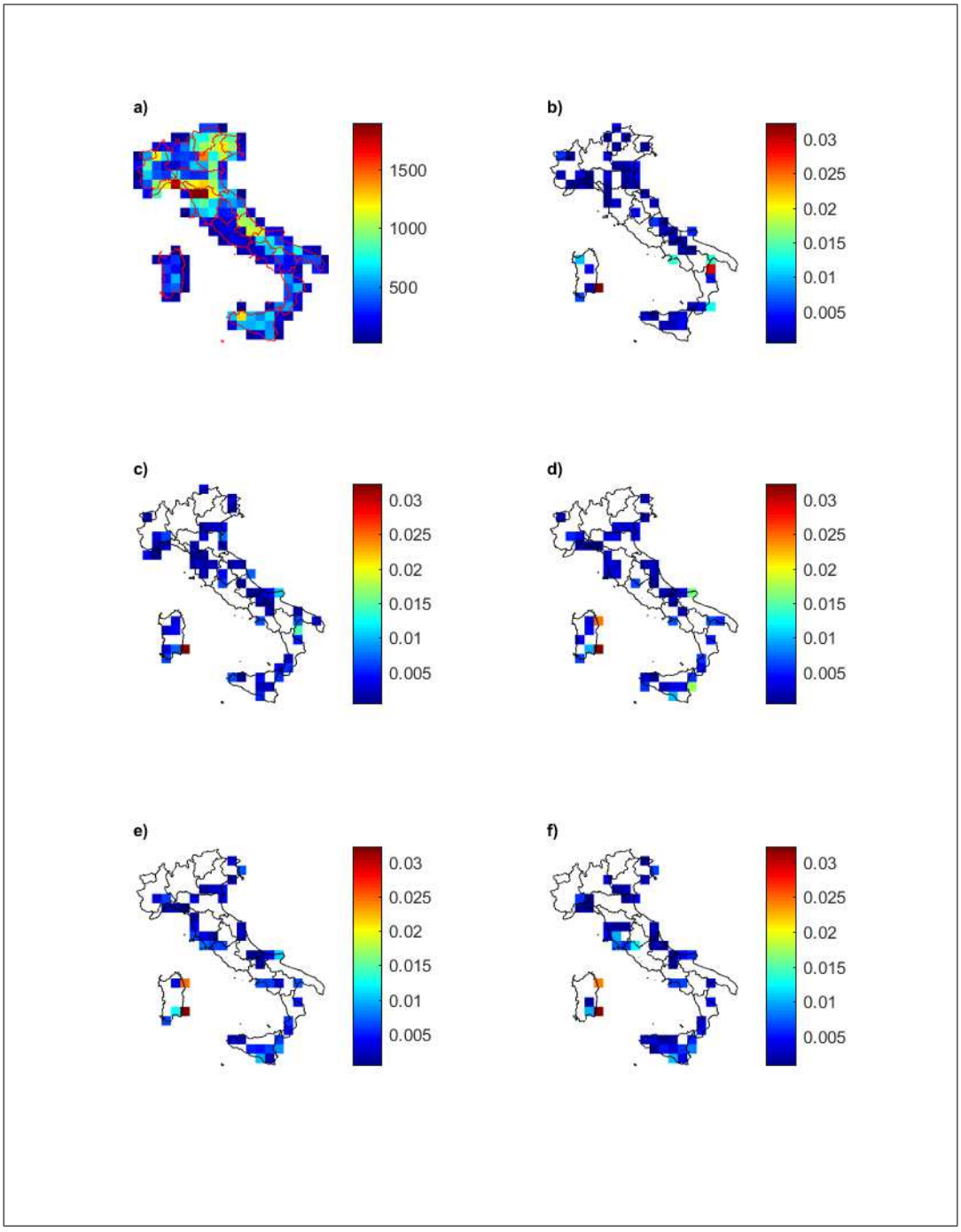


Fig. D.17 (a) Number of station-years per cell on a 50 km gridded domain. Number of normalized rainfall amounts exceeding the 0.999 empirical quantile normalized with the number of station-years for the duration from 1 (b) to 24 (f) hours

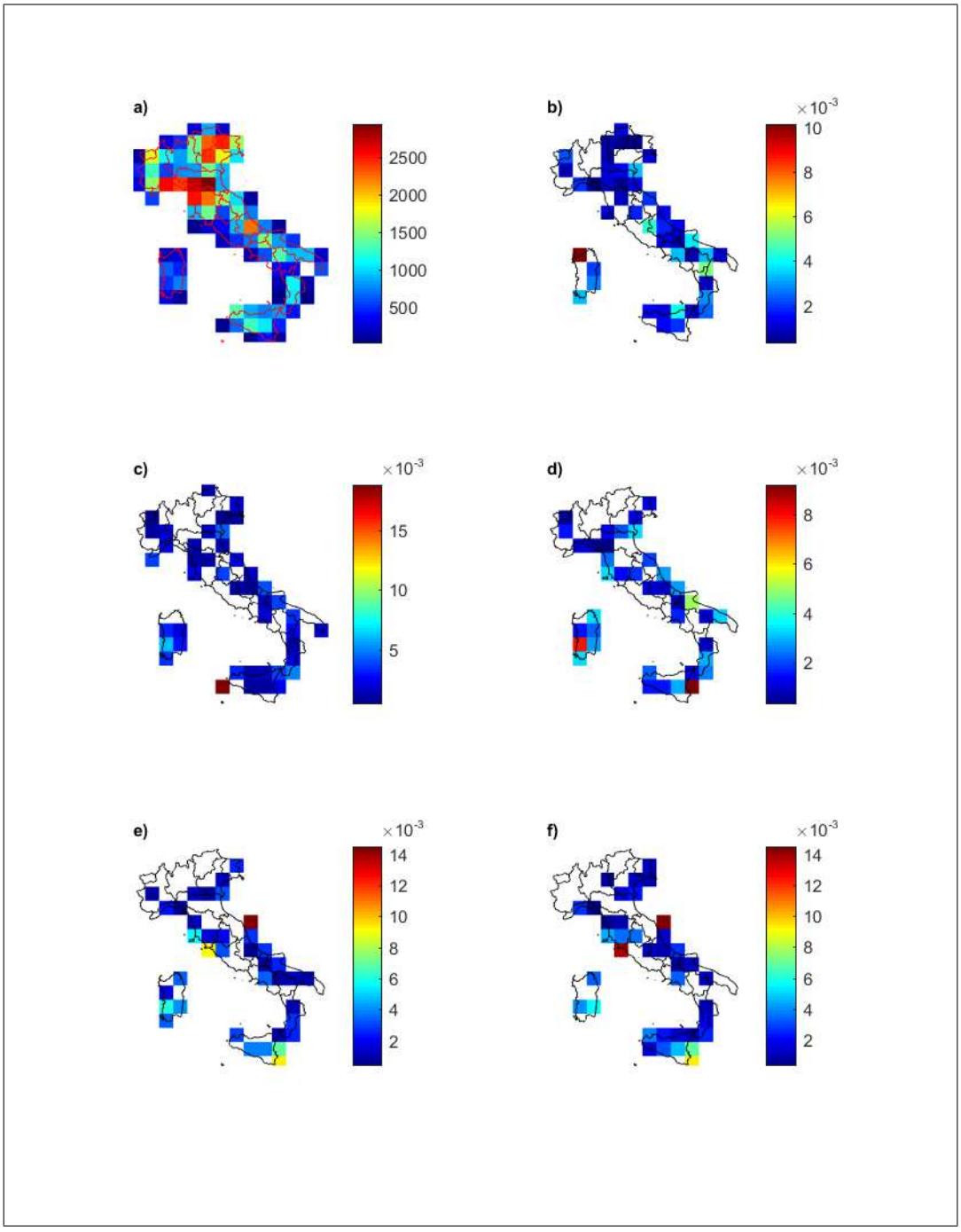


Fig. D.18 (a) Number of station-years per cell on a 75 km gridded domain. Number of normalized rainfall amounts exceeding the 0.999 empirical quantile normalized with the number of station-years for the duration from 1 (b) to 24 (f) hours

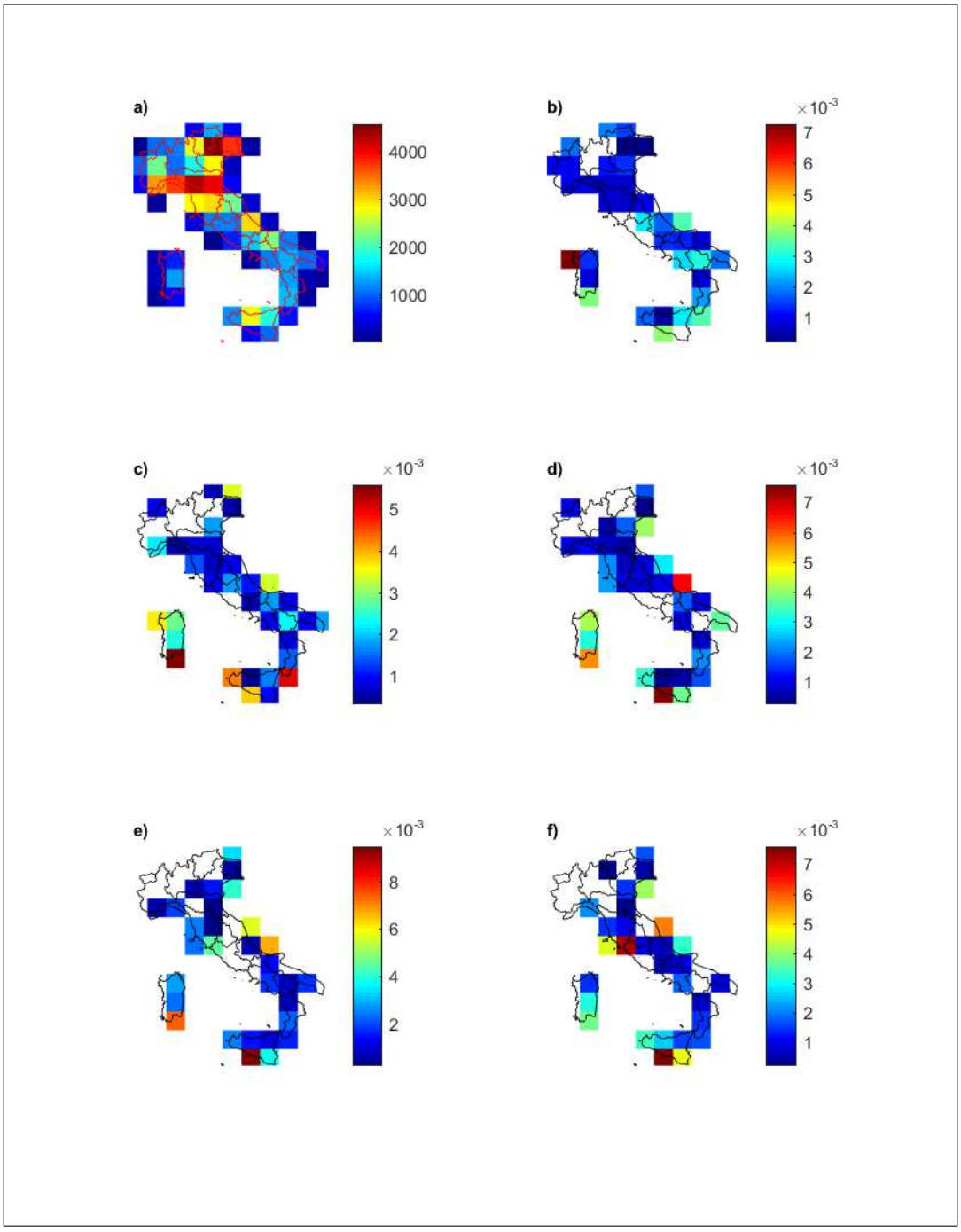


Fig. D.19 (a) Number of station-years per cell on a 100 km gridded domain. Number of normalized rainfall amounts exceeding the 0.999 empirical quantile normalized with the number of station-years for the duration from 1 (b) to 24 (f) hours

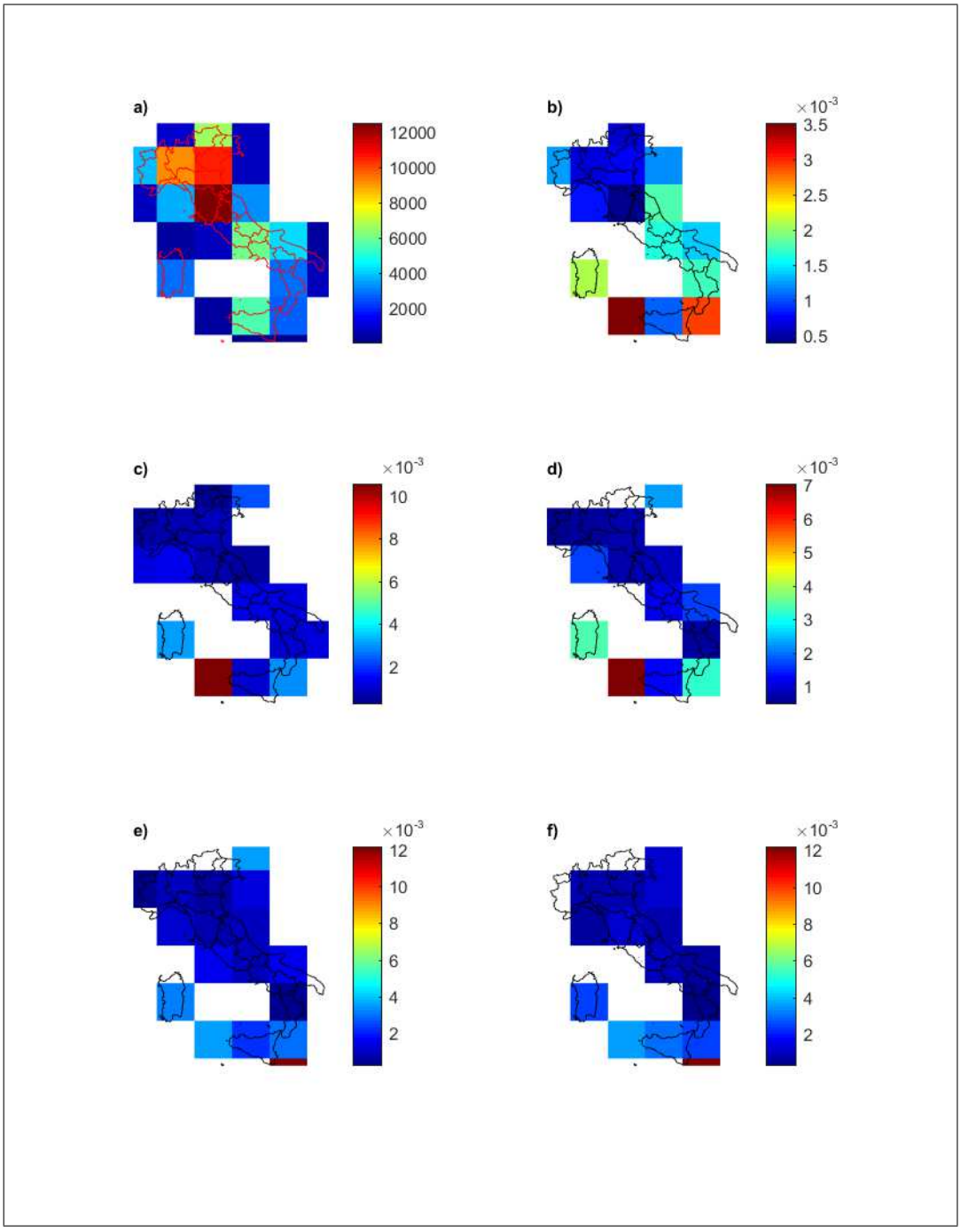


Fig. D.20 (a) Number of station-years per cell on a 200 km gridded domain. Number of normalized rainfall amounts exceeding the 0.999 empirical quantile normalized with the number of station-years for the duration from 1 (b) to 24 (f) hours

D.4.3 Stations considered in the morpho-climatological analysis of the over-threshold extremes

This section reports the name and coordinates of the stations considered in the morpho-climatological analysis (i.e., the ones exceeding the 0.999 empirical quantile considering the normalized rainfall for the duration 1 and 24 hours). Coordinates are reported in the WGS84-UTM32N reference system.

Table D.1 Stations with normalized values overcoming the cumulative frequency 0.999 for the 1 hour duration.

COD	NAME	X (m)	Y (m)	REGION
1009	Barisciano	876911	4695848	Abruzzo
1016	Bussi_Officine	899669	4682997	Abruzzo
1067	Isola_Del_G_Sasso	883438	4714121	Abruzzo
1071	Lentella	970181	4664768	Abruzzo
1102	Pescosansonesco	902955	4689499	Abruzzo
1118	Rosello	943636	4652428	Abruzzo
1137	Teramo	886675	4733272	Abruzzo
1140	Tornimparte	854671	4690789	Abruzzo
1154	Poggio_Cancelli_Simn	855894	4718102	Abruzzo
2048	Metaponto	1163684	4497870	Basilicata
2058	Nova_Siri_Scalo	1150390	4470568	Basilicata
2069	Ripacandida	1067396	4551342	Basilicata
3043	Schiavonea_(Cs)	1147137	4416547	Calabria
3112	Stilo_(Rc)	1161038	4281655	Calabria
3129	San_Luca_(Rc)	1119604	4245951	Calabria
3146	Villa_San_Giovanni_(Rc)	1081125	4250762	Calabria
3165	Giffone_(Rc)	1124016	4278929	Calabria
4256	S_Gregorio_Matese_(En_)_Simn	948895	4595122	Campania
4259	Salerno_Simn	987519	4519516	Campania
5023	Benvignante	718465	4949771	Emilia-Romagna
5078	Castellarano	636634	4928634	Emilia-Romagna

5113	Faenza	730270	4908274	Emilia-Romagna
5152	Ligonchio_Centrale	607253	4907960	Emilia-Romagna
5166	Marozzo	746706	4964541	Emilia-Romagna
5246	Poviglio	621535	4966533	Emilia-Romagna
5255	Ravenna	753521	4921370	Emilia-Romagna
6073	Cormor_Paradiso	822710	5089098	Friuli Venezia Giulia
8093	Ellera_Ligure (Già_Ellera_Ed_Ellera_-_Foglietto)	457646	4912238	Liguria
8122	Isola_Del_Cantone	496717	4943516	Liguria
8175	Osiglia_-_Diga	435974	4906697	Liguria
8207	Rossiglione	472826	4935499	Liguria
8215	S_Pietro_Di_Novella	516001	4912170	Liguria
9033	Brescia	598525	5042101	Lombardia
9349	Vallazza	676829	4986940	Lombardia
10089	Senigallia	839781	4847200	Marche
10115	Osimo	862572	4824262	Marche
11017	Larino	991823	4644597	Marche
12089	Lombriasco	392119	4965430	Piemonte
12353	Noasca	368232	5034858	Piemonte
13031	Castellaneta	1171174	4527920	Puglia
13032	Castelluccio_Dei_Sauri	1042309	4592631	Puglia
13073	Massafra	1186638	4523358	Puglia
13082	Monte_Santangelo	1079373	4640562	Puglia
13121	Santagata_Di_Puglia	1035610	4574994	Puglia
14014	Is_Cannoneris_(Caserma)	487275	4320684	Sardegna
14054	Bunnari	467825	4507540	Sardegna

14055	Sassari_R_U_	462405	4508030	Sardegna
14073	Lula	541314	4480081	Sardegna
14095	Muravera	548464	4363993	Sardegna
14098	Monte_Acuto_(C_Ra)	541074	4355943	Sardegna
15011	Montalbano_Elicona	1028384	4225638	Sicilia
15077	Risalaimi	887957	4212229	Sicilia
15110	Lentina_(Cont_Da)	822522	4218181	Sicilia
15120	Birgi_Nuovo	807456	4198898	Sicilia
15188	Racalmuto	918732	4151029	Sicilia
15229	Pietraperzia	954832	4154123	Sicilia
15316	Bronte	1013866	4198229	Sicilia
15321	Capizzi_Bis	982287	4203316	Sicilia
15333	Paterno	1022197	4174487	Sicilia
16064	Fornovolasco	608910	4876100	Toscana
16237	Casole_Delsa	665675	4800970	Toscana
16349	Castelluccio	631430	4787190	Toscana
16401	Vagliagli	689954	4812132	Toscana
17042	Pozzolago_(Centrale)	672300	5114847	Prov. Trento
18012	Silandro_-_Schlanders	636292	5164894	Prov. Bolzano
18026	S_Leonardo_In_Passiria_-_St_Leonhard_In_Passeier	671444	5186398	Prov. Bolzano
19015	Villastrada_Umbra	746600	4770775	Umbria
20008	Verres_A_	397580	5057950	Valle D'Aosta
20014	Chatillon	392544	5067290	Valle D'Aosta
20037	Promiod_Covalou_Centrale	391310	5071020	Valle D'Aosta
21097	Podestagno_(Cortina_Dampezzo)	738480	5164910	Veneto
21256	Piove_Di_Sacco_Simn	737981	5020854	Veneto
21279	Santa_Margherita_Di_Cod_Simn	744727	5015549	Veneto

Table D.2 Stations with normalized values overcoming the cumulative frequency 0.999 for the 24 hours duration.

COD	NAME	X	Y	REGION
1035	Castel_Del_Monte	888984	4701337	Abruzzo
1050	Collepietro	894540	4685399	Abruzzo
1082	Nerito	867671	4720386	Abruzzo
1141	Torrebruna	959900	4650055	Abruzzo
3043	Schiavonea_(Cs)	1147137	4416547	Calabria
3126	Ardore_Superiore_(Rc)	1126815	4252943	Calabria
3165	Giffone_(Rc)	1124016	4278929	Calabria
3173	Vibo_Valentia_(Vv)	1118484	4305005	Calabria
4254	S_Croce_Del_Sannio_Simn	979697	4595250	Campania
4259	Salerno_Simn	987519	4519516	Campania
4263	Senerchia_Simn	1023772	4527576	Campania
5025	Berra	736356	4984966	Emilia-Romagna
5166	Marozzo	746706	4964541	Emilia-Romagna
5170	Martinella	728470	4954938	Emilia-Romagna
6094	Gorizia_Presa_C_B_P_I_	856289	5097008	Friuli Venezia Giulia
6158	Pontebba	830468	5159299	Friuli Venezia Giulia
7017	Bolsena	744881	4725240	Lazio
8030	Busalla	495944	4935153	Liguria
8091	Diga_Val_Noci	502763	4926622	Liguria
8101	Genova_-_Bolzaneto	491695	4922449	Liguria
8140	Madonna_Della_Guardia	489104	4926202	Liguria
8159	Monte_Cappellino	496574	4933033	Liguria
8207	Rossiglione	472826	4935499	Liguria
8247	Tigliolo	528771	4921171	Liguria

9349	Vallazza	676829	4986940	Lombardia
10085	Pedaso	893456	4781419	Marche
10087	Porto_Sant_Elpidio	885619	4799543	Marche
12093	Lavagnina_Centrale	479739	4939750	Piemonte
13031	Castellaneta	1171174	4527920	Puglia
13114	San_Marco_In_Lamis	1052216	4638929	Puglia
13144	Chieuti_Scalo	1010555	4659423	Puglia
14008	Nuraminis	501155	4365903	Sardegna
14011	Decimomannu_(Viv_Forest_)	497475	4349943	Sardegna
14070	Budoni_(C_Ra)	558973	4506840	Sardegna
14093	Seui_F_C_	527914	4409752	Sardegna
14094	Ballao	530954	4377843	Sardegna
14095	Muravera	548464	4363993	Sardegna
14098	Monte_Acuto_(C_Ra)	541074	4355943	Sardegna
15057	Vicari	902314	4196728	Sicilia
15081	Pioppo	873494	4220507	Sicilia
15093	Palermo_Villa_Trabia	881072	4229498	Sicilia
15096	Palermo_Piazza_Verdi	882026	4228109	Sicilia
15102	Fellamonica	863157	4209703	Sicilia
15110	Lentina_(Cont_Da)	822522	4218181	Sicilia
15111	S_Andrea_Bonagia	817208	4218191	Sicilia
15151	Sciacca	864845	4159862	Sicilia
15180	Marianopoli	932729	4173281	Sicilia
15201	Agrigento_(Isp_Agrario)	906765	4138622	Sicilia
15204	Canicatti	929668	4145808	Sicilia
15222	Villarosa	957091	4172643	Sicilia
15225	Caltanissetta	947215	4162134	Sicilia
15252	Diga_Comunelli	957525	4124786	Sicilia
15264	Vizzini	1010710	4128476	Sicilia
15276	Noto	1040657	4099866	Sicilia
15287	Augusta	1052443	4138658	Sicilia
15298	Lentini_Bonifica	1038936	4150500	Sicilia
15310	Diga_Ancipa	990986	4201656	Sicilia
15329	Centuripe	1006921	4179388	Sicilia

15370	Alcantara_(Teleferiva)	1050713	4204968	Sicilia
16153	Renacci_(Fattoria)	703885	4829580	Toscana
16318	Vada_(Bonifica)	618325	4798385	Toscana
16349	Castelluccio	631430	4787190	Toscana
16355	S-Carlo_Solvay	629700	4773350	Toscana
16377	Massa_Marittima	653850	4768500	Toscana
16387	Roccastrada	676800	4764375	Toscana
16393	Batignano	677040	4748740	Toscana
16394	Acquisti	668365	4746240	Toscana
16396	Grosseto	672990	4736055	Toscana
16407	Siena_Universita	688975	4799420	Toscana
16443	Ponte_Tura	677355	4737095	Toscana
16446	Poggio_Perotto	685065	4720545	Toscana
16455	Manciano	708940	4717840	Toscana
16458	S-Donato	682670	4711235	Toscana
16466	Portoferraio_Citta	609100	4741410	Toscana
17069	Forte_Dampolla	627788	5080243	Prov.Trento
21222	Legnaro_Simn	732547	5026214	Veneto
21290	Termine_Simn	796508	5056735	Veneto
21302	Villa_Simn	807474	5072111	Veneto

Statistical estimation of design rainfall is considered a consolidated topic in hydrology. However, extreme rainfalls and their consequences still constitute one of the most critical natural risks worldwide, particularly in urban environments. Additional efforts for improving the spatio-temporal analysis of extreme rainfall are therefore required, particularly at the regional scale.

This dissertation investigates empirical evidences, systematic data, tools and methodologies for improving the spatial statistical analysis of extreme precipitations. Italy is considered a challenging case study, due to its peculiar geographic-orographic-urbanistic setting, often associated with storm-induced disasters.

The six chapters of this work cover different spatial and technical domains, describing the efforts made in different directions, aimed at providing advances for the definition of a comprehensive framework for better assessing the hazard related to severe rainstorms, at different spatial scales, according to different technical needs.

TORINO, MAY 2017