ALMA MATER STUDIORUM - UNIVERSITÀ DI BOLOGNA

FACOLTA' DI INGEGNERIA

Corso di Laurea in Ingegneria per l'Ambiente e il Territorio

Insegnamento di Costruzioni Idrauliche e Protezione Idraulica del Territorio

DISTART - Dipartimento di Ingegneria delle Strutture, dei Trasporti, delle Acque, del Rilevamento, del Territorio

PROBABILISTIC ENVELOPE CURVES FOR EXTREMES RAINFALL EVENTS

CURVE INVILUPPO PROBABILISTICHE PER PRECIPITAZIONI ESTREME

TESI DI LAUREA DI:
Lorenza Tagliaferri

RELATORI:
Chiar.mo. Prof. Ing. Armando Brath
Chiar.mo Prof. Ing. Günter Blöschl

CORRELATORI:
Dott. Ing. Ralf Merz
Dott. Ing. Attilio Castellarin

Anno Accademico 2006/2007

Sessione III

Parole Chiave:

- Analisi di frequenza
- Austria
- Italia
- Precipitazioni estreme
- Curve inviluppo

Water is an image of the evolution and the cycle of life and flows over borders and breaks limits.

A genuine symbol against all preconceived tracks.

PANTHA REI – *All Flows*

INDEX

INTRODUZIONE 7

1 INTRODUCTION 10

2 METHODOLOGIES FOR THE DESIGN FLOOD ESTIMATION 14

2.1 INDIRECT METHODS 16
2.2 ESTIMATION OF THE DESIGN RAINFALL EVENT 17
2.3 OTHER FLOOD ESTIMATION METHODS 19

3 FREQUENCY ANALYSIS OF EXTREME EVENTS 20

3.1 PROBABILITY CONCEPTS 20
3.2 PROBABILITY DISTRIBUTIONS FOR EXTREME EVENTS 24
3.2.1 The Normal Distribution 24
3.2.2 The Lognormal Distribution 26
3.2.3 The Gumbel Distribution 27
3.2.4 The Generalized Gumbel Distribution (GEV) 32
3.2.5 The Weibull Distribution 34
3.3 PLOTTING POSITIONS AND PROBABILITY PLOTS 34
3.3.1 Choice of Plotting Position 35

4 PROBABLE MAXIMUM PRECIPITATION 38

5 PROBABILISTIC ENVELOPE CURVES FOR EXTREME FLOODS 40

5.1 DEFINITION OF REGIONAL ENVELOPE CURVES 40
5.2 A PROBABILISTIC INTERPRETATION OF ENVELOPE CURVES 43
5.3 DESCRIPTION OF THE METHOD 43
5.3.1 Estimating the Effective Number of Sites 44
5.3.2 Evaluation of the Intersite Correlation 45
5.3.3 Estimating the Effective Sample Years of Data 46

6 STUDY REGIONS AND LOCAL REGIME OF RAINFALL EXTREMES 48

6.1 CLIMATE AND MORPHOLOGY OF AUSTRIA 48
6.2 AVAILABLE RAINFALL DATA 52
6.2.1 Peaks-Over-Threshold Procedure 53
6.2.2 Identification of Regions 55
6.3 CLIMATE AND MORPHOLOGY OF ITALIAN REGION 58

6.4 AVAILABLE RAINFALL DATA 63

7 METHODOLOGY 65

7.1 AUSTRIAN CATCHMENTS 65

7.1.1 Uncorrelated Case 66

7.1.2 Correlated Case 68

7.2 ITALIAN CATCHMENTS 73

7.2.1 Analysis and Modelling of Intersite Correlation 78

7.2.2 Validation Procedure 84

8 DISCUSSIONS AND CONCLUSIONS 86

CONCLUSIONI 89

APPENDIX 92

ANNEXES 94

REFERENCES 119

ACKNOWLEDGMENTS 121

INTRODUZIONE

Gli eventi naturali estremi, quali portate di piena, siccità, forti piogge, mareggiate, terremoti o venti di particolare intensità, possono generare conseguenze catastrofiche per l'uomo e per la società. E' pertanto evidente come la stima della frequenza di accadimento di un particolare evento sia un problema di grande importanza ed interesse scientifico.

Le attività di pianificazione e controllo di emergenze climatiche, le attività di protezione civile, il progetto di strutture di ingegneria civile, la gestione delle riserve naturali, il controllo ambientale e il calcolo per la protezione da rischi ambientali si fondano in buona misura sulla conoscenza del regime di frequenza di eventi estremi. La stima di tali frequenze non risulta però agevole in quanto gli eventi estremi sono rari per definizione e la loro osservazione è assai sporadica, a ciò va sommato il fatto che molto spesso le serie storiche disponibili hanno lunghezza assai limitata.

In particolare la valutazione di eventi pluviometrici di progetto è una problematica che desta molto interesse nell'idrologia. La comunità scientifica negli ultimi anni, ha dedicato numerosi sforzi alla messa a punto di tecniche affidabili, per la stima di portate fluviali o di altezze di precipitazione aventi assegnato livello di rischio.

Gli eventi idrologici di progetto (ad esempio portate e precipitazioni) rappresentano ipotetici eventi associati a una data probabilità di superamento, in genere espressa in termini di tempo di ritorno. Per esempio, l'informazione relativa alla portata di progetto è necessaria al fine di identificare le misure di protezione del territorio e delle costruzioni idrauliche dal rischio di esondazione.

Per quanto riguarda le portate estreme la letteratura documenta la diffusa utilizzazione di curve inviluppo regionali (RECs). Tali curve riassumono l'attuale limite di portate estreme verificatesi in una data regione. L'idea di limitare le portate sperimentate tramite una curva inviluppo è classica nell'ambito dell'idrologia e, per gli Stati Uniti, si rifà a *Jarvis* (1925), che presentò una REC basata sulle portate registrate su 888 siti

Solo 50 anni più tardi, *Crippen e Blue*, 1977, e *Crippen,* 1982 aggiorono lo studio fatto da *Jarvis* nel 1925 creando 17 diverse REC, una per ogni diversa regione idrologica statunitense, basate su un totale di 883 siti. *Matalas*, 1997 e *Vogel*, 2001 hanno mostrato come le REC identificate da *Crippen e Blue*, 1977, e *Crippen*, 1982 rappresentino il limite delle portate estreme anche nel periodo che va dal 1977 al 1994 per 740 dei 883 siti analizzati da *Crippen e Blue*.

Enzel et al., 1993 esaminarono le serie storiche di portate limitate da una REC per il bacino del fiume Colorado e videro che la stessa REC corrispondeva anche al limite delle stime delle paleo-portate disponibili per il bacino.

Lo sviluppo e la costruzione delle RECs non è rimasto confinato agli Stati Uniti; sono state sviluppate per l' Italia (*Marchett*i, 1955), per la Grecia occidentale (*Mimihou*, 1984), per il Giappone (*Kadoya*, 1992) e per altre regioni. Le RECs sono state usate per confrontare le portate manifestatesi negli Stati Uniti, in Cina e nel mondo da *Costa*, 1987 e, più recentemente, da *Herschy*, 2002.

Le RECs hanno continuato ad essere costruite e viste soprattutto quali resoconto delle portate manifestatesi, piuttosto quali strumento conoscitivo per il progetto di misure di tutela nei confronti di portate "catastrofiche". Si pensava che non ci fosse la possibilità di associare a una REC un valore di probabilità (*Crippen e Blue*, 1977; *Crippen*, 1982; *Vogel et al*., 2001).

Water Science and Technology Board, Commission on Geosciences, Environment and Resources (1999) sancirono che la determinazione della probabilità di superamento di una REC era difficile principalmente a causa della correlazione spaziale tra i siti. Di conseguenza le REC non avevano grande utilità nonostante l' U.S. Interagency Advisory Committee on Water Data (1986) ne riconobbero la necessità per "mostrare e riassumere i dati di portata estremi attualmente in accadimento".

Sarebbe pertanto interessante fornire un'interpretazione probabilistica di una REC, considerato anche il fatto che negli 80 anni passati da quando *Jarvis* (1925) ha proposto una curva inviluppo, nessuna interpretazione probabilistica è mai stata presa in esame.

Le precipitazioni costituiscono la componente del ciclo idrologico che maggiormente concorre alla formazione del deflusso superficiale, pertanto lo studio del loro regime di frequenza è un requisito essenziale per valutare il rischio idrogeologico in una determinata regione.

Scopo del presente lavoro di Tesi è l'estensione dell'interpretazione delle curve inviluppo probabilistiche agli eventi di precipitazione estremi, nei bacini italiani e austriaci. Il primo obiettivo dello studio consiste nell'identificazione di un'affidabile metodologia di valutazione, che ci consenta di fornire la stima accurata, in una regione precisa, dell'altezza di precipitazione relativa ad una data durata e posizione, per un assegnato tempo di ritorno, elevato o molto elevato.

In particolare tale analisi punta l'attenzione sulle curve inviluppo regionali, già sviluppate per le portate da *Castellarin et al*. (2003 e 2007). Gli autori hanno sviluppato uno stimatore empirico del tempo di ritorno *T* associato a una determinata REC che, il linea di principio, consente l'utilizzo delle RECs a fini ingegneristici di progetto in bacini strumentati e non.

Il seguente lavoro propone l'estensione del concetto di REC agli eventi estremi di precipitazione introducendo la Curva di Durata-Altezza di precipitazione (DDEC), definita come il limite superiore regionale di tutti gli eventi meteorici registrati per diverse durate di precipitazione (qui i dati rappresentano massimi annuali).

Si adatterà inoltre l'interpretazione probabilistica proposta per le REc alle DDEC e all'adattamento di tali curve per la stima dell'evento di pioggia corrispondente ad un elevato tempo di ritorno *T* e a una assegnata durata.

Saranno due i datasets nazionali considerati, la serie di picchi di precipitazione al di sopra di una soglia (POT) per le durate di 30 minuti, 1, 3, 6, 9, 12 e 24 ore ottenuti da 700 stazioni di registrazione in Austria, e le serie dei massimi annuali (AMS) per gli eventi meteorici di durata compresa tra i 15 minuti e le 24 ore raccolti da 220 stazioni di registrazione nell' Italia centro-settentrionale.

L'approccio alla REC proposto sarà diverso per le due distinte regioni, così come saranno diverse le modifiche apportate al metodo per la stima empirica della probabilità di superamento *p*.

1 INTRODUCTION

Extreme environmental events, such as floods, droughts, rainstorms, and high winds, may have severe consequences for human society. How frequently an event of a given magnitude may be expected to occur is of great importance. Planning for weather-related emergencies, design of civil engineering structures, reservoir management, pollution control, and insurance risk calculations, all rely on knowledge of the frequency of these extreme events. Estimation of these frequencies is difficult because extreme events are rare by definition and, therefore, available information are often very sparse or short.

In particular the evaluation of extreme design hydrological events is a fundamental and highly debated topic in hydrology. The scientific community gave great importance to the optimization of reliable techniques for the estimation of extreme rainfall events and resulting floods in the last decades.

Design hydrological events (e.g., floods and rainstorms) are hypothetical events associated with a given exceedance probability, generally expressed in terms of recurrence interval. For instance, design flood information is needed for the identification of flood protection measures in a river basin and the design of the related structures, such as levee systems that need to be designed in order to prevent failure or overtopping (hydrologic failure). Design floods are also required for planning measures and structures as well as for safety control of existing structures.

Concerning extreme floods, the literature documents the diffuse utilization of regional envelope curves (RECs). These curves summarize the current bound on our experience of extreme floods in a region. The idea of bounding our flood experience with an envelope curve is classical in hydrology and, for Unites States dates back to Jarvis [1925], who presented a REC based on record floods at 888 sites in the conterminous United States. Roughly 50 years later, *Crippen and Bue*, 1977 and *Crippen*, 1982 updated the study by *Jarvis*, 1925 by creating 17 different RECs, each

for a different hydrologic region within the United States, based on a total of 883 sites. *Matalas,* 1997 and *Vogel et al*., 2001 document that the

RECs identified by *Crippen and Bue,* 1977 and *Crippen,* 1982 still bound our flood experience gained from 1977– 1994 at 740 of the 883 sites compiled by *Crippen and Bue*. *Enzel et al.,* 1993 examine the REC bounding the historical flood experience for the Colorado river basin and show that the same REC also bounds the paleoflood discharge estimates available for the basin.

The construction of RECs is not confined to the United States; they have been developed for Italy (*Marchetti, 1955*), western Greece *(Mimikou, 1984)*, Japan (*Kadoya, 1992),* and elsewhere. RECs have been used to compare record flood experience in the United States, China, and the world by *Costa*, 1987 and, more recently, by *Herschy*, 2002.

The REC provides an effective summary of our regional flood experience. The pioneering work of *Hazen,* 1914, who formalized flood frequency analysis, a formalism still in use, and who was among the first to suggest a method for improving information at a site through the transfer of information from other sites (i.e., substitution of space for time), has tempered the use of a flood magnitude as a design flood without an accompanying probability statement. Our objective is to provide a probabilistic interpretation of the REC. In the almost 80 years since Jarvis [1925] introduced the envelope curve, a probabilistic interpretation of a REC has never been seriously addressed.

RECs have continued to be constructed and viewed mainly as summary accounts of record floods, rather than as meaningful tools for the design of measures to protect against ‘‘catastrophic’’ floods. It has been suggested that there is no obvious way to assign a probabilistic statement to a REC (*see, e.g., Crippen and Bue, 1977; Crippen, 1982; Vogel et al., 2001*). Water Science and Technology Board, Commission on Geosciences, Environment and Resources (1999) argued that the determination of the exceedance probability of a REC is difficult due to the impact of intersite correlation. As a consequence, RECs are assumed to have little utility beyond the suggestion of the U.S. Interagency Advisory Committee on Water Data (1986) that they are useful for ‘‘displaying and summarizing data on the actual occurrence of extreme floods.’’ A probabilistic interpretation of the REC offers opportunities for several engineering

applications which seek to exploit regional flood information to augment the effective record length associated with design flood estimates.

A potential advantage of assigning a probabilistic statement to a REC is that this approach avoids the need to extrapolate an assumed at-site flood frequency distribution hen estimating a design event.

Rain is the most important component of the water cycle with respect to the formation of runoff for a wide portion of European climates, the analysis of its frequency regime is fundamental for the assessment of flooding potential of a given area.

Aim of this work is the extension of the probabilistic interpretation of regional envelope curves to extremes rainfall events in Italian and Austrian catchments. The primary objective of the present study is the identification of a reliable methodology that enables us to accurately evaluate the rainfall depth for a given duration and location associated with high and very high recurrence intervals.

In particular this study focuses on probabilistic regional envelope curves proposed by *Castellarin et al.,* (*2005*) and *Castellarin* (*2007*) for flood flows. The authors formulated an empirical estimator of the recurrence interval *T* associated with a given REC, which, in principle, enables us to use RECs for design purposes in ungauged basins.

This work extends the REC concept to extreme rainstorm events by introducing the Depth-Duration Envelope Curves (DDECs). DDECs are defined as the regional upper bounds on all the record rainfall depths at present for various rainfall duration (here record indicates historical maxima). It also adapts the probabilistic interpretation proposed for RECs to DDECs and it assesses the suitability of these curves for estimating the *T*-year rainfall event associated with a given duration and large *T* values

The study focuses on two different national datasets, the peak over threshold (POT) series of rainfall depths with duration 30 minutes, 1, 3, 6, 9 and 24 hours obtained for 700 Austrian raingauges and the Annual Maximum Series (AMS) of rainfall depths with duration spanning from 15 minutes to 24 hours collected at 220

raingauges located in northern-central Italy. The approach to the REC is different for the two catchments as well as the adjustment of the empirical estimator for the exceedance probability p.

2 METHODOLOGIES FOR THE DESIGN FLOOD ESTIMATION

This paragraph aims to presenting the possible methodologies for the design-flood estimation. This task, as already mentioned, is fundamental in the evaluation of hydrologic risk, and in particular, of flooding potential of a given site.

Both floods and rainfall data can be used as hydrological variables for the estimation of the design flood. The identification of an hydrological variable is developed by two different methodologies, the direct methods that directly calculate the floods and the indirect methods that identify the rainfalls to be given as input in a rainfall-runoff model, able to convert the rainfalls into floods thanks to appropriate calibrations.

Although this work gives particular attention to rainfall events, we are going to illustrate the common methodologies for the design flood estimation.

A common approach for estimating design flood consists of modeling hydrological events as random variables, allowing the determination of the flood exceeded with given probability. Usually the problem is that of information: if one had a sufficiently long record of flood flows, rainfall, low flows or pollutant loadings, then a frequency distribution for a site could be precisely determined, so long as change over time due to urbanization or natural processes did not alter the relationship of concern. In most situations, available data are not sufficient to precisely define the risk of large floods, rainfall, pollutant loadings or low flows.

A fundamental step for the identification of a design event is the evaluation of the relation $x = x(T)$. Then the design event is identified by selecting the return period T believed to be adapt to the importance of the structure itself and to the implications that its failure would involve.

The *direct methods* determine the expression $x = x(T)$, from the available rainfall data analysis in neighbourhood catchments, otherwise from the extrapolation of

statistical analysis known for neighbourhood catchments, similar from the hydrological point of view.

The *indirect methods* derive the same expression from the statistical analysis of the extreme rainfall events occurring in the hydrographical catchment, applying later a transformation of these rainfall events in superficial runoff, that means they evaluate the design flood from extreme rainfall events.

The choice between the two methods can be made according to the reliability and validity of the available rainfall data. The use of direct methods requires hydrometrical observations of annual maxima floods for a sufficiently long period of time; this information is later analysed by statistical techniques in order to indentify the probability distribution that better represents the collected data. As described direct methods seem to be the more suitable for their validity and simplicity in application: the variable on which the measurement are mead is the flood itself that is the variable to be estimate for design purposes.

In practical the problem stands in their application: measures are often uncertain or related to hydrometric sections different from that of design, thus their utilization requires extrapolations and hypothesis not easy to formulate.

Moreover, rarely it is possible to arrange sufficiently long flood data sets (20-30 years) able to allow the analysis of return periods connected with the design. If is T larger than the mean length of the available series, the error in the corresponding flood estimation can be significant and consequentially the extrapolations became to have uncertain reliability (it would be better not to have estimations of T greater than 2-3 times the series length (*Cunnane, 1986*)). Series of limited length can be represented by different probability distributions, from which the obtained extrapolations are very different.

As said, a possible alternative is the recurs to the indirect methods which, based on the rainfall-runoff transformation, are strong of the statistical distributions of the rainfall data, much more than the hydrometrical data. Anyway, the transformation of the design rainfall event in design flood needs the formulation and estimation of model parameters which have the aim of simulating soil infiltration, storage in surface and interception by vegetation. Such complex models introduce great

uncertainty being a rough approximation of realty. It is for this reason important, for the using of indirect methods, the accurate knowledge of the hydrological characteristics of the analysed catchment.

Definitely, the indirect methodology has the power of being correct for a great number of cases, much more than the direct methods, thank to the density of the rainfall net. This typology presents all the same some limiting factors:

- The difficult in arranging sufficient rainfall data for the calibration of the rainfall-runoff model, model that can be very complex;
- Supposing the same frequency for rainfall events and floods, the method gives the same probability to floods and rainfall events, assumption that can be criticised;

In spite of these critics, the indirect one is by now a valid methodology, fundamental for drainage systems and hydrographical nets.

Next paragraph focuses on the indirect methodology.

2.1 INDIRECT METHODS

A different approach is to view flood flows as the product of a deterministic transformation of rainfall events, seen as random variables. Because is rain cause of floods, it is also supposed to be strictly related to them.

The consequent transformation of rainfall in floods will then entrusted to rainfall-runoff models. Several are the typologies of models, all able to represent hydrological phenomena that make the basin as a deterministic system by which rainfall events and floods becomes inputs and outputs.

Different models can be distinguished according to the applied transformation model (event-based or continuous simulation) and to the way the design rainfall event is represented.

We will focus on this representation of rainfall events, without giving importance to the rainfall-runoff simulation.

2.2 ESTIMATION OF THE DESIGN RAINFALL EVENT

Input to a rainfall-runoff model can be design hyetographs, historical series of rainfall data, or synthetic series of rainfall events generated by stochastic rainfall models: each input is connected to a different probabilistic structure thus a different bond between *x* and *T*. According to the defined input indirect methods can be classified as:

- Methods based on depth-duration-frequency curves
- Methods derived by simulation
- Analytic-derived methods

The first ones consist of using a single event model as input to the rainfall-runoff model; these give the bond between the rainfall depth h with occurrence period *T* and duration *t*, and the duration t itself. These methods can be found by statistical interference methods by the analysis of the extreme historical series in the neighbourhood of the region of concern. Data are supplied by the National Hydrographical Service of Italy (SIMN) that provide series of maximum rainfall depth data for durations of days or less than a day. For each duration are estimated the parameters of the suitable probability distribution, so to identify the correct quantiles of rainfall depth. Rainfall depths for durations different from the estimates ones, will be calculated by interpolation of the obtained data.

Some other lows are useful for expressing this expression, in Italy we find an exponential relation that describes the grow of the expected rainfall in a certain site with the growing of the duration *t*, function of the return period *T*:

$$h(t,T) = a_T \cdot t^{n_T} \qquad (2.2\text{-}1)$$

where h represents the rainfall depth, and a_T and n_T are the parameters to be estimates as function of the return period.

Figure 2.2-1 shows that equation (2.2-1) assumes a convex exponential form on a *t*/*h* graph, actually for a given frequency the rainfall rate decreases with duration and the cumulative rain also progressively decreases whit the duration itself.

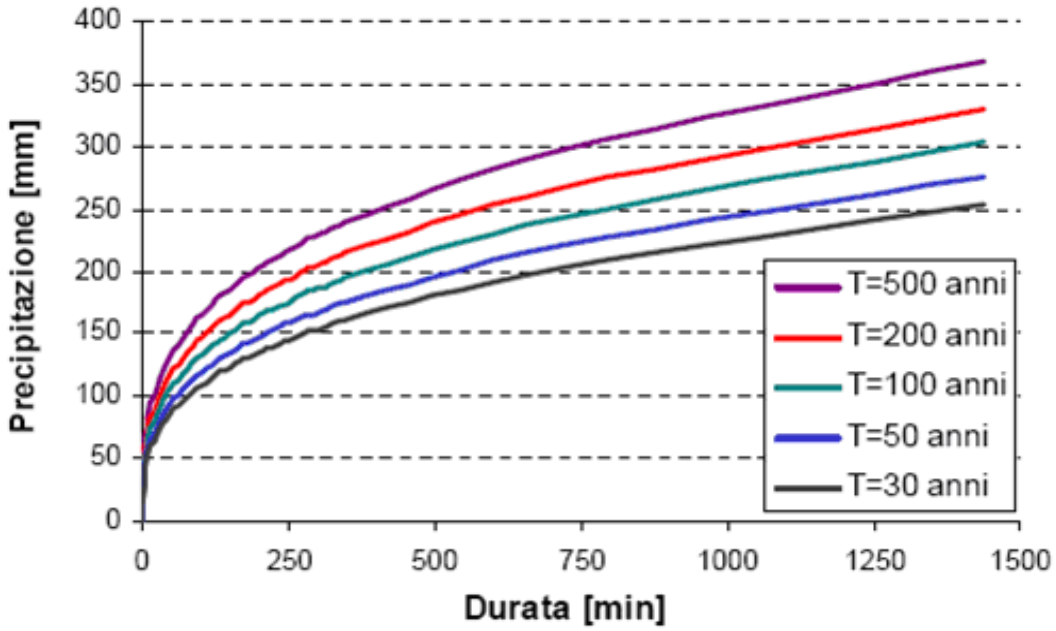


Figure 3.2.1-1 Representation of the rainfall depth versus duration for different values of return period, it is easily recognizable a convex exponential form

Limit of this method is that for each single site is important to define a depth-duration-frequency curve that require the estimation of many parameters: two parameters for each duration, and two for deriving the exponential law. Considering that the length of the series and the available data are not sufficient, usually smaller than 50-60 years of registration, the estimation of the parameters is not possible.

In order to supply the lack of information, some simplifications can be introduced. For example an constant scale that means rainfall depth for the duration *t* equal to one corresponding to a given duration, simply multiplied by an estimated scale factor. Although this assumption, not always applicable and to be verified, there is the problem of obtaining rainfall data for ungauged basins: we need a more general expression applicable on regional scale. The introduction of envelope curves let the extension of information to ungauged basins, in the next paragraphs this concept will be studied in depth.

Methods derived by simulation use the rainfall data for a limited number of years to generate, by appropriate rainfall stochastic models, synthetic rainfall series of greater dimensions to be used as input in rainfall-runoff models that give as output

simulated series of floods. By applying to these series the usual statistical techniques we will obtain estimations of design floods. Definitely the design flood is estimated by a simulated series of rainfall data, not really observed. Thus the method can be applied to generate simulated series of given length , also 1000 years, without having such a complete database.

Analytic-derived methods use simplified assumptions for the transformation process of rainfall in runoff, in order to be able to analytically derive the statistical proprieties of the runoffs as functions of similar proprieties in rainfall events.

2.3 OTHER FLOOD ESTIMATION METHODS

Besides the methods mentioned above, many others methodologies have been developed for estimating design floods. Another approach that can be very useful is based on consideration of the largest floods that have been observed in the region of interest. The usual procedure is to draw an *envelope curve* on a regional plot of maximum recorded flood at each gauging station against drainage basin area. Logarithmic values are normally plotted with discharge in m^3/km^2. The graph provides a useful summary of flood experience in a region. Plotting and labelling of the maximum flood for each drainage basin makes the scatter of the data obvious. Trends in flood characteristics in a region can be examined, as with elevation, latitude, stream slope, distance from the ocean and other moisture source, or different record length.

It may be possible to draw envelope curves for different subregions. However, as time proceeds, higher floods are recorded and the envelope curve moves to higher discharges. Probabilities of floods can not be estimated objectively by this method. Anyway where data are sparse and other methods can not be used, envelope curves are better employed for either checking that estimates by other methods are of the correct order of magnitude of providing preliminary estimates.

3 FREQUENCY ANALYSIS OF EXTREME EVENTS

At present the description of design flood derives from a probabilistic approach that models hydrological events as random variable, allowing the determination of the flood exceeded with given probability. Usually the problem is that of information: in most situations, available data are not sufficient to precisely define the risk of large floods, rainfall, pollutant loadings or low flows

Frequency analysis is an information problem: in order to supply the lack of information explained above, such as a not sufficiently long record of flood flows or rainfall, hydrologist are forced to use practical knowledge of the processes involved and efficient and robust statistical techniques, to develop the best estimator they can.

These techniques generally restricted, with 10 to 100 sample observations to estimate events exceedance with a chance of at least 1 in 100, corresponding to exceedance probability of 1 per cent or more.

The hydrologist should be aware that in practice the true probability distribution of phenomena in question are not known. Even if they were , their functional representation would likely have too many parameters to be of much practical use.

The practical issue is how to select a reasonable and simple distribution to describe the phenomena of interest, to estimate that distribution's parameters, and thus to obtain risk estimates of satisfactory accuracy for the problem at hand.

3.1 PROBABILITY CONCEPTS

We introduce here some probabilistic concepts that stands at the basis of frequency analysis.

Let the upper case letter X denote a *random variable*, and the lower case letter x be a possible value of X. For a random variable X, its *Cumulative Distribution* Function (cdf), denoted *Fx(x)*, is the probability the random variable X is less than or equal to x:

$$Fx(x) = P(X \leq x)$$

Fx(x) is the nonexceedance probability for the value x.

Continuous random variables take on values in continuum. For example, the magnitude of floods and low flows is described by positive real values, so that $X \geq 0$.

The *Probability Density Function* (pdf) describes the relative likelihood that a continuous random variable X takes on different values, and is the derivative of the cumulative distribution function:

$$fx(x) = \frac{dFx(x)}{dx}$$

In hydrology the *percentiles* or *quantiles* of a distribution are often used as design events. The 100*p* percentile or the *p*th quantile x_p is the value with cumulative probability p:

$$Fx(x) = p$$

The 100*p* percentile x_p is often called the *100(1-p)* percent exceedance event because it will be exceeded with probability *1-p.*

The *Return Period* (sometimes called *Recurrence Interval*) is often specified rather than the exceedance probability. For example, the annual maximum flood-flow exceedance with a 1 percent probability in any year, or chance 1 in 100, is called the 100-year flood. In general, x_p is the *T*-year flood for

$$T = \frac{1}{1-p}$$

Here are two ways that the return period can be understood. First , in a fixed x_p *T*-year period the expected number of exceedance of the *T*-year event is exactly 1, if the distribution of floods does not changeover that period; thus on average one flood greater than the *T*-year flood level occurs in a *T*-years period.

Alternatively, if floods are independent from year to year, the probability that the first exceedance of level x_p occurs in year k is the probability of (k-1) years without an exceedance followed by a year in which the value of X exceeds x_p:

$$P \text{ (exactly k years until } X \geq x_{p)} = p^{k-1} (1-p)$$

This is a geometrical distribution with mean 1 / (1-*p*). Thus the *average* time until the level x_p is exceeded equals *T* years. However, the probability that x_p is not exceeded in a *T*-year period is $p^T = (1-1/T)^T$, which for $1/(1-p) = T \geq 25$ is approximately 36,7 percent, or about a chance of 1 in 3.

Return period is a means of expressing the exceedance probability. Hydrologists pften speak of the 20-year flood or of the 1000-year rainfall, rather than events exceeded with probabilities of 5 or 0,1 percent in any year, corresponding to chances of 1 in 20, or 1 in 1000. Return period has been incorrectly understood to mean that one and only one *T*-year event should occur every *T* years. Actually , the probability of the *T*-year flood being exceeded is 1/*T* in every year. The awkwardness of small probabilities and the incorrect implication of return periods can both be avoided by reporting *odds ratios*: thus the 1 percent exceedance event can be described as a value with a 1 in 100 chance of being exceeded each year.

Several summary statistics can describe the character of the probability distribution of a random variable. Moments and quantiles are used to describe the location or central tendency of a random variable, and its spread.

The *mean* of a random variable *X* is defined as

$$\mu x = E[X]$$

The second moment about the mean is the *variance*, denoted Var(*X*) or σx^2 where

$$\sigma_x^2 = \text{Var}(X) = E[(X - \mu x)^2]$$

The standard deviation σx is the square root of the variance and describes the width of scale of a distribution. These are examples of *product moments* because they depend upon powers of *X*.

A dimensionless measure of the variability in *X*, appropriate for use with positive random variables $X \geq 0$, is the coefficient of variation defined as

$$CVx = \frac{\sigma x}{\mu x} \qquad (3.1)$$

Also the coefficient of skewness γx is a dimensionless measure that describes the relative asymmetry of a distribution, it is defined as

$$\gamma x = \frac{E(X - \mu x)^3}{\sigma_x^3} \qquad (3.12)$$

and the coefficient of kurtosis which describes the thickness of a distribution's tails:

$$\frac{E(x - \mu x)^4}{\sigma_x^4} \qquad (2.1.3)$$

From a set of observations (*X1, …, Xn*) the moments of a distribution can be estimated. Estimators of the mean, variance and coefficient of skewnwss are

$$\overline{\mu x} = \overline{X} = \sum_{i=1}^{n} \frac{Xi}{n} \qquad (2.1.4)$$

$$\overline{\sigma_x^2} = S^2 = \frac{\sum_{i=1}^{n} (Xi - \overline{X})^2}{n-1} \qquad (2.1.5)$$

$$\overline{\gamma x} = G = \frac{n\sum_{i=1}^{n} (Xi - \overline{X})^3}{(n-1)(n-2)S^3} \qquad (2.1.6)$$

When data vary widely in magnitude, as often happens, the sample product moments of the logarithms of the data are often employed to summarize the characteristics of a data set or to estimate parameters of distributions. A logarithmic transformation is an effective vehicle for normalizing values which vary by orders of magnitude, and also for keeping occasionally large values from dominating the calculation of product-moments estimators. However, the danger with use of logarithmic transformations is that unusually small observations (or low outliers) are given greatly increased weight. This is a concern if it is the large events that are of interest, small values are poorly measured, small values reflect rounding, or small values are reported as zero if they fall below some threshold.

3.2 PROBABILITY DISTRIBUTIONS FOR EXTREME EVENTS

This paragraph provides descriptions of several families of distributions commonly used in hydrology. These include the normal/lognormal family, the Gumbel/Weibull generalized extreme value family, and the exponential/Pearson/log Pearson type 3 family. Table 2.1-1 in Appendix provides a summary of the pdf and cdf of these probability distributions, and their means and variances.

Many other distributions have also been successfully employed in hydrologic applications.

3.2.1 The Normal Distribution

The normal distribution is useful in hydrology for describing well-behaved phenomena such as average annual stream flow, or average annual pollutant loadings. The *central limit theorem* demonstrated that if a random variable *X* in the sum of *n* independent and identically distributed random variables with a finite variance, then with increasing *n* the distribution of *X* becomes normal regardless of the distribution of the original random variables.

The pdf for a normal random variable *X* is

$$fx(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu x}{\sigma x} \right)^2 \right]$$

X is unbounded both above and below, with mean μx and variance σx^2. The normal distribution's skew coefficient is zero because the distribution is symmetric. The product-moment coefficient of kurtosis, $E[(X-\mu x)^4]/\sigma^4$, equals 3.

The two moments of the normal distribution, μx and σx^2 are its natural parameters. They are generally estimates by the sample mean and variance in Eq. (2.1.4) and (2.1.5); these are the maximum likelihood estimates if (*n*-1) is replaced by *n* in the denominator of the sample variance. The cdf of the normal distribution is not available in closed form. Selected points *zp* for the *standard normal distribution*

with zero mean and unit variance are given in Table 2.1-1; because the normal distribution is symmetric, z_p=-z_{1-p}.

p	0.5	0.6	0.75	0.8	0.9	0.95	0.975	0.99	0.998	0.999
z_p	0.000	0.253	0.675	0.842	1.282	1.645	1.960	2.326	2.878	3.090

Table 3.2-1 Quantiles of the Standard Normal Distribution

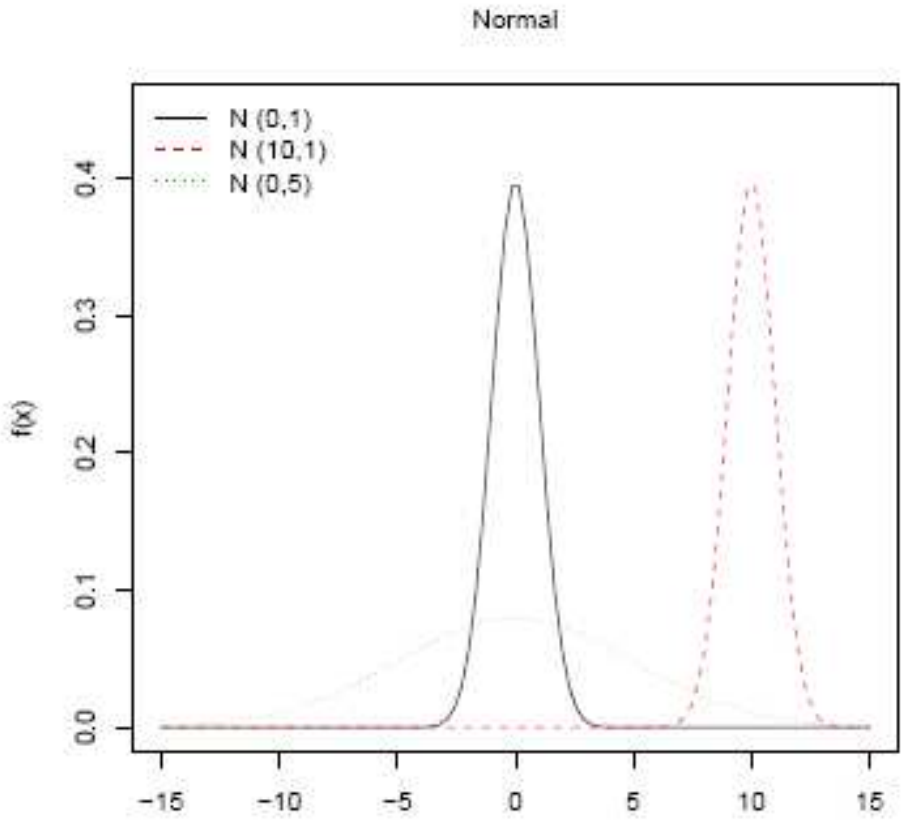


Figure 3.2.1-1 Effect of parameters on the Normal pdf, we consider (1) μ = 0, σ = 1; (2) μ = 10, σ = 1; (3) μ = 0, σ = 5.

3.2.2 The Lognormal Distribution

Many hydrologic processes are positively skewed and are not normally distributed. However, in many case for *strictly positive random variables X*>0, their logarithm

$$Y = \ln (X)$$

is described by a normal distribution. This is particularly true if the hydrologic sample results from some multiplicative processes, such as dilution.

Inverting the equation above:

$$X = \exp (Y)$$

If *X* has a lognormal distribution, the cdf for *X* is

$$Fx(x) = P(X \le x) = P[Y \le \ln(x)] = P\left[\frac{Y - \mu y}{\sigma_Y} \le \frac{\ln(x) - \mu y}{\sigma_Y}\right] = \Phi\left[\frac{\ln(x) - \mu y}{\sigma_Y}\right]$$

where Φ is the cdf of the standard normal distribution.

As a function of the coefficient of variation CV*x*, the skew coefficient is

$$\gamma x = 3\mathrm{CV}x + \mathrm{CV}^3 x$$

As the coefficient of variation and skewness go to zero, the lognormal distribution approaches a normal distribution.

The sample mean and variance of the observed (*yi*) obtained by using Eq. (2.1.4) and (2.1.5) are the maximum likelihood estimators of the lognormal distribution's parameters if (*n*-1) is replaced by *n* in the denominator of S^3y.

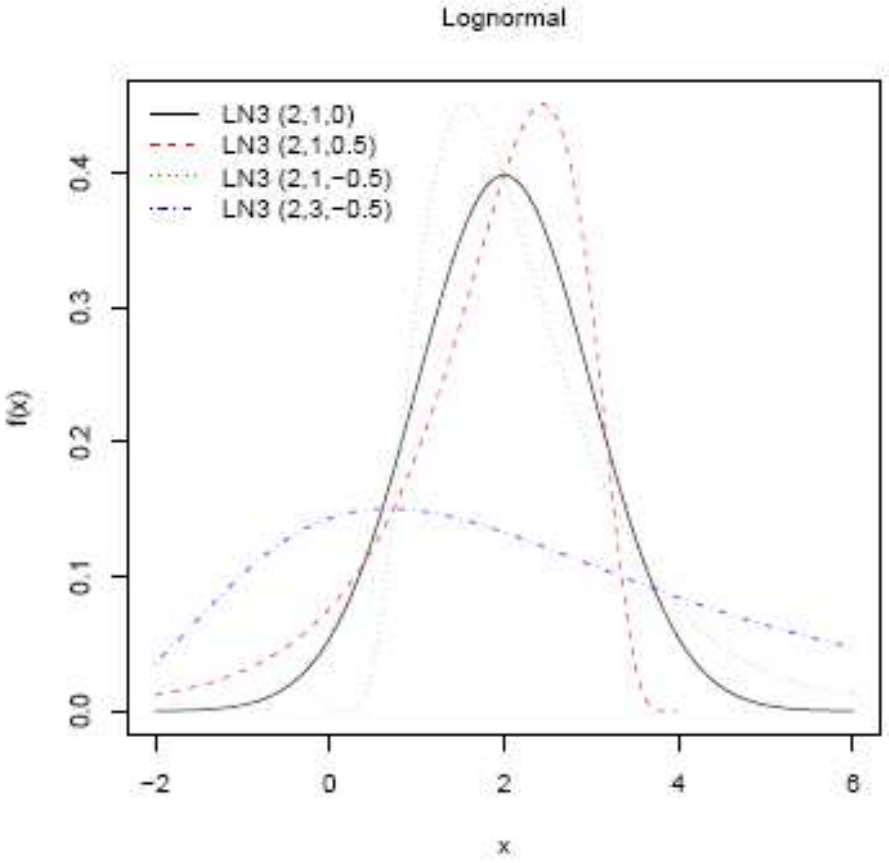


Figure 3.2.2-1 **2** Effect of parameters on the Normal pdf, we consider (1) ξ = 2, α = 1, k = 0; (2) ξ = 2, α = 1, k = 0. 5; (3) ξ = 2, α = 1, k = − 0. 5; (4) ξ = 2, α = 3, k =− 0. 5.

3.2.3 The Gumbel Distribution

Many random variables in hydrology correspond to the maximum of several similar processes, such as the maximum rainfall or flood discharge in a year. The physical origin of such random variables suggest that their distribution is likely to be one of several *extreme value* (*EV*) *distributions* described by Gumbel. The cdf of the largest of *n* independent variates with common cdf F(x) is simply F(x)^n. For large n and many choices for F(x), F(x)^n converges to one of three extreme value distributions, called type I,II and III. Unfortunately, for many hydrologic variables this convergence is too slow for this argument alone to justify adoption of an extreme value distribution as a model of annual maxima and minima.

The *EV* type I distribution is called *Gumbel Distribution*.

Let M_1, …, M_n be a set of daily rainfall data, and let the random variable X=max(M_i) be the maximum for he year. If the M_i are independent and identically distributed random variables unbounded above, with an exponential like upper tail, then for large n the variate X has an extreme value type Gumbel distribution. For example the annual maximum 24-h rainfall depth are often described by a Gumbel distribution.

The Gumbel distribution has the cdf, mean and variance described in Appendix,

the Gumbel pdf is

$$f(x) = \frac{1}{\alpha} e^{\frac{x-\xi}{\alpha}} e^{-e^{\frac{x-\xi}{\alpha}}}$$

where ξ is the location parameter and α is the scale parameter.

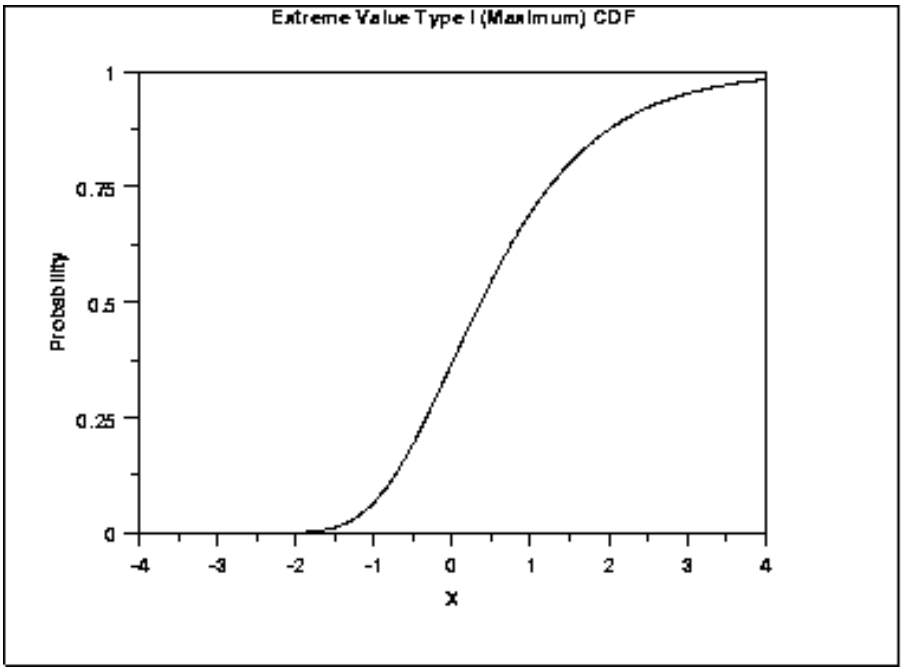
The case where ξ =0 and α =1 is called the *standard Gumbel distribution*. The equation for the standard Gumbel distribution (minimum) reduces to

$$f(x) = e^x e^{-e^x}$$

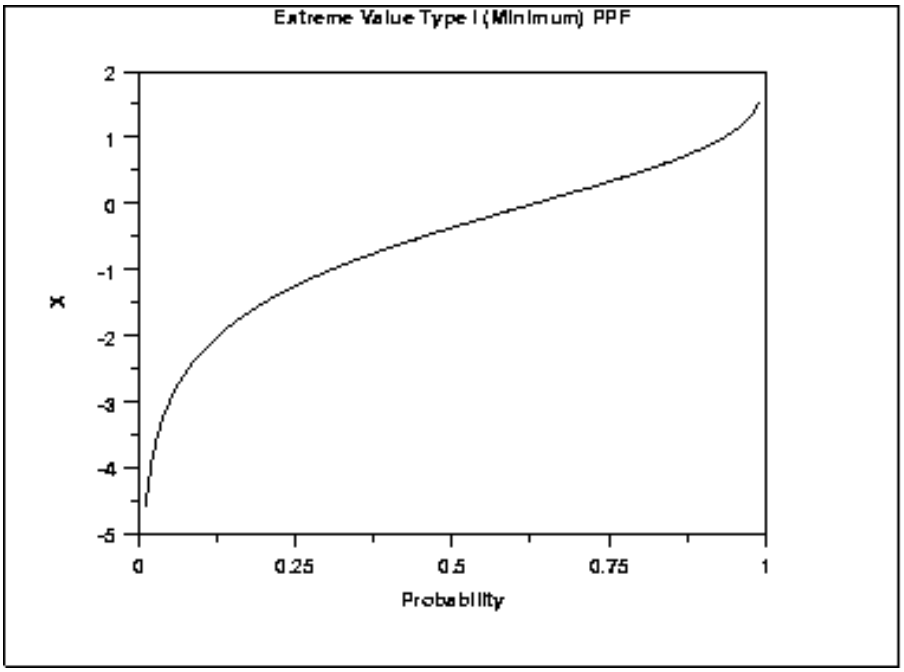
The Gumbel distribution's density function is very similar to that of the lognormal distribution with γ=1,13. Changing ξ and α moves the centre of the Gumbel pdf and changes its width, but does not change the shape of the distribution. The Gumbel distribution is asymptotically equivalent to the exponential distribution with cdf :

$$Fx(x) = \exp\left[-\exp\left(-\frac{x-\xi}{\alpha}\right)\right]. \quad (3.2.3\text{-}5)$$

The following is the plot of the Gumbel cumulative distribution function for the maximum case.



While the following is the plot of the Gumbel percent point function for the minimum case.



The cdf is easily inverted to obtain

$$x_p = \xi - \alpha \ln[-\ln(p)] \qquad (3.2.3\text{-}1)$$

The estimator of α obtained by using the second sample L moment is

$$\hat{\alpha} = \frac{\hat{\lambda}_2}{\ln(2)} = 1{,}443\hat{\lambda}_2 \qquad (3.2.3\text{-}2)$$

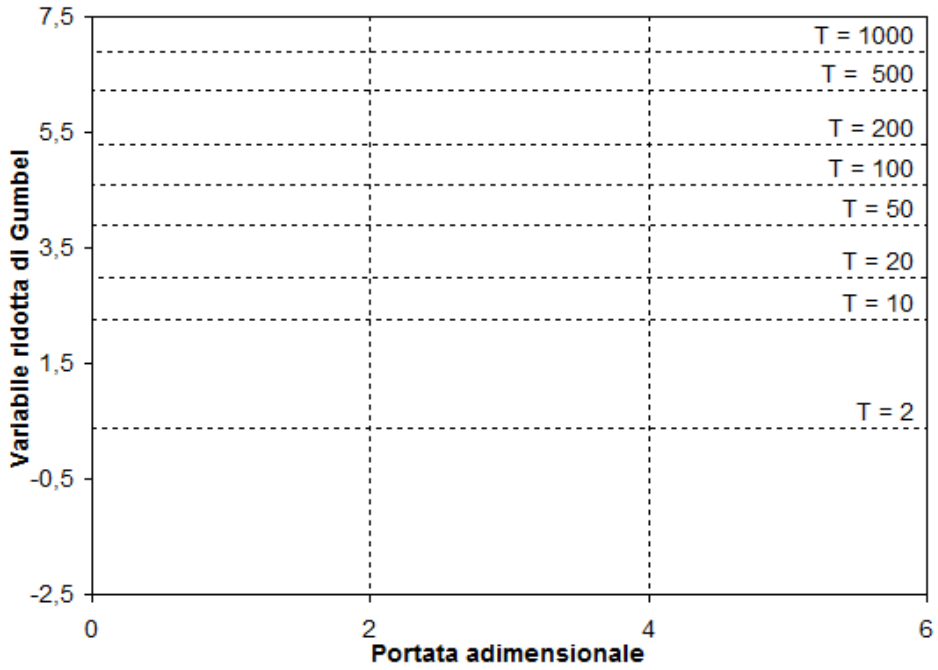
If the sample variance s^2 was employed, one obtains

$$\hat{\alpha} = \frac{s\sqrt{6}}{\pi} = 0{,}7797s \qquad (3.2.3\text{-}3)$$

The corresponding estimator of ξ in either case is

$$\hat{\xi} = \bar{x} - 0{,}5772\hat{\alpha} \quad (3.2.3\text{-}4)$$

The form of the Gumbel probability paper is based on a linearization of the cdf. From Equation (3.2.3-5), the Gumbel probability paper resulting from this linearized cdf function is shown next.



L-moments estimators for the Gumbel distribution are generally so good or better than method-of-moment estimators when the observations are actually drawn from case. However, L-moments estimators have been shown to be robust , providing more accurate quantile estimators than product moment and maximum-likelihood estimators when observations are drown from a range of reasonable distributions for flood flows.

Some of the specific characteristics of the Gumbel distribution are the following:

- The shape of the Gumbel distribution is skewed to the left. The Gumbel pdf (*Gumbel probability density function*) has no shape parameter. This means that the Gumbel pdf has only one shape, which does not change.
- The Gumbel pdf has location parameter μ, which is equal to the mode, but it differs from median and mean. This is because the Gumbel distribution is not symmetrical about its ξ.
- As ξ decreases, the pdf is shifted to the left.
- As ξ increases, the pdf is shifted to the right.

- As σ increases, the pdf spreads out and becomes shallower.
- As σ decreases, the pdf becomes taller and narrower.
- For $T = \pm\infty$, pdf = 0. For T = ξ, the pdf reaches its maximum point $\left(\frac{1}{\sigma e}\right)$

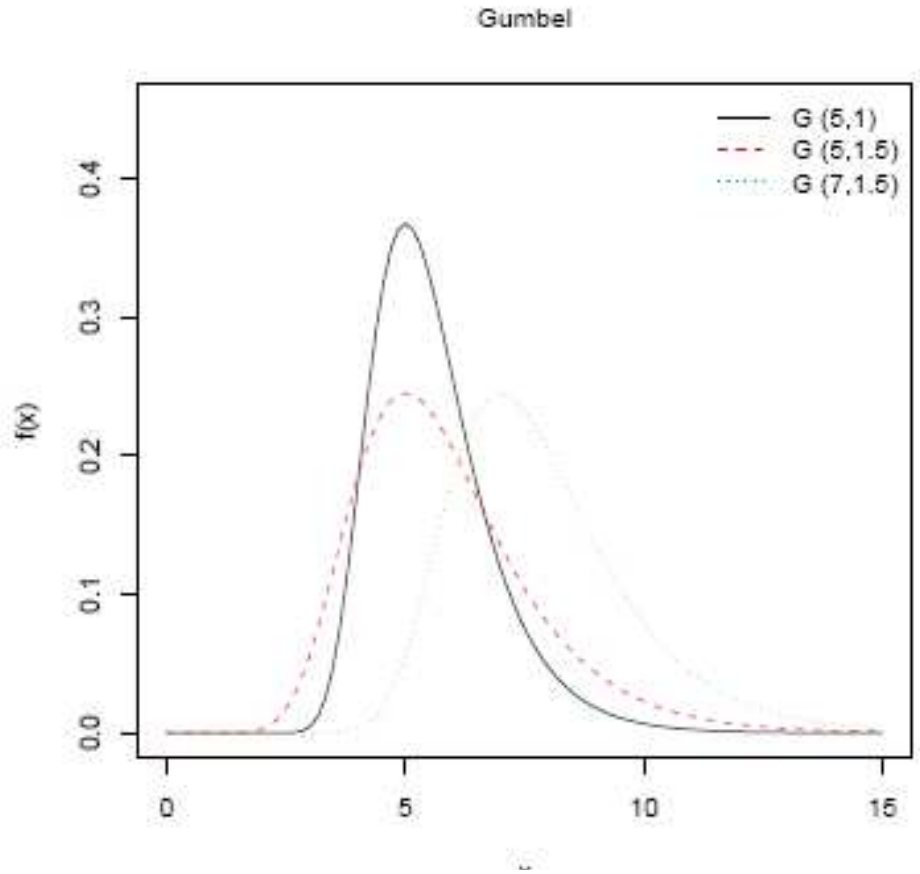


Figure 3.2.3-1 Effect of parameters on Gumbel pdf, we consider (1) ξ = 5, α = 1; (2) ξ = 5, α = 1. 5; (3) ξ = 7, α = 1. 5

3.2.4 The Generalized Gumbel Distribution (GEV)

The *Generalized Extreme Value* (GEV) *Distribution* is a general mathematical formulation which incorporates Gumbel's type I,II and III extreme value distributions for maxima. The GEV distribution's cdf can be written

$$Fx(x) = \exp\left\{-\left[1-\frac{k(x-\xi)}{\alpha}\right]^{1/k}\right\}$$

The Gumbel distribution is obtained when k=0. For |k|<0,3 the general shape of the GEV distribution is similar to the Gumbel distribution, though the right-hand tail is thicker for k<0 and thinner for k>0.

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

The moments of the GEV distribution can be expressed in terms of the gamma function Γ(·), or k>-1/3 the mean and variance are given in Appendix.

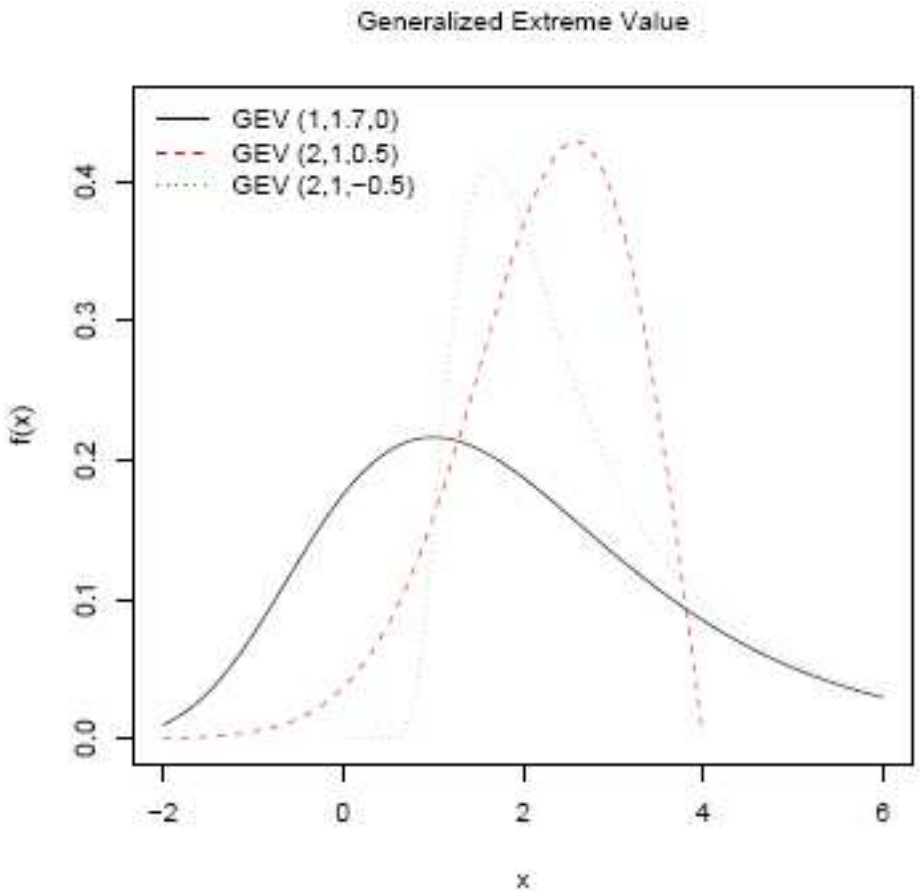


Figure 3.2.4-1 Effect of parameters on GEV pdf, we consider (1) ξ = 1, α = 1.7, k = 0; (2) ξ = 2, α = 1, k = 0.5; (3) ξ = 2, α = 1, k = −0.5.

3.2.5 The Weibull Distribution

The *Weibull Distribution* is developed as the extreme value type III distribution for minima bounded below by zero.

If *Wi* are the minimum stream flows in different days of the year, then the annual minimum is the smallest of the *Wi*, each of which is bounded below by zero. In this case the random variable X=min(Wi) may be well described by the EV type III distribution for minima, the Weibull distribution's cdf, mean and variance are included in Appendix. The skewness coefficient is the negative of that of GEV distribution, the second L moment is

$$\lambda_2 = \alpha(1-2^{-1/k})\ \Gamma(1+1/k)$$

There are important relationship between the Weibull, Gumbel and GEV distributions. If X has a Weibull distribution, then Y= -ln(X) has a Gumbel distribution. This allows parameter estimation procedures and goodness-on-fit tests available for the Gumbel distribution to be used for the Weibull, thus if ln(X) has mean $\lambda_{1,(\ln X)}$ and L-moment $\lambda_{2,(\ln X)}$, X has Weibull parameters

$$k=\ln(2)/\ \lambda_{2,(\ln X)} \qquad \text{and} \qquad \alpha = \exp(\lambda_{1,(\ln X)} + 0{,}5772/k)$$

3.3 PLOTTING POSITIONS AND PROBABILITY PLOTS

Probabilistic plots are extremely useful visually revealing the character of data sets. The graphical evaluation of the adequacy of a fitted distribution is generally performed by plotting the observations so that they would fall approximately on a straight line if a postulated distribution were the true distribution from which the observations were drown. This can be done with the use of special commercially available probability papers for some distributions, or with the more general techniques presented here, on which such special papers are based.

Let $\{Xi\}$ denote the observed values and $X_{(i)}$ the ith largest value in a sample, so that $X_{(n)} \leq X_{(n-1)} \leq \ldots \leq X_{(i)}$. The random variable U_i defined as

$$U_i = 1\text{-Fx}[X_{(i)}] \qquad (2.1.3.1)$$

correspond to the the *exceedance probability* associated with the *i*th largest observation. If the original observations were independent, in repeated sampling the U_i have *beta distribution* with mean

$$E[U_i] = i/(n+1) \qquad (2.1.3.2)$$

and variance

$$Var(Ui) = \frac{i(n-i+1)}{(n+1)^2(n+2)} \quad (2.1.3.3)$$

Knowing the distribution of the exceedance probability $U_{i,}$ one can develop estimators *qi* of their values which can be used to plot each $X_{(i)}$ against a probability scale.

Let $G(x)$ be a proposed cdf for the events. A visual comparison of the data and a fitted distribution is provided by a plot of the *i*th largest observed event $X_{(i)}$ versus an estimate of what its true value should be. If $G(x)$ is the distribution of X, the value of $X_{(i)} = G^{-1}(1- U_i)$ should be nearly $G^{-1}(1-qi)$, where the *probability plotting position qi* is our estimate of $U_{i.}$ Thus the points $[G^{-1}(1-qi), X_{(i)}]$ when plotted would, apart from sampling fluctuation, lie on a straight line through the origin.

The exceedance probability of the *i*th largest event is often estimated using the *Weibull plotting position*:

$$qi = \frac{i}{n+1} \quad (2.1.3.4)$$

corresponding to the mean of U_i.

3.3.1 Choice of Plotting Position

Hazen originally developed probability paper and imagined the probability scale divided into *n* equal intervals with midpoints $qi = (i\text{-}0,5)/n$, $i=1, \ldots,$ n; these served as his plotting positions. Gumbel rejected this formula in part because it assigned a return period of 2*n* years to the largest observation, Gumbel promoted Eq. (2.1.3.4).

Cunanne argued that plotting positions *qi* should be assigned so that on average $X_{(i)}$ would equal $G^{-1}(1-qi)$, that is *qi* would capture the mean of $X_{(i)}$ so that

$$\mathrm{E}[X_{(i)}] \approx G^{-1}(1\text{-}qi)$$

Such plotting position would be almost quantile-unbiased. The Weibull plotting position $i/(n+1)$ equal the average exceedance probability of the ranked observation $X_{(i)}$, and hence are probability-unbiased plotting positions. The two criteria are different because of the nonlinear relationship between $X_{(i)}$ and U_i.

Different plotting positions attempt to achieve almost quantile-unbiasedness for different distributions; many can be written

$$qi = \frac{i-a}{n+1-2a}$$

which is symmetric so that $q_i = 1\text{-}q_{n+1-i}$. Cunanne recommended a=0,40 for obtaining nearly quantile-unbiased plotting positions for a range of distributions.

Other alternatives are Blom's plottin position (a=3/8) which gives nearly unbiased quantiles for the normal distribution, and the Gringorten position (a=0,44) which yields optimized plotting positions for the largest observations from a Gumbel distribution. These are summarized in Table 2.1.-2, which also reports the return period $T_1 = 1/q_1$, assigned to the largest observation.

The difference between the Hazen formula, Cunanne's recommendation and the Weibull formula is modest for i of 1 of 3 or more. However, differences can be appreciable for i=1, corresponding to the largest observation. It is important to remember that the actual exceedance probability associated with the largest observation is a random variable with mean $1/(n+1)$ and a standard deviation of nearly $1/(n+1)$, see Eq. (2.1.3.2) and (2.1.3.3). Thus all plotting positions give crude estimates of the unknown exceedance probabilities associated with the largest (and smallest) event.

Table 3.3-1 Alternative Plotting Positions

Name	*Formula*	*a*	T_1	*Motivation*
Weibull	$\frac{i}{n+1}$	0	n+1	Unbiased exceedance probability for all distributions
Median	$\frac{i-0.3175}{n+0.365}$	0.3175	1.47n+0.5	Median exceedance probability for all distributions
APL	$\frac{i-0.35}{n}$	0.35	1.54n	Used with PWMs
Blom	$\frac{i-3/8}{n+1/4}$	0.375	1.60n+0.4	Unbiased normal quantiles
Cunnane	$\frac{i-0.40}{n+0.2}$	0.40	1.67n+0.3	Approximately quantile-unbiased
Gringorten	$\frac{i-0.44}{n+0.12}$	0.44	1.79n+0.2	Optimized for Gumbel distribution
Hazen	$\frac{i-0.5}{n}$	0.50	2n	A traditional choice

4 PROBABLE MAXIMUM PRECIPITATION

For design of high-hazard structures as spillways on large dams it is necessary to use precipitation values with very low risk of exceedance. Ideally a hydrologist would like to choose design storms for which there is no risk of exceedance. A theoretical problem that has plagued the search for such a storm is determining whether there is indeed an upper limit of rainfall amount.

The existence of an upper limit on rainfall is both mathematically and physically realistic (*Gilman,1964*). The spatial and temporal context of the upper bond on rainfall amount is incorporated into the definition of probable maximum precipitation (PMP), which is defined as "theoretically the greatest depth of precipitation for a given duration that is physically possible over a given size storm area at a particular geographical location at a certain time of year." A more troublesome problem than ascertaining whether an upper bond exists is determining what it is.

Observed rainfall totals worldwide provide a broad indication of maximum possible rainfall totals at a point as function of duration.

The estimation of probable maximum precipitation was developed and has evolved in the United States as a hydrometeorological procedure. Three meteorological components determine maximum probable precipitation:

1. amount of precipitable water
2. rate of convergence
3. vertical motion

Meteorological models representing all three components have been developed, but it has proved quite difficult to specify maximum rates of areal convergence and vertical motions for the models. In the standard approach to estimating PMP observed storm totals for extreme storms are used as indicators of the maximum values of convergence and vertical motions. The two major components of the

standard approach to the PMP computation are *moisture maximization* and *storm transposition*. In the first step of moisture maximization, the goal is to increase storm rainfall amounts to reflect the maximum possible moisture availability. In the storm transposition step, it is determined whether a given storm, which occurred in a broad region around the basin of interest, can be transposed to represent rainfall over the basin.

The principal data required for standard PMP computations are:

1. a catalogue for extreme storms
2. surface dew-point temperature observations

The storm catalogue for Austria and Italy contains storm date, location and depth. Additional meteorological information, including surface and upper-air maps, is typically used subjectively in determining storm transposition regions and for other quality-control purpose. Surface dew-point is used as a moisture index for moisture maximization . Precipitable water can be computed from surface dew-point values under the assumption that saturation levels extend to the ground.

Storm transposition is based on the assumption that for a given storm meteorologically homogeneous regions exist over which the storm is equally likely to occur. The transposition procedure involves meteorological analysis of the strom to be transposed, determination of transposition limits, and application of adjustments for changes in storm location. Meteorological analysis provides a characterization of key aspects of storm type. Transposition limits are determined from a long series of daily weather charts by identifying the boundary of the region over which meteorologically similar storm type have occurred. Adjustments account for differences in moisture maxima for the storm location and transposition sites. Adjustments are sometimes also made for topographic effects, although objective procedures for determining orographic adjustments are not widely accepted.

Having obtained a series of storms, PMP is determined by *envelopment*. Envelopment entails selection of the storm which has the largest maximised storm rainfall for a given time interval.

The envelopment process is used because a single historical storm is generally not the critical event the entire range of time scales required.

5 PROBABILISTIC ENVELOPE CURVES FOR EXTREME FLOODS

This chapter wants to introduce the method developed by *Castellarin et.al, 2005* for the estimation of extreme floods by envelope curves and for the assumption of their connected probability. This method has been rearranged in order to make it applicable to extreme rainfall events, both to annual maximum series and peaks-over-threshold series . The following methodology will be then applied not to floods as explained, but to extreme rainfall events in different ways for Austrian and Italian catchments, as discussed in Chapter 7.

Next paragraphs propose the original probabilistic interpretation of envelope curves and the formulation of an empirical estimator of the recurrence interval *T* associated with a REC.

5.1 DEFINITION OF REGIONAL ENVELOPE CURVES

A regional envelope curve (REC) summarizes the current bound on our experience of extreme floods in a region. Although RECs are available for many regions of the world, their traditional deterministic interpretation limits the use of the curves for design purposes, as magnitude, but not frequency of extreme flood events, can be quantified. A probabilistic interpretation of a REC is introduced via an estimate of its exceedance probability.

The bound on our experience of extreme floods gained up to the present through systematic observation of flood discharges in a region is defined in terms of the largest floods observed at all gauging stations in a region. Herein, the largest flood is termed the record flood, and gauging stations are referred to as sites. An example of a REC is illustrated in Figure 5-1 which plots, for each site, the normalized record flood, defined as the logarithm of the ratio of the record flood to its basin area, versus

the logarithm of the basin area. The REC is the line drawn on Figure 5-1 which provides an upper bound on all the normalized record floods at present.

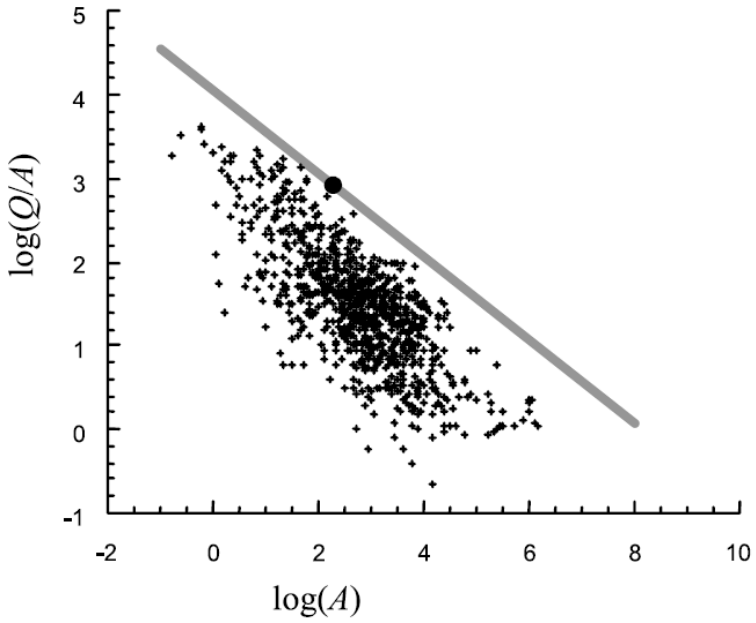


Figure 5-3.3.1-1 Q in feet^3/s (1foot=0.3048 m) and A in miles^2 (1 mile=1.609km) [Jarvis, 1925]: elements of experience (pluses) and element of experience (circle) defining the intercept of the envelope curve (shaded line).

The idea of bounding our flood experience dates back to Jarvis [1925], who presented a REC based on record floods at 888 sites in the conterminous United States. Roughly 50 years later, Crippen and Bue [1977] and Crippen [1982] updated the study by Jarvis [1925] by creating 17 different RECs, each for a different hydrologic region within the United States, based on a total of 883 sites. Matalas [1997] and Vogel et al. [2001] document that the

RECs identified by Crippen and Bue [1977] and Crippen [1982] still bound our flood experience gained from 1977– 1994 at 740 of the 883 sites compiled by Crippen and Bue. Enzel et al. [1993] examine the REC bounding the historical flood experience for the Colorado river basin and show that the same REC also bounds the paleoflood discharge estimates available for the basin.

The development of RECs is not confined to the United States; they have been developed for Italy [*Marchetti, 1955], western Greece [Mimikou, 1984*], Japan

[*Kadoya, 1992*], and elsewhere. RECs have been used to compare record flood experience in the United States, China, and the world by Costa [1987] and, more recently, by Herschy [2002].

The REC provides an effective summary of our regional flood experience. The pioneering work of Hazen [1914], who formalized flood frequency analysis, a formalism still in use, and who was among the first to suggest a method for improving information at a site through the transfer of information from other sites (i.e., substitution of space for time), has tempered the use of a flood magnitude as a design flood without an accompanying probability statement. Our objective is to provide a probabilistic interpretation of the REC. In the almost 80 years since Jarvis [1925] introduced the envelope curve, a probabilistic interpretation of a REC has never been seriously addressed.

RECs have continued to be constructed and viewed mainly as summary accounts of record floods, rather than as meaningful tools for the design of measures to protect against ‘‘catastrophic’’ floods. It has been suggested that there is no obvious way to assign a probabilistic statement to a REC [*see, e.g., Crippen and Bue, 1977; Crippen, 1982; Vogel et al., 2001*]. Water Science and Technology Board, Commission on Geosciences, Environment and Resources [1999] argued that the determination of the exceedance probability of a REC is difficult due to the impact of intersite correlation. As a consequence, RECs are assumed to have little utility beyond the suggestion of the U.S. Interagency Advisory Committee on Water Data [1986] that they are useful for ‘‘displaying and summarizing data on the actual occurrence of extreme floods.’’ A probabilistic interpretation of the REC offers opportunities for several engineering applications which seek to exploit regional flood information to augment the effective record length associated with design flood estimates.

A potential advantage of assigning a probabilistic statement to a REC is that this approach avoids the need to extrapolate an assumed at-site flood frequency distribution hen estimating a design event.

This work would provide a probabilistic interpretation of the REC, to approximate its exceedance probability, and to quantify the effect of intersite correlation on estimates of the exceedance probability and extend their utilisation from floods to rainfall events.

5.2 A PROBABILISTIC INTERPRETATION OF ENVELOPE CURVES

It is common practice to construct a REC, as in Figure 5-1, which plots the logarithm of the ratio of the record flood to the drainage area, ln(Q/A) versus ln(A). Jarvis [1925] suggested modeling the REC for the United States using,

$$\ln\frac{Q}{A} = a + b\ln(A) \qquad (5.1\text{-}1)$$

with a = 9.37 (or a = 4.07 if log is used in (5.1-1) instead of ln) and b = -0.50, where Q and A are in cubic feet per second and square miles, respectively. Together with Jarvis [1925], other empirical studies showed that b is negative and greater than 2/3 for various portions of the world [*Linsley et al., 1949; Marchetti, 1955; Crippen and Bue, 1977; Matalas, 1997; Herschy, 2002*].

Assuming a fixed value of b, the intercept a in (5.1.1) may be estimated by forcing the REC to bound all record floods to the present, say up to the year n. Let $X_{j\,i}$ denote the annual maximum flood in year i = 1, 2, . . .,n at site j = 1, 2, . . .M, where M is the number of sites in the region.

Let $X_{j\,(i)}$ denote the flood flow of rank (i) at site j, where ranking is from smallest (1) to largest (n). The REC's intercept up to the year n can then be expressed as

$$a^{(n)} = \max_{j=1,\ldots,M}\left\{\ln\left(\frac{X_j^{(n)}}{A_j}\right) - b\ln(A_j)\right\}$$

where A_j is the area of site j = 1, 2, . . .M.

5.3 DESCRIPTION OF THE METHOD

The probabilistic regional envelope curves (PRECs) developed by *Castellarin et al, 2005* presented a probabilistic interpretation of the envelope curve presented in equation (5.1-1). The proposed probabilistic regional envelope curves (PRECs) are based upon two fundamental assumptions:

1. the group of river basins (i.e., geographical region or pooling group of sites is homogeneous in the sense of the index flood hypothesis
2. the relationship between the index flood μx (e.g. mean annual flood) and A is of the form,

$$\mu x = CA^{b+1}$$

where b and C are constants and b is the same as in equation (5.1-1).

Under these assumptions has been developed an empirical estimator of the exceedance probability p of the expected PREC. The expected PREC is the envelope curve that, on average, is expected to bound the flooding experience for a region of given characteristics (i.e., M sites with record length n and mean intersite correlation $\bar{\rho}$). The expected PREC is identified through a series of Monte Carlo simulations experiments by generating a number of synthetic cross-correlated regions consisting of M concurrent sequences of annual maximum flood with record length n. It is proposed an estimator of p of the expected PREC that is a function of the effective number of sites, $M_{EC} \leq M$ (i.e., number of independent flood series with an equivalent information content).

The recurrence interval of the expected PREC can then be computed as the inverse of p. Assigning an exceedance probability, or equivalently a recurrence interval, to a PREC makes it a practical tool for estimating a design event at gauged and ungauged river basins.

We present an algorithm for the application of the empirical estimator of p to historical annual maximum series of unequal length.

5.3.1 Estimating the Effective Number of Sites

The problem of estimating the exceedance probability p of the expected PREC reduces to estimating the exceedance probability of the largest value in a regional sample of standardized annual maximum peak flows.

The proposed estimator of *p* considers the spatial correlation between the regional data. The estimator evaluates the equivalent number of independent sequences M_{EC} for *M* concurrent and cross-correlated series of annual maxima with length *n*.

The estimators reads,

$$\hat{M}_{EC} = \frac{M}{1+\overline{\rho^{\beta}}(M-1)} \qquad \text{with} \qquad \beta = 1.4\frac{(nM)^{0.176}}{(1-\rho)^{0.376}} \qquad (5.3\text{-}1)$$

where $\overline{\rho^{\beta}}$ and $\overline{(1-\rho)^{0.376}}$ are average values of the corresponding functions of the correlation coefficients (i.e. $\overline{\rho^{\beta}}$ is the average of the M(M-1)/2 values of $\rho^{\beta}_{k,j}$, where $\rho_{k,j}$ is the correlation coefficient between sites k and j, with $1 \leq k < j \leq M$).

Equation (5.3-1) was identified by referring to M_{EC} values obtained through Monte Carlo simulations, *Var* $[\rho_M]$, is a function of *M*. The empirical values of *Var* $[\rho_M]$ are computed for *M* concurrent and cross-correlated synthetic series with length n generated from the multivariate normal distribution with zero mean, unit variance and average correlation among the series $\overline{\rho}$. Then are computed M_{EC} values relative to a number of combinations of *M*, *n* and $\overline{\rho}$ values by finding the zeros of the theoretical expression of *Var* $[\rho_M]$ for the corresponding empirical values obtained trough Monte Carlo simulation. Finally equation (5.3-1) is referred to these zeros.

The *p* value of the expected PREC can be computed by estimating M_{EC} trough equation (5.3-1) and using a suitable plotting position by setting the overall sample years of data to the equivalent number of independent annual maxima $n \cdot M_{EC}$.

5.3.2 Evaluation of the Intersite Correlation

The application of equation (5.3-1) requires the representation of the distribution of theoretical correlation coefficients for the study region through a suitable cross-correlated formula.

A possible approach to modelling cross-correlation between rainfall events is to compute the sample correlation coefficients using sample estimators proposed in the scientific literature (*Stedinger, 1981*) and, in absence of more specific information

about the spatial structure of the intersite correlation model, to approximate the true correlation coefficient $\rho_{k,j}$ through empirical correlation functions of the distance $d_{k,j}$ among sites *k* and *j* (see, *Tasker and Stedinger, 1989: Hosking and Wallis, 1988; Troutman and Karlinger, 2003*).

A simple correlation function that can been used is as follow:

$$\rho k, j = \exp\left(-\frac{\lambda_1 d_{k,j}}{1+\lambda_2 d_{i,j}}\right) \qquad (5.3.2\text{-}1)$$

where $\lambda_1 > 0$ and $\lambda_2 \geq 0$ are the formula regional parameters of the formula.

5.3.3 Estimating the Effective Sample Years of Data

The condition of equal length for all series is seldom encountered in real-world data sets, hence equation (5.3-1) is generally inapplicable.

In order to supply this problem let us suppose that the actual regional data sets consist of *M* AMS and globally spans *n* years. First, one identifies the number of years n_1, for which the original data set includes only one observation of the annual maxima discharge, that is the years in which *M-1* observations are missing. These n_1 observations are effective by definition. Then, the data set containing the $n\text{-}n_1$ remaining years is subdivided into $N_{sub} \leq (n\text{-}n_1)$ subsets; each one of them (say subset *s*) is selected in such a way that all its $Ls \leq M$ sequences are concurrent and of equal length *ls* and therefore suitable for the application of the estimator in equation (5.3-1). The overall effective sample of years of data, n_{eff}, coincides with the sum of n_1 and the effective sample of years of data of all *Ns* subsets.

Therefore $\hat{n}_{eff}$ can be calculated as

$$\hat{n}_{eff} = n_1 + \sum_{s=1}^{Nsub} \hat{n}_{eff,s} = n_n + \sum \frac{Lsls}{1+[\overline{\rho^{\beta}}](Ls-1)_{Ls}} \qquad (5.3.3\text{-}1)$$

with $\beta = 1.4\frac{(Lsls)^{0.176}}{\left[(1-\rho)^{0.376}\right]_{Ls}}$

where the notation $[\cdot]_{Ls}$ indicates that the average terms $\overline{\rho^{\beta}}$ and $(1-\rho)^{0.376}$, which have the same meaning as in equation (5.3-1), are to be computed with respect to the *Ls* > 1 annual flood sequences which form subset *s*. Also β exponent in equation (5.3.3.-1) coincides formally with β in equation (5.3-1).

This is consistent with the fact that the *Ls* sequences forming each subset *s* are concurrent and of equal length *ls*, which was the condition adopted for the identification of the empirical relationship in equation (5.3-1).

6 STUDY REGIONS AND LOCAL REGIME OF RAINFALL EXTREMES

6.1 CLIMATE AND MORPHOLOGY OF AUSTRIA

The study region is Austria which is hydrologically quite diverse, ranging from lowlands in the east to high alpine catchments in the west. Elevations range from less than 200 m above sea level (asl) to more than 3000m asl. Mean annual precipitation is less than 400 mm/year in the east and almost 3000 mm/year in the west. Land use is mainly agricultural in the lowlands, forested in the medium elevation ranges, while alpine vegetation and rocks prevail in the highest catchments.

The analysis of timing of floods (*Merz and Blöschl, 2003*) outlines that in the high alpine catchments in the west of the country floods tend to occur in summer showing strong seasonality, which suggest the presence of snow and glacier melt floods. In the north, weak seasonality suggests that there is no single dominant process for a given catchment. The spring floods in the very north of the country suggest that both early snowmelt and rain-on-snow are likely to occur in these catchments. In Carinthia, in the very south of Austria, floods tends to occur in late autumn which may be related to the advection of most air from the Mediterranean south of Austria.

Considering the storm duration it is outlined that short-durations of storms causing floods mainly occur in the south-eastern Austria where short convective storms are known to be important for flood generation. In contrast, long duration storms occur in the north-western Austria at the northern fringe of the Alps. As most flood producing weather system approach from the north-west, this zone is an area of orographic enhancement effects where synoptic rainfalls tend to be most important for flood generation.

From the point of view of rainfall depth and snowmelt, rainfall depths are usually largest in the north west of the country, mainly as result of the orographic effects mentioned above.

Figures 6.1-1a and 6.1-1b show the spatial patterns of the frequency of long-rain floods, short-rain floods, flash-floods and snowmelt floods in Austria. This frequency is the number of years a maximum annual flood is classified as a certain process type, scaled by the total number of years for each catchment. Figure 6.1-1 indicates that long rain floods are the main causative process type of annual maximum floods in most catchments in Austria, as the frequencies are on the order of 0.5 (see Table 6.1-1 below).

Process Type	*Long-rain floods*	*Short-rain floods*	*Flash-floods*	*Rain-on-snow floods*	*Snowmelt floods*	*All types*
Number of events	783	597	302	430	154	2266
Number of peaks < MAF	2511 (56.6%)	1281 (39.7%)	274 (50.3%)	1398 (57.4%)	248 (71.5%)	5712 (49.6%)
Number of peaks > MAF and < 10 yr flood	2051 (41.3%)	1541 (47.8%)	225 (41.3%)	957 (39.3%)	94 (27.1%)	4868 (42.3%)
Number of peaks > 10 yr flood	404 (8.1%)	403 (12.5%)	46 (8.4%)	80 (3.3%)	5 (1.4%)	938 (8.1%)
Total number of flood peaks	4966 (100%)	3225 (100%)	545 (100%)	2435 (100%)	347 (100%)	11518 (100%)

Table 6.1-1 Flood type classification of annual maximum flood in Austria, MAF is the mean annual precipitation

In catchments at the northern fringe of the high Alps, long rain floods are particularly common. The high Alps tend to act a as a topographic barrier to north-westerly airflows, and orographic enhancement often produces persistent rainfall which can result in floods. The regions of the highest relative importance of long-rain floods are identical with the regions of the highest mean annual rainfall in Austria. Short-rain floods, see Figure 6.1-1b, also occur quite commonly with a frequency of the order of 0.3.

There are, again, significant spatial differences. Short-rain floods occur more frequently in southern Austria than north of the Alps. This is likely due to two mechanism. The main ridge of the Alps tends to block weather systems approaching from the northwest which reduces the advection of moist air and hence the persistence of the rainfall. Also, in the south of the Alps southern airflows may produce floods that are associated with high-intensity short-duration storms. There is likely the tendency for a quicker response of some of the catchments in the south of the main Alpine ridges as compared to the north of it, which tends to enhance the role of short duration storms in flood generation.

Flash floods occur significantly, see Figure 6.1-1c, than long-rain and short-rain floods. Flash floods are only important in eastern Austria, specifically in the hilly region of Styria in south-eastern Austria and in the hilly region of Waldviertel in north-eastern Austria. The hilly terrain appears to increase the instability of the boundary layer and hence the likelihood of convective storm. Throughout Austria, the spatial pattern in Figure 6.1-1c is rather patchy which reflects the random and local nature of flash floods causing maximum annual floods.

Rain-on-snow floods Figure 6.1-1d rarely produce the maximum annual flood. Those catchments with frequency of snowmelt floods of more than 0.1 are mainly located in the high Alps where both snow and glacier melt can be important for flooding, and in the northern Austria where early spring snowmelt may produce floods. However, these are usually minor floods.

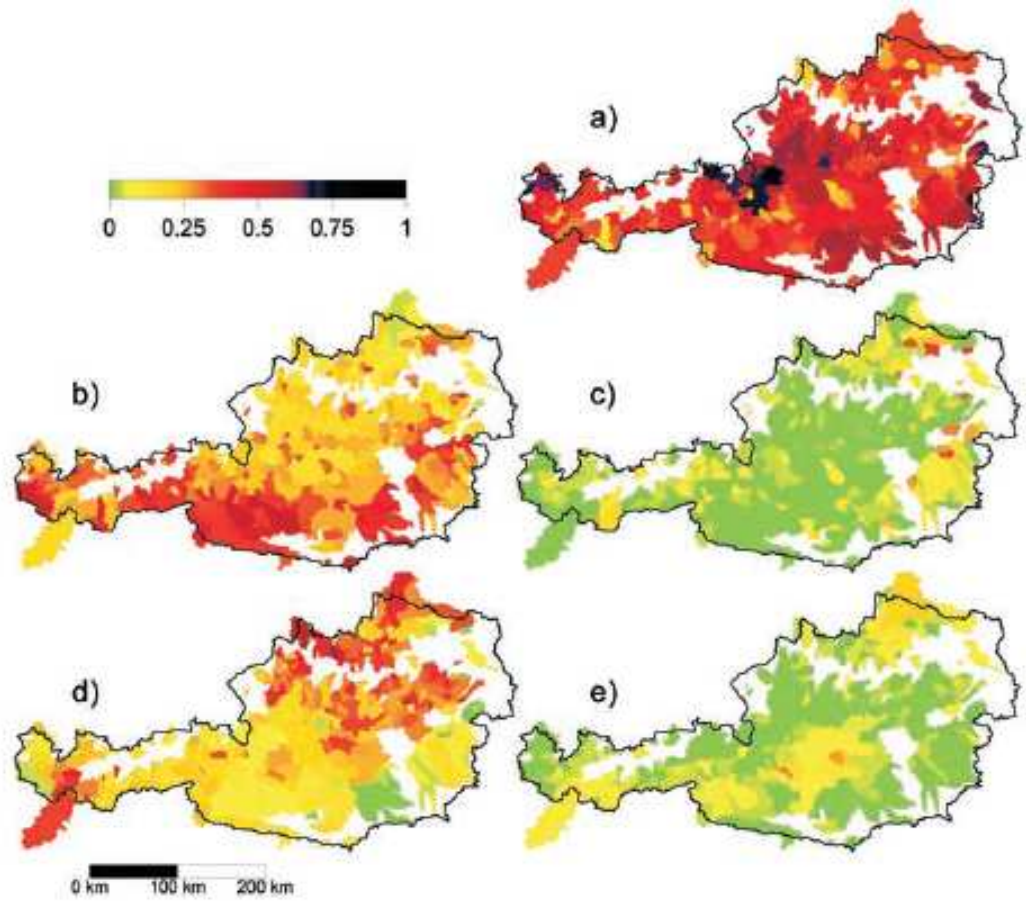


Figure 5.3.3-1 Regional patterns of the frequency of flood process types. A frequency of unity indicates that in a catchment all the maximum annual floods are due to one particular process while a frequency of zero indicates that this process never leads to a maximum annual flood. (a) Long-rain floods, (b) short-rain floods, (c) flash floods, (d) rain-on-snow floods, and (e) snowmelt floods. For nested catchments the frequencies of the smaller catchments have been plotted on top of those of the larger catchments, and only catchments smaller than 5000 km2 are shown.

Considering *seasonality* it is possible to assume that long-rain floods occur throughout the year but there is a tendency for more events and more extreme events to occur in summer, particularly in June and July. This is because heavy rainfall events occur more frequently in the summer month than in the rest of the year.

Short-rain floods also mainly occur in summer and there is a tendency for soe of the major events to also occur in autumn. These are events that have occurred in southern Austria, see Figure 6.1-2.

Flash-floods only occur in summer when enough energy is available for convective storms.

Rain-on-snow floods occur throughout the year with the exception of late summer and early autumn. The largest Rain-on-snow floods occur in late December.

Similarly snowmelt floods occur throughout the year with the exception of late summer and autumn when all the catchments are snow free.

The seasonal pattern of the flood occurrence is mainly related to the annual fluctuation in air temperature. Air temperature also exhibits a strong altitudinal dependence. There is a very pronounced seasonal patterns for all floods types. Long-rain floods and short-rain floods exhibits an upper envelope of about 1000 m als from January to April. In late spring and summer these floods can occur in catchments of any elevation. At the end of the year the upper envelope gradually decreases to 100 m asl. This pattern of the upper envelope closely follows the snow fall line. In the high altitude catchments no annual maximum flood occur in the cold months of the year as most of the precipitation falls as snow.

Flash floods exhibitts a strong seasonal pattern following the annual pattern of global radiation. It is interesting that, in July and August, flash floods can occur in catchments with elevation of up to 3000 m als. These are likely a result of convective storms that can occur at any altitude during summer.

Rain-on-snow floods and snowmelt floods exhibits a narrow altitudinal range of occurrence which varies with time of the year. In winter both types occur in catchments lower than 1000 m als.

This pattern is clearly related to the seasonal pattern of air temperature. Both Rain-on-snow floods and snowmelt floods appear to occur only within a limited range of temperature conditions for which a snow cover exists, but snowmelt and/or rain may occur.

6.2 AVAILABLE RAINFALL DATA

Hydrologic processes such as precipitations evolve on a continuous time scale. However, most hydrologic processes of practical interest are defined in a discrete time scale. The considered data derive from a discrete time series by sampling the continuous process of rainfall at discrete points in time. Hydrologic time series may

also be classified into categories depending on a number of factors, in particular this work will give importance to Uncorrelated and Correlated time series, the discussed is reported in Paragraphs 7.1.1 and 7.1.2.

The analysed data are the rainfall daily data with duration 30 minutes, 1 hour, 3 hours, 9 hours and 24 hours (1948-1997) recorded in 700 hydrological stations in Austria and obtained by applying a filtered peak over threshold procedure to the daily time series.

Actually two general approaches are available for modelling flood, rainfall and many other hydrological series. Using an annual maximum series, one considers the largest event in each year; using a partial duration (PDS) series or peaks-over-threshold (POT) approach, the analysis includes all peaks above a truncation or threshold level.

6.2.1 Peaks-Over-Threshold Procedure

Partial duration series considers all independent peaks which exceed a specific threshold: a threshold is applied to the sequence to retain only the large peaks.

Fortunately one can estimate annual exceedance probability from the analysis of POT by empirical relationships or equations. The basic idea is to extract from the daily discharge sequences a sample of peaks containing more than one flood peak per year, in order to increase the available information with respect to the annual maxima analysis. However, in the practical applications ambiguous criteria for peak selection affect the efficiency of the method: for example, the average number of peaks per year, λ, is usually forced to remain in the range between 2 and 3 in order to preserve independence among subsequent flood peaks (*e.g., Lang et al., 1999; Madsen et al., 1997*), despite the fact that such low values contrast with physical and statistical considerations.

However usually POT series are relatively long and reliable POT records are often available and, if the arrival rate for peaks over threshold is large enough, POT analysis should yield very accurate estimates of extreme quantiles. Still, a drawback

of POT analysis is that one must have criteria to identify only independent peaks (and not multiple peaks of the same event).

The appropriate threshold is given for this study by the *Wussow Criteria*. There are several deterministic criteria for heavy rain in common use, Wussow (1922) gives a simple analytical expression for identifying the critical rainfall depth *Ncr* so that the inequality rainfall depth $N \geq Ncr$ can be used as a criterion to assess heavy point rainfall. *Ncr* is function of several physical parameters; in particularly assuming geometrically similar catchments, the critical rainfall depth *Ncr* during the time of duration *D* appears to depend on the acceleration of gravity *g*, a coefficient *K* of turbulent diffusion of water in air and the precipitation intensity *Rcr*

The estimate of *Ncr* can be determined as

$$Ncr = CD^{\frac{1}{2}}$$

where *C* is the Wussow's empirical value: $C = \sqrt{5}mm \cdot \min^{-\frac{1}{2}}$

The adopted criterion fixes the threshold including all peaks which rainfall depth is larger than the square root of 5 times the corresponding duration.

$$N(mm) \geq \sqrt{5D(\min)}$$

6.2.2 Identification of Regions

We have identified circular subregions centred on points located as to create a grid, 6000 gridpoints derive by the subdivision of the region in 100 parts on x (dx=5.6 km) and in 60 parts on y (dy=4.8 km). This grid covers the entire analysed area and will be fundamental in following studies. The available data are taken by 700 rainfall stations distributed on Austria and span by the duration of 30 minutes to 24 hours, considering durations as belonging to a range of time as shown in Table 2.2.1-1 below.

Table 6.2.1-6.2-1 Considered duration classes for the rainfall data

Duration Class *(min)*	From *(min)*	To *(min)*
30	1	44
60	45	89
180	90	358
360	90	899
540	365	899
1440	900	4310

The rainfall stations position and the 6000 gridpoints analysed are showed in *Figure 6.2.2-1*.

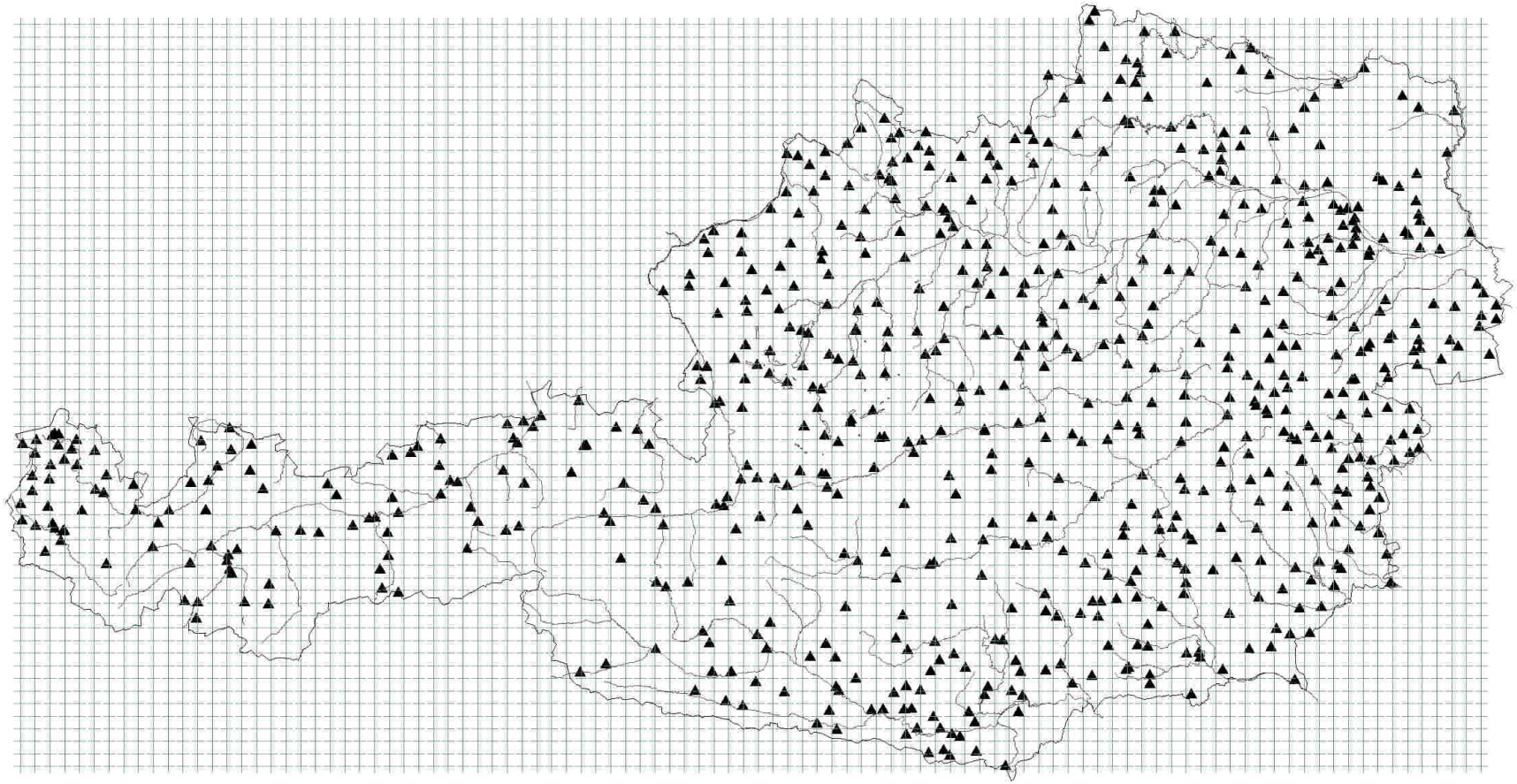


Figure 6.2.2-1– Location of the 693 stations in Austria. On the entire area a grid has been created, 6000 gridpoints are so defined with spacing of 5,6 km on X coordinates and 0,5 km on Y**.**

The first representation of the available data is given by looking for the largest rainfall event in given distance by each gridpoint. 50 km has been selected as the reference distance in which to consider stations, so to create circular regions with radius 50 km and consequent approximate area of 8000km^2. Only the rainfall stations belonging to the region are considered, and the first analysis consist of selecting the stations for each region and then considering in it the maximum rainfall data.

This procedure is extended to each region and the maximum selected data are plotted in order to generate, by using the spatial interpolating *Technique of Kriging*, a *Surfer* colour scaled map.

Kriging is a group of geostatistical techniques to interpolate the value $Z(x_0)$ of a random field $Z(x)$ (e.g. the rainfall depth as a function of the geographic location x) at an unobserved location x_0 from observations $z_i = Z(x_i), i = 1,...,n$ of the random field at nearby locations $x_1,...,x_n$. Kriging computes the best linear unbiased estimator $\hat{Z}(x_0)$ of $Z(x_0)$ based on a stochastic model of the spatial dependence quantified either by the variogram $\gamma(x,y)$ or by expectation $\mu(x) = E[Z(x)]$ and the covariance function $c(x,y)$ of the random field.

The kriging estimator is given by a linear combinationf the observed values $z_i = Z(x_i)$ with weights $w_i(x_0), i = 1,...,n$ chosen such that the variance (also called *kriging variance* or *kriging error*): is minimized subject to the unbiasedness condition:

$$E\left[\hat{Z}(x) - Z(x)\right] = \sum_{i=1}^{n} w_i(x_0)\mu(x_i) - \mu(x_o) = 0$$

Figure 6.2.2-2 represents the used rainfall data for the duration class of 1 hour, collected by all stations and the maximum rainfall event recorded in distance 50 km from each gridpoint:

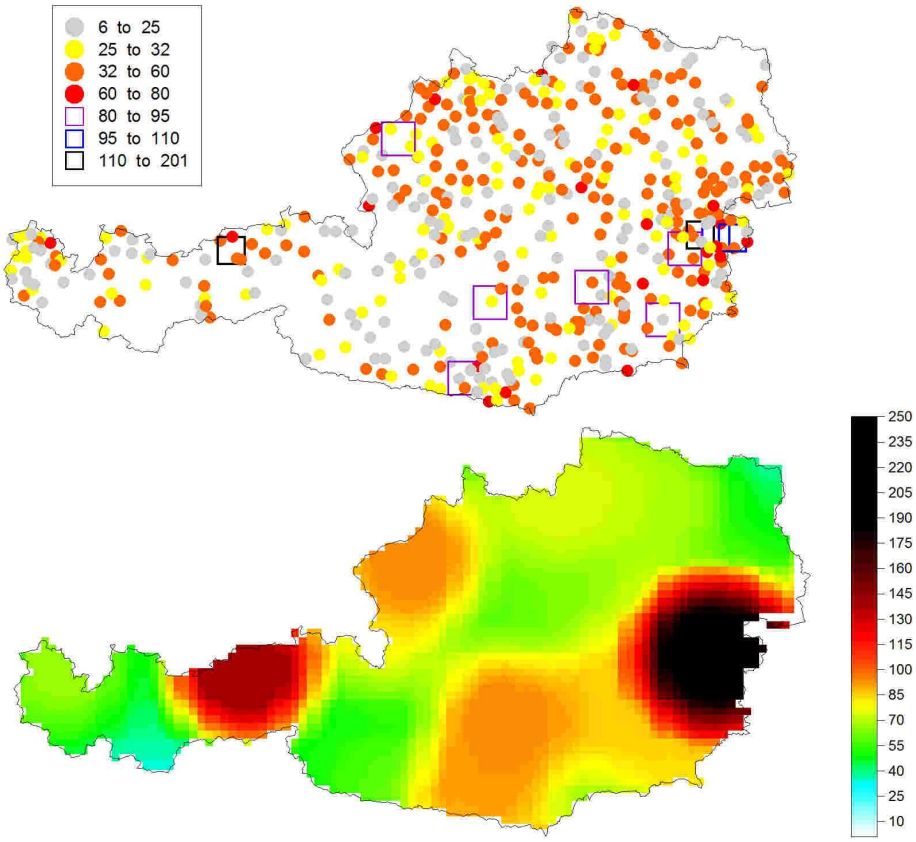


Figure 6.2.2-2 Representation of the available rainfall data (mm) for the duration class of 1 hour; top: pointy collected data by each rainfall station – bottom: maximum rainfall in distance 50 km by each gridpoint**.**

In Annexes(1) and (2) show the other durations, as the pointy data and as maximum rainfall in distance 50 km.

The length in years of the series of data vary significantly among stations, this tendency is showed in Annexes (3).

6.3 CLIMATE AND MORPHOLOGY OF ITALIAN REGION

The study area includes the administrative Districts of Emilia-Romagna and Marche, in northern central Italy, and occupies 37.200 km^2. The area is bounded by the Po River to the north, the Adriatic Sea to the east, and the divide of the Apennines to the southwest (see Figure 6.3-1). The north-eastern portion of the study

region is mainly flat, whereas the south-western and coastal parts are predominantly hilly and mountainous.

The mean annual precipitation (MAP) varies on the study region from about 500 to 2500 mm. Altitude is the factor that most affects the MAP, which exceeds 1500 mm starting from altitudes higher than 400 m above sea level (asl) and exhibits the highest values along the divide of the Apennines (see Figure 6.3-1).

The storm duration range spans from the short and localized convective events to the long and large areal coverage storms associated with large cyclonic weather systems. A regional analysis of the dates of occurrence of short-duration rainfall extremes (i.e., t equal to 1,…,3 hours) showed remarkable consistency and a mean timing which virtually coincided with the beginning of August for the entire study area; the dates of occurrence of long-duration rainfall extremes (i.e., t equal to 24 hours or 1 day) showed, for the same area, less consistency and a mean timing which varied from the beginning of September to the beginning of November [*Castellarin and Brath, 2002*].

Several regional frequency analyses of rainfall extremes were performed over the study area. These studies, according to the logic of the traditional ‘‘index flood’’ hypothesis [*Dalrymple, 1960*], proposed subdivisions of the region shown in Figure 6.4-1 into homogeneous climatic regions, within which the statistics of rainfall extremes for a given storm duration t are assumed to be constant [*Brath et al., 1998; Brath and Franchini, 1999; Brath and Castellarin, 2001*]. This assumption contrasts with the findings of other studies, which demonstrated for different geographical areas and climatic contexts that the statistics of rainfall extremes vary systematically with location [*Schaefer, 1990; Alila, 1999*]. These studies also identified statistically significant relationships between these statistics and the MAP, which was used as a surrogate of geographical location. Schaefer [1990] and Alila [1999] showed, for two rather large geographical areas, that the coefficients of variation and skewness of rainfall extremes tend to decrease as the local value of MAP increases, and they proposed to enhance the index flood hypothesis for the regional frequency analysis of rainfall extremes by dispensing with the delineation of geographical areas and defining, instead, as climatically homogeneous subregions those areas which have a small MAP range.

The applicability of the findings of Schaefer [1990] and Alila [1999] to the particular geographical and climatic context considered herein was investigated, and the available annual maximum series (AMS) of rainfall depth were analyzed, making use of a rain gage network with a higher resolution than the networks considered in the above mentioned studies. In particular, the variability of the sample L moment ratios [*Hosking, 1990*] of skewness (L-C_s) and variation (L-C_v) was examined against the variability of MAP. The results of the analysis show that the findings of Schaefer [1990] and Alila [1999] hold for a rather large portion of the study region, as the values of L-C_v and L-Cs of rainfall extremes are low in humid areas, while both statistics tend to increase, more or less markedly depending on the considered storm duration, as the local MAP value decreases. Nevertheless, the merger of two river basins indicated in Figure 6.3-1 as the Tyrrhenian Region, due to its closeness to the Tyrrhenian coast, exhibits a conflicting behaviour. The region is mountainous and therefore humid (average MAP ≈1400 mm), yet it shows high L-Cv and L-Cs values for all the storm durations considered in this study. This anomaly is illustrated in Figure 2 for the annual maxima of daily rainfall.

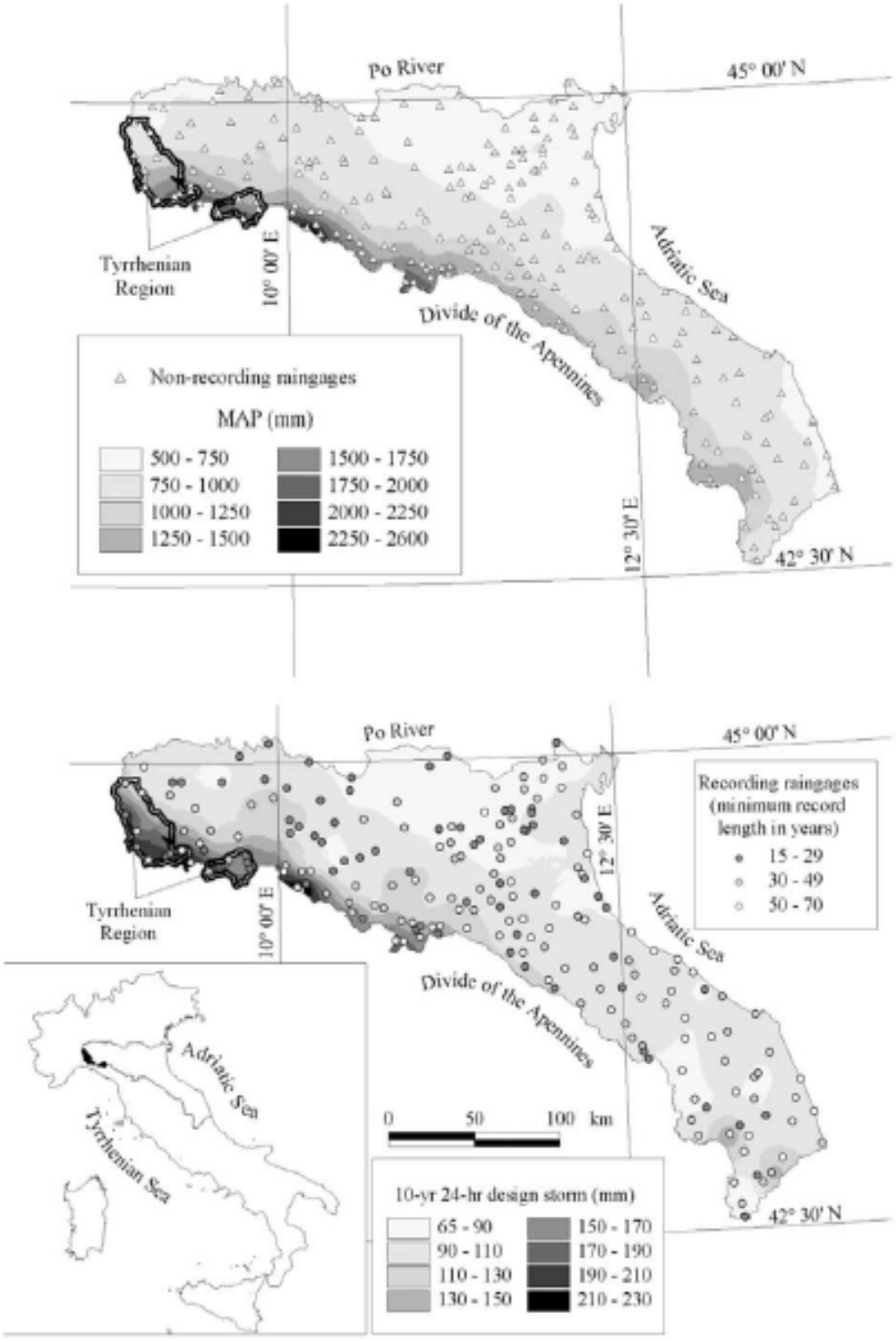
The figure shows

1. the sample L moment ratios against the MAP for 179 rain gages with at least 30 years of observations;
2. the weighted average MAP against the weighted average L-Cv and L-Cs for the Tyrrhenian Region;
3. the relationship between L-Cv and MAP identified by Alila for a storm duration of 24 hours;
4. the moving weighted average curves of L-Cv and L-Cs, based on a moving window including 15 data points outside the Tyrrhenian Region (shaded curve). The average MAP, L-Cv, or L-Cs values are obtained by weighing each measure proportionally to the recorded length of the corresponding rain gage [*see Hosking and Wallis, 1997*]. Figure 2 shows the anomalous behavior of the Tyrrhenian Region and the consistency with the results obtained by Schaefer [1990] and Alila [1999] for the remainder of the region.

The anomaly of the Tyrrhenian Region was already pointed out by Castellarin et al. [2001] and Castellarin and Brath [2002]. This atypical behavior may be accounted for by the proximity of the region to the Tyrrhenian shoreline and two windows in the Apennine divide, which locally drops below 1000 m asl, being normally above 1400–1500 m asl. These windows, known in literature as the Genoa gap, produce a fundamental topographic control, channelling the most severe disturbances coming from the south and originating over the Tyrrhenian Sea, and allowing them to have significant climatic control beyond the Apennine divide [*Tripoli et al., 2002*]. The fact that the northern coastal area of the Tyrrhenian Sea exhibits a maritime rainfall regime and rather high coefficients of variation and skewness of rainfall extremes [*Brath and Rosso, 1995*] may partly explain the high L-Cv and L-Cs values observed for the Tyrrhenian Region, despite its rather high MAP values.

Because of its particular behaviour, the Tyrrhenian Region has been excluded from this study .

Figure 6.2.2-1 Emilia-Romagna and Marche administrative regions; location of recording gages for different values of the minimum record length; Tyrrhenian Region; isoline maps of MAP and R10yr,24hr.



6.4 AVAILABLE RAINFALL DATA

This second paragraph focuses on Italy, it analyzes the annual series of precipitation maxima observed by a rather dense gaging network located in a wide geographical area of northern central Italy with duration spanning from 15 minutes to 24 hours.

This work adapts the probabilistic interpretation proposed for RECs to DDECs and it assesses the suitability of these curves for estimating the *T*-year rainfall event associated with a given duration and large *T* values.

The available extreme rainfall data consist of the annual series of precipitation maxima with duration t equal to 1, 3, 6, 12, 24 hours and 1 day (i.e., from 9:00 A.M. to 9:00 a.m. of the following day) that were obtained for a rather dense network of recording (hourly rainfall) and observational-day (daily rainfall) rain gages from the

National Hydrographic Service of Italy (SIMN). *Table 6.4-1* summarizes the available rainfall data, whereas Figure 6.4-1 shows the location of the recording gages, along with a subset of 226 nonrecording gages for which recent observation of MAP values are available (i.e., 1950–1991).

Table 6.4-1 Study Area: Number of raingauges and number of observations for different durations.

Duration (min)	*Number of Sites*	*Number of Observations*
1440	223	8052
720	223	8059
360	223	6667
180	223	7090
60	223	8047
45	189	941
30	222	3766
15	220	2187

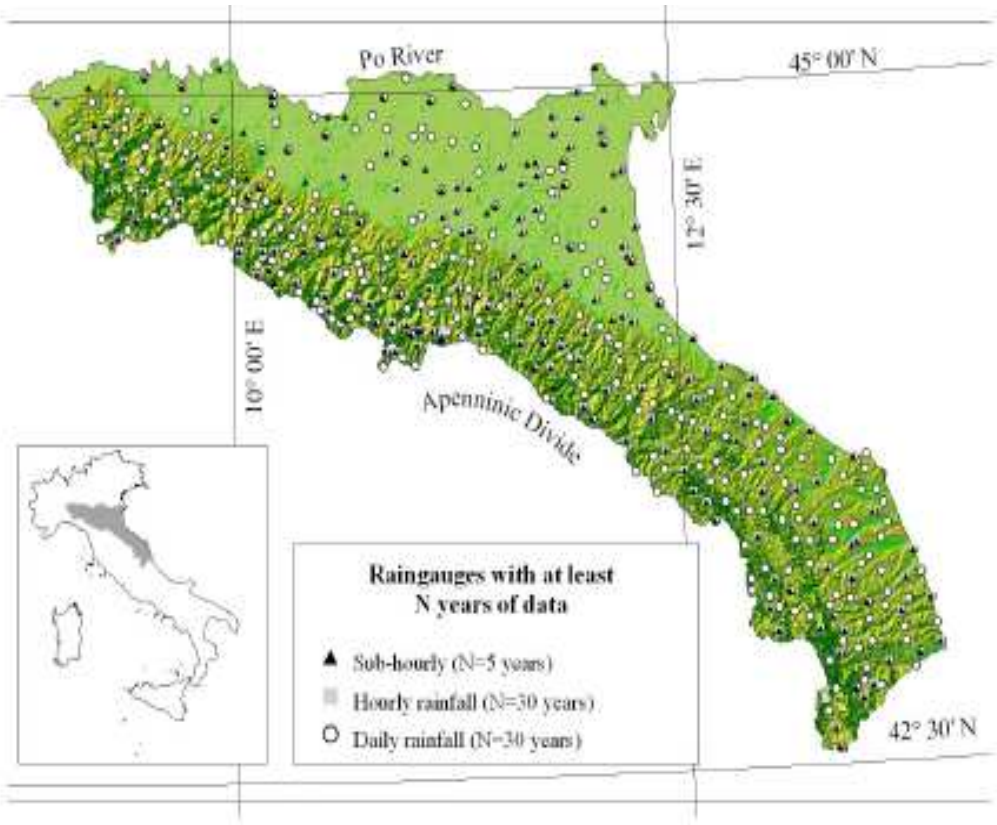


Figure 6.2.2-1 Study area and location of stations for different values of the minimum record length.

7 METHODOLOGY

7.1 AUSTRIAN CATCHMENTS

This first part focuses on Austria, in that Region has been analysed the peak over threshold (POT) series of rainfall depths, in order to find out the recurrence interval T and create a REC for the region Austria in which 6000 points are analysed so to define a grid covering the entire area.

The first analysis considers no spatial correlation between the rainfall data but a suitable plotting position is applied on the all sample years of data; the effect of intersite correlation on the exceedance probability is later discussed in order to make the probabilistic interpretation of envelope curves actually applicable to the real world data sets

A single time series x_t is correlated in time if the x at time t depend (linearly) on the x at time t-k, for k=1,2,…Time series of precipitation are usually uncorrelated, but several combinations of *cross-correlation* may exist. Let us consider the two series x_t and y_t, if the y at time t depend (linearly) on the x at time t-k, for k=02,1,… then the two time series are cross-correlated. It is possible that each series can be not correlated in time but cross-correlated between them. Examples are precipitation series at two nearby sites, one would expect that the time series will be cross-correlated because the sites are relatively close to each other and therefore subject to similar climatic and hydrologic events.

Next paragraphs deal with two different cases: the first one does not consider cross-correlation among series while the second deal with the searching for significant correlation among series and consequently revises the methodology of analysis.

7.1.1 Uncorrelated Case

If we consider no correlation between rainfall data collected by stations, the evaluation of the recurrence interval in years consist in the applying of a suitable plotting position. The chosen one is the Hazen plotting position.

If one considers the relation between the *T*-year event and the return period *T* as a representation of the frequency curve, one needs to attribute a return period to each measured peak value. This is done by assigning to the *j*-th peak in the ordered sample (*j* is then the rank of the peak) a return period [*Claps and Laio, 2003*]:

$$T_j = \left(\frac{tq}{\lambda q \cdot tq + 0.5 - j} \right)$$

where *tq* is the record length in years, λq the average number of flood peaks per year, and the 0.5 value comes from assuming the Hazen plotting position.

The rank of the peak, *j*, goes from 1, smallest event, to (λq *tq*)for the largest one.

This formula changes, considering a region, in:

$$T_j^R = \left(\frac{tq \cdot k}{\lambda q \cdot tq \cdot k + 0.5 - j} \right)$$

where *tq* is the record length in years, *k* the number of stations in the region, λq the average number of flood peaks per year, and the 0.5 value comes from assuming the Hazen plotting position. The rank of the peak, *j*, goes from 1, smallest event, to (λq *tq*·*k*) for the largest one.

We consider actually separated circular regions constructed around each gridpoint with radius 50 km. In order to apply the plotting position to a region is important to refer to the total number of station as they would be one. For this reason instead of *tq*, the record length in years for the single station, we refer to *tq*·*k*, equivalent length in year given by all stations.

In a single region let us suppose *k* = 10 stations, *tq* = 20 years as record length for each of the *k* stations and λq = 5 events/year.

- For the largest event, the rank is $j = \lambda q\ tq \cdot k = 20 \cdot 10 \cdot 5$ and the return period becames:

  $T^r_j = (20 \cdot 10) / (20 \cdot 10 \cdot 5 + 0{,}5 - 20 \cdot 10 \cdot 5 + 0{,}5) = 400$ *years*

- For the smallest event, the rank is $j = 1$ and the return period becames:

  $T^{r}_{j} = (20 \cdot 10) / (20 \cdot 10 \cdot 5 + 0{,}5 - 1) = 0{,}2$ *years*

The Hazen plotting position, if applied to each point of the grid, gives for the maximum rainfall event, the results showed in Annexes (4).

The following *Figure 7.1.1-1* shows the return period in years for the all study region for the duration class of 24 hours.

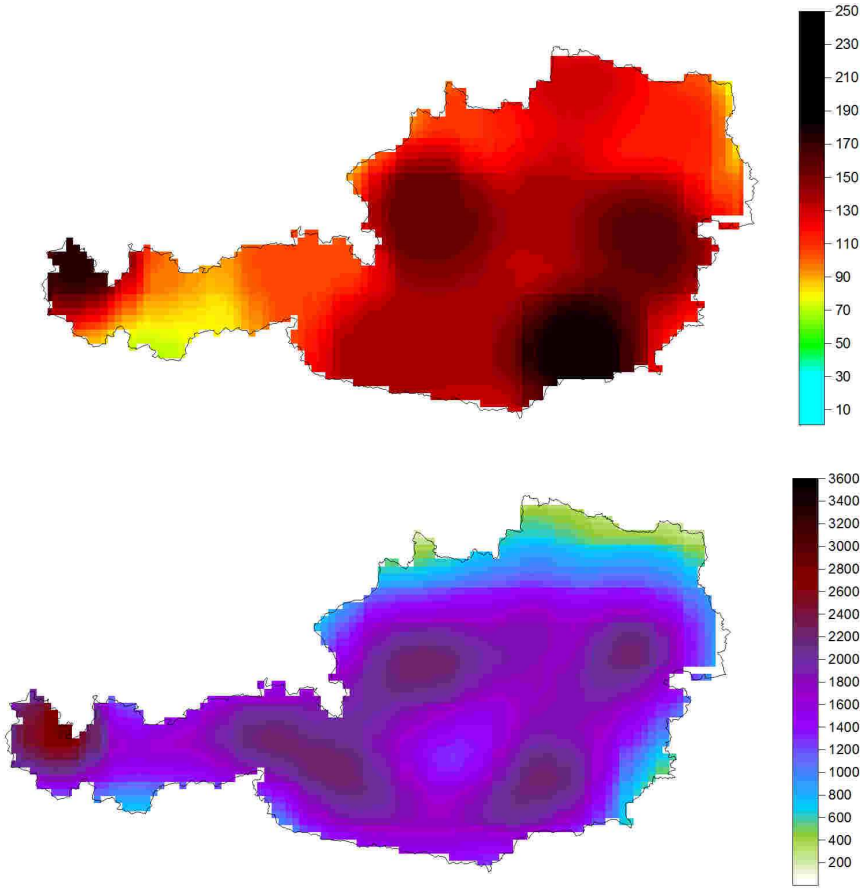


Figure 7.1.1-1 Daily rainfall depth (mm) and corresponding return period (years).

The REC concept is extended to extreme rainstorm events by introducing the Depth-Duration Envelope Curves (DDEC), which is defined as the regional upper bound on all the record rainfall depths at present for various rainfall duration.

Figure 7.1.1-2 presents the DDEC for the Austrian study region.

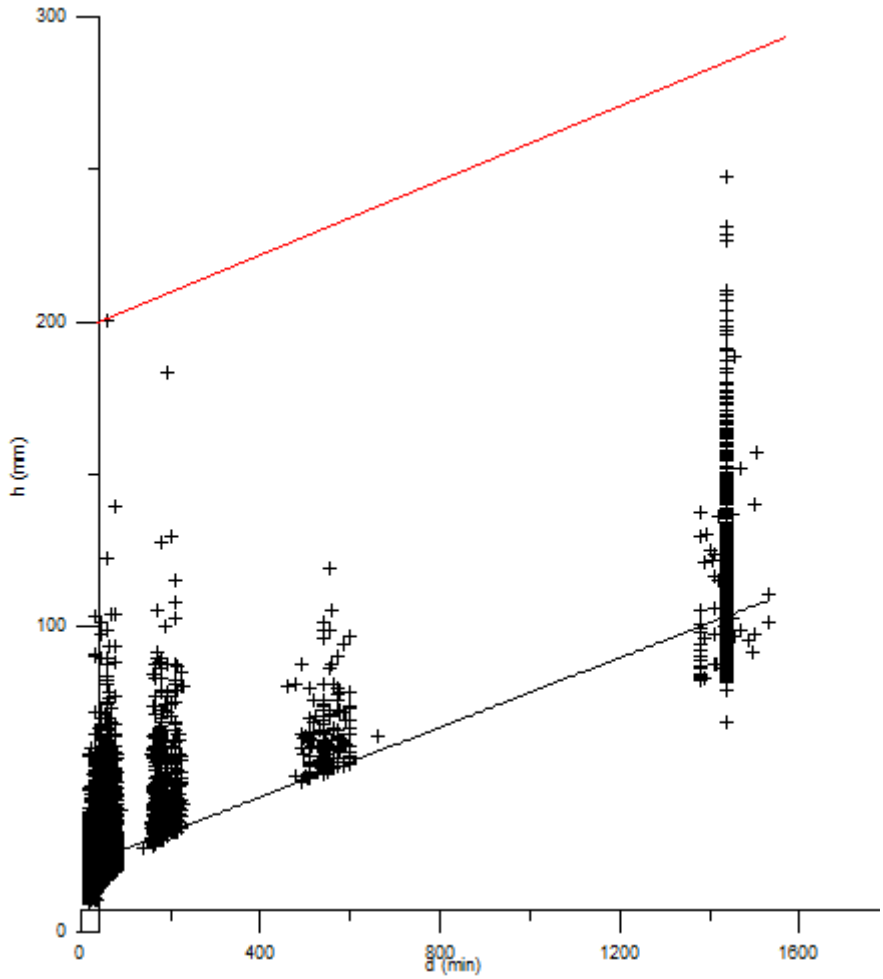


Figure 7.1.1-2 Depth-Duration Envelope Curve (red line); annual maximum rainfall depths (black +); trend-line for log(duration)-log(average rainfall depth).

7.1.2 Correlated Case

The estimation of the recurrence interval for cross-correlated series requires the quantification of the equivalent number of independent data which, in turn, is a function of the cross-correlation among sequences.

The effect of intersite correlation on the exceedance probability is discussed in order to make the probabilistic interpretation of envelope curves applicable to the real world data sets. This point is important because real world data sets generally

consist of sequences of rainfall data that are not concurrent nor with equal record lengths.

It is well known that intersite correlation leads to increases in the variance of flood statistics [see, e.g., *Hosking and Wallis*, 1988], thus the probabilistic interpretation of envelope curves has to be corrected considering a suitable correlation and plotting position formula.

Castellarin et al., [2005] proposed an empirical estimator of the exceedance probability *p* that considers the spatial correlation between the regional data. The estimator evaluates the equivalent number of independent sequences M_{EC} for M concurrent and correlated sequences with equal length *n*.

The estimator reads,

$$\hat{M}_{EC} = \frac{M}{1+\overline{\rho^{\beta}}(M-1)} \qquad \text{with} \qquad \beta = 1.4\frac{(nM)^{0.176}}{(1-\rho)^{0.376}} \qquad (7.1.2\text{-}1)$$

where $\overline{\rho^{\beta}}$ and $\overline{(1-\rho)^{0.376}}$ are average values of the corresponding functions of the correlation coefficients (i.e. $\overline{\rho^{\beta}}$ is the average of the M(M-1)/2 values of $\rho^{\beta}_{k,j}$, where $\rho_{k,j}$ is the correlation coefficient between sites k and j, with $1 \le k < j \le M$).

The *p* value of the expected PREC can be computed by estimating M_{EC} through equation (7.1.2-1) and using an appropriate plotting position by setting the overall sample years of data to the equivalent number of independent years $n \cdot M_{EC}$.

A possible approach to modelling cross-correlation between rainfall events is to compute the sample correlation coefficients using sample estimators proposed in the scientific literature [*Stedinger, 1981*] and, in absence of more specific information about the spatial structure of the intersite correlation model, to approximate the true correlation coefficient $\rho_{k,j}$ through empirical correlation functions of the distance $d_{k,j}$ among sites *k* and *j* [see, *Tasker and Stedinger, 1989: Hosking and Wallis, 1988; Troutman and Karlinger, 2003*].

A simple correlation function that can be used is as follow:

$$\rho k, j = \exp\left(-\frac{dk, j}{\alpha}\right) \qquad (7.1.2\text{-}2)$$

where α is the formula regional parameter.

From the regression of the empirical values of $\rho_{k,j}$ and $d_{k,j}$ we assume α = 100km for the duration of 24 hours and α = 10km for the duration of 1 hour. *Figure 7.1.2-1* and *Figure 7.1.2-1* show the correlation coefficient trend for the durations of 1 hour and 24 hours.

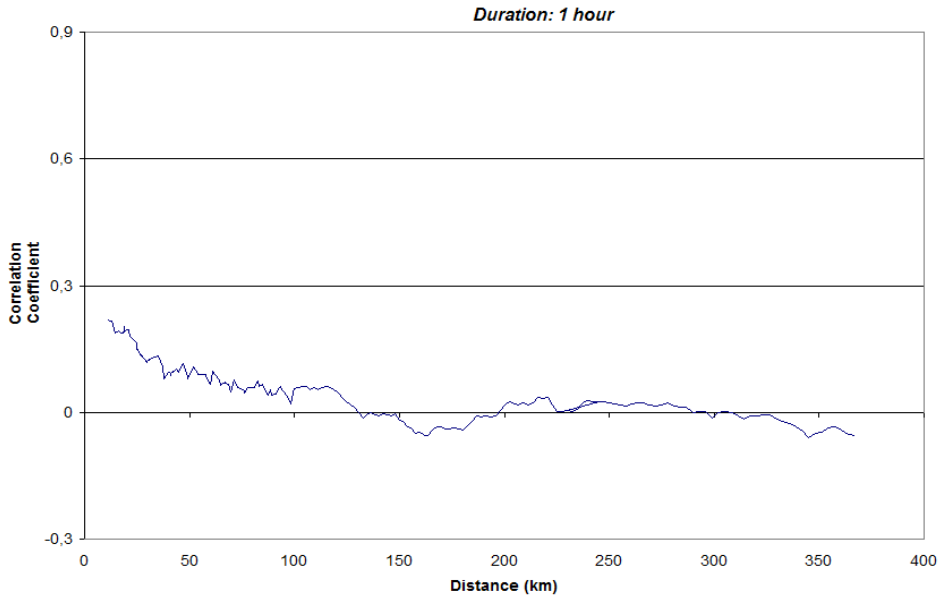


Figure 7.1.2-1 Empirical correlation coefficients, considered duration of 1 hour

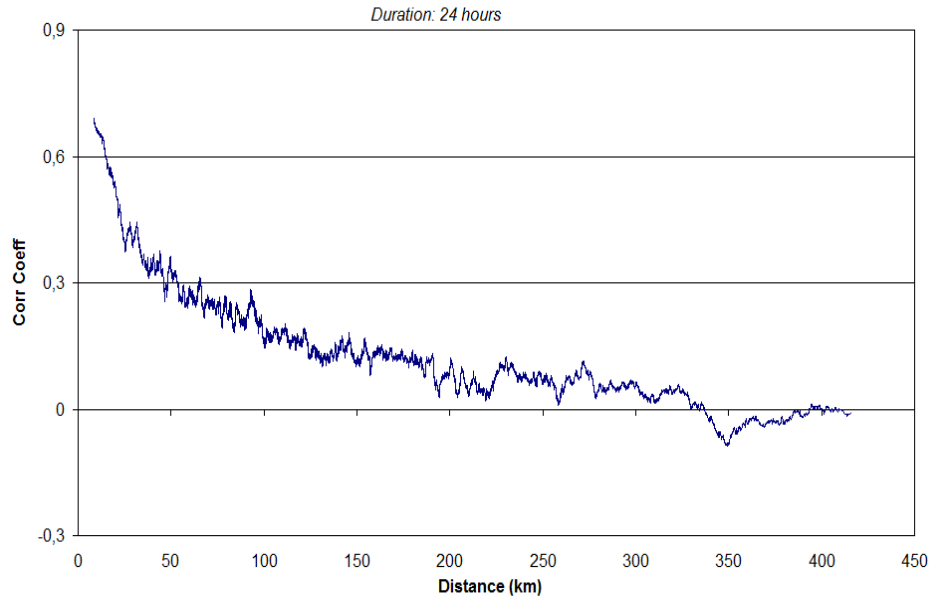


Figure 7.1.2-2 Empirical correlation coefficients, considered duration of 24 hours.

The parameter α for different durations will be extrapolated by a graph as in *Figure 7.1.2-3*:

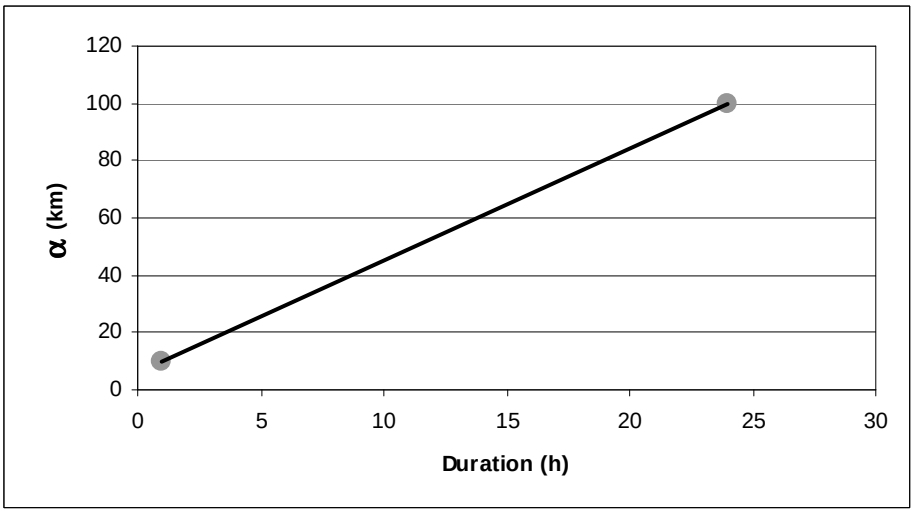


Figure 7.1.2-3 Visualisation the parameter α to be used in (7.1.2-2) for different durations in hours.

Duration (d)	Parameter α
0.5	8
1	10
3	18
9	42
24	100

Table 7.1-1 Estimates of the parameter α for different durations

The Hazen plotting position, if applied to each point of the grid and considering the intersite correlation, gives for the maximum rainfall event, the results showed in (5) and (6) in Annexes.

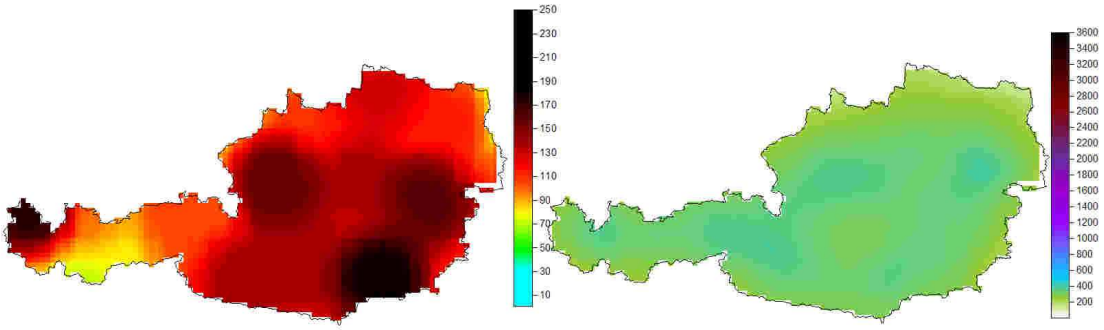


Figure 7.1.2-3: Daily rainfall depth (mm) and corresponding return period (years).

The following *Figure 7.1.2-3* and *Figure 7.1.2-4* show the return period in years for the all study region for the duration class of 24 hours, considering the intersite correlation, and the reduction in return period between the uncorrelated and correlated case.

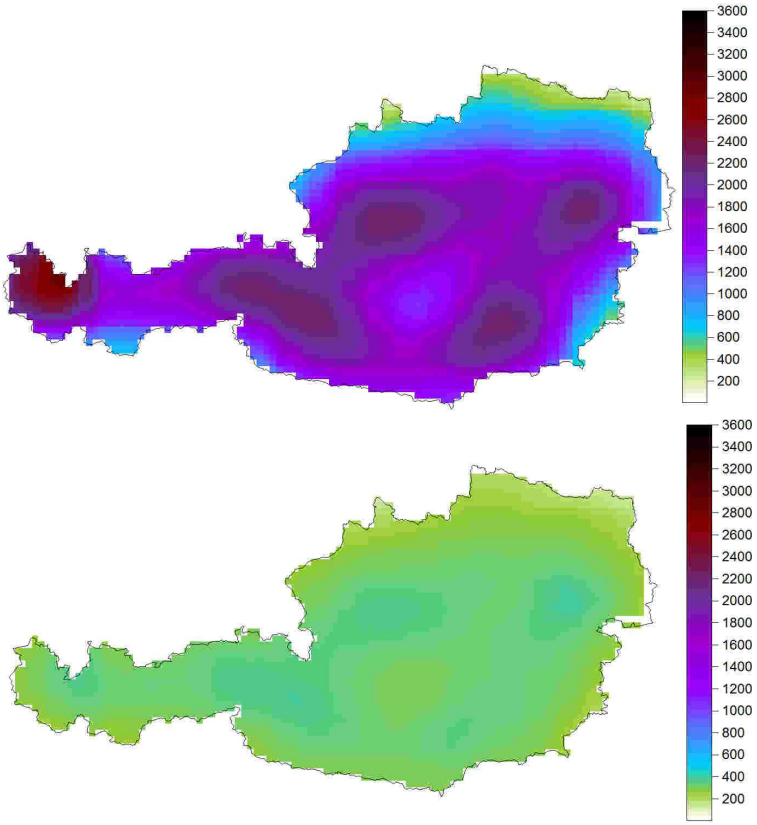


Figure 7.1.2-4: Return period in years for the duration class of 24 hours. Top: not correlated case; Bottom: correlated case.

Considering the spatial correlation between the series and the effective number M_{EC} of uncorrelated series in the region, the return period decreases.

This reduction will be higher with the increasing of the correlation between stations, that means that high durations will have higher reductions in terms of return period while low durations will have less reductions, for instance for the duration class of 30 minutes the recurrence interval in almost the same both in uncorrelated case and the correlated one.

The following *Table 7.1.2-2* illustrates the effects of spatial correlation on the recurrence interval:

Duration class (minutes)	Uncorrelated case: maximum T (years)	Correlated case: maximum T (years)	Decrease (%)
30	2960	2866	3,2
60	3040	2902	4,5
180	3240	2672	17
360	3560	2094	41
540	2400	1155	52
1440	3200	418	87

Table 7.1-2 Comparison for return periods between the not correlated and the correlated case, the effects of intersite correlation increase with high durations**.**

7.2 ITALIAN CATCHMENTS

For Italian catchments the procedure described in Paragraph 5.3 for flood flows probabilistic envelope curves is modified and reviewed in order to make it applicable to extremes rainfall events.

Instead of plotting for each site the normalized record flood, defined as the logarithm of the ratio of the record flood to its basin area versus the logarithm of the basin area, we plot for each site the normalized annual maxima rainfall versus the logarithm of the mean annual precipitation (MAP). This PREC considers also the intersite correlation among stations by using the Tasker and Stedinger regional cross-correlation formula.

The construction of the empirical PREC and the estimation of p starts by the definition of a homogeneous region, we consider a unique homogeneous region corresponding to the entire study region. An estimate of the PREC slope $\hat{b}$ is

obtained by regressing the empirical values of rainfall against the MAP value of the corresponding site. The value of the intercept *a* in equation (5.1-1) is computed as

$$a = \max_{j=1,..,M} \left\{ \ln\left(\frac{h_j}{MAP_j} \right) - \hat{b} \ln(MAP_j) \right\}$$

where h_j denotes the maximum rainfall depth observed at site ***j***=1,…,*M* and *M* is the number of sites in the region while MAP_j is the mean annual precipitation associated t site *j*. A suitable regional cross-correlation function is identified (the estimation of coefficients λ_1 and λ_2 of equation (5.3.2-1) is studied in deep in paragraph 7.2.1). Then $\hat{n}_{eff}$ is computed trough equation (5.3.3-1). A suitable regional parental distribution is chosen and a quantile-unbiased plotting position suitable for this distribution is utilized for estimating *p* as a function of $\hat{n}_{eff}$.

The GEV distribution was shown to be a suitable probabilistic model for representing the annual maximum rainfall event sequences in the study area.

Table 7.2-1 reports the estimates of the parameters λ_1 and λ_2 of the cross-correlation formula in equation (5.3.2-1) obtained for the hole study area and for each duration class. The estimates were obtained by applying a weighted least squares regression algorithm that uses, as weight for each sample cross-correlation coefficient between two sequences (sites), for the corresponding number of concurrent rainfall events, see *Figure 7.2-1*. *Table 7.2-1* also reports the estimates of the PREC's slope, $\hat{b}$, along with the values of the intercept *a*. Finally *Table 7.2-1* lists the plotting position parameters and the number of total observations and effective observations, out coming by the application of equation (5.3-1).

Figure 7.2-2 illustrate the envelope curves for the region for the duration class of 24 hours.

Duration (min)	1440	720	360	180	60	45	30	20	15
No. Of sites	223	223	223	223	223	189	222	213	220
No. Of observations	8052	8059	6667	7090	8047	941	3766	2057	2187
No. Of single observations	1	1	7	8	1	4	0	0	0
No. Of effective observations	3241.6	4769.6	5225.0	6283.3	7558.9	844.90	3645.60	1933.40	2044.90
Tasker & Stedinger λ_1	0.0359	0.0562	0.0948	0.1460	0.2060	0.1960	0.3290	0.8520	0.2120
Tasker & Stedinger λ_2	0.0100	0.0175	0.0329	0.0508	0.0665	0.0622	0.1040	0.3520	0.0662
Mean cross-correlation (model)	0.1736	0.1349	0.1087	0.0912	0.0703	0.1178	0.0645	0.0966	0.0719
Recurence interval (years)	6483	9539	10450	12567	15118	1690	7291	3867	4090
Plotting-position parameter	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
DDEC slope	-0.282	-0.361	-0.530	-0.581	-0.763	-0.962	-0.9132	-0.038	-0.9866
DDEC intercept	1.0555	1.3645	2.2719	2.5417	3.1724	4.2034	3.5763	4.0452	3.8419

Table 7.2-1 Characteristics of the study region

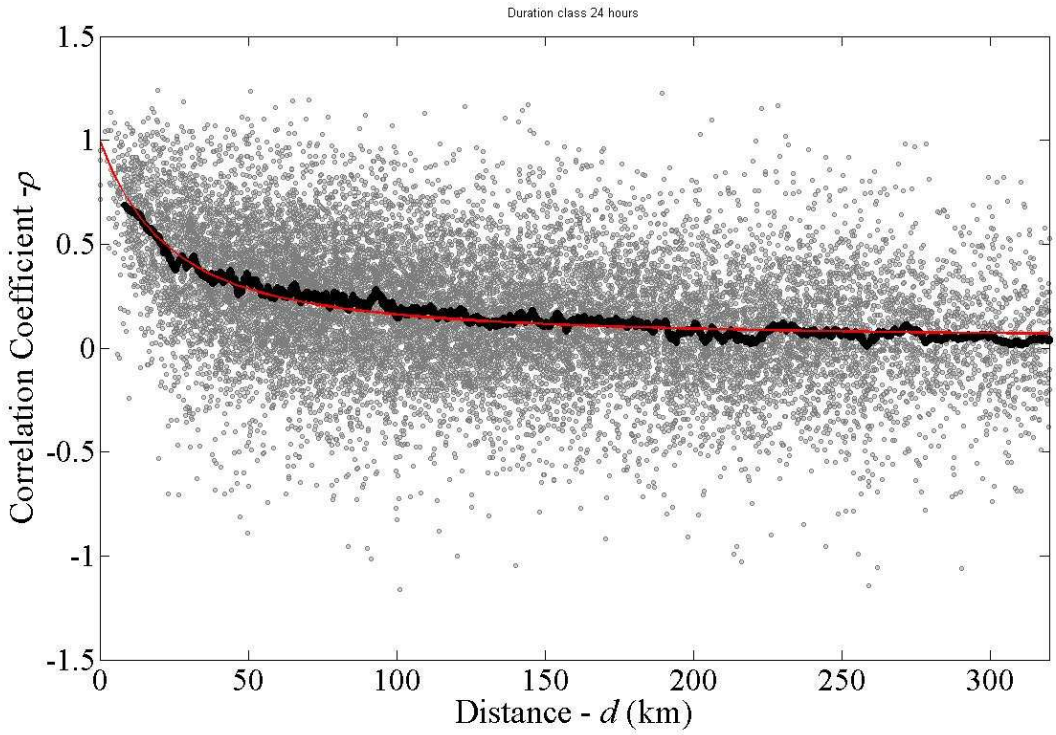


Figure 7.1.2-1 Tasker & Stedinger cross-correlation model (red line); empirical values (grey dots); moving average with width = 100 (black thick line)

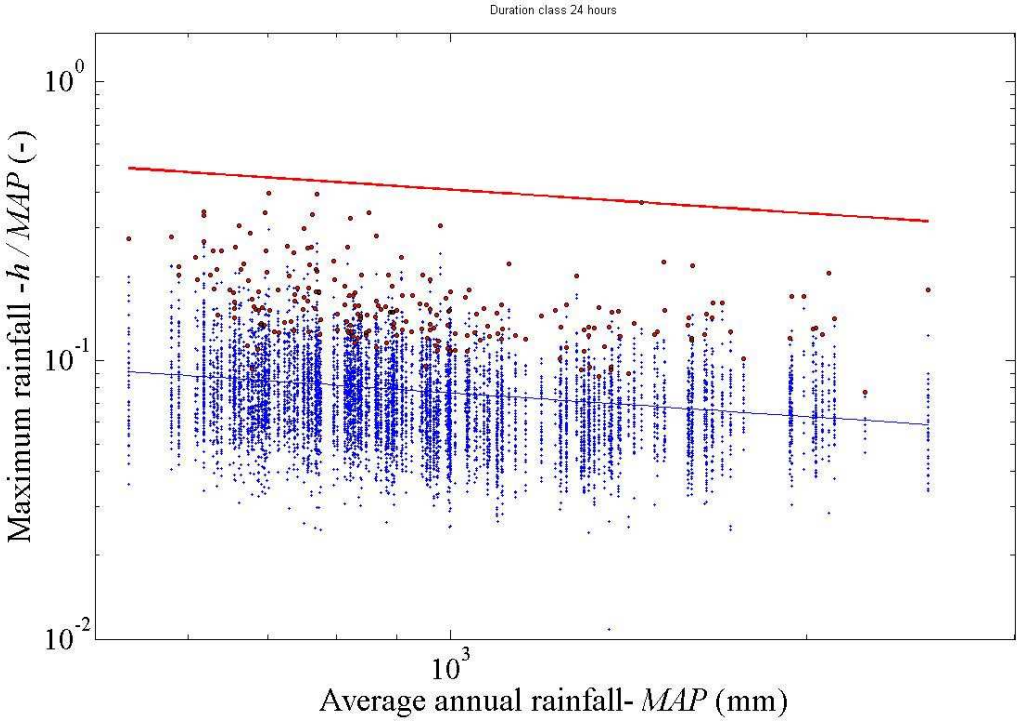


Figure 7.1.2-2 Empirical regional envelope curve for the study area and for the duration of 24 hours.

In Annexes (7) and (8) show the empirical envelope curves and the empirical cross-correlation coefficient for couple of stations in the hole study area, for all durations.

It has also been analysed the trends for DDEC intercept and slope with duration. While the DDEC slope increases, that means that the absolute value decreases, the intercept increases with the increasing of duration. Figures 7.1.2-3 and 7.1.2-4 present slopes and intercepts for hourly durations, it is possible in this way to extrapolate slope and intercept for every other duration. The two analysed DDEC parameters are approximated by a logarithmic model that minimize the gap between the empirical and values (grey dots) and the mathematically modelled values (black logarithmic line).

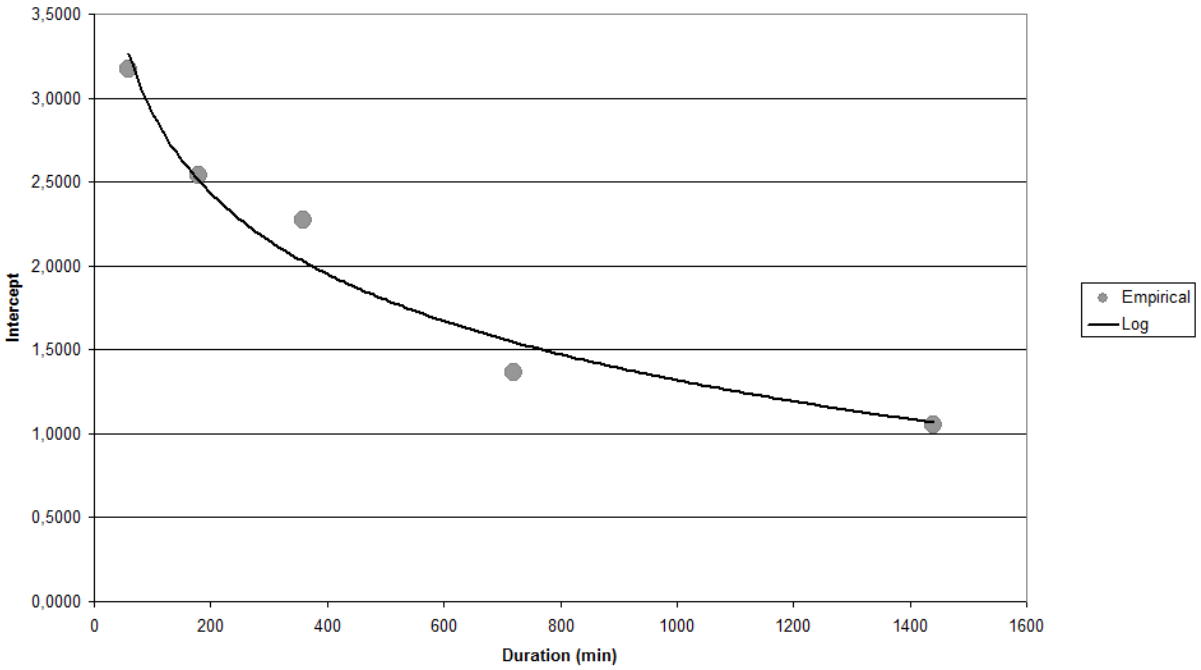


Figure 7.1.2-3 Empirical DDEC intercept (grey dots) versus duration, logarithmic modelling (continuous black line).

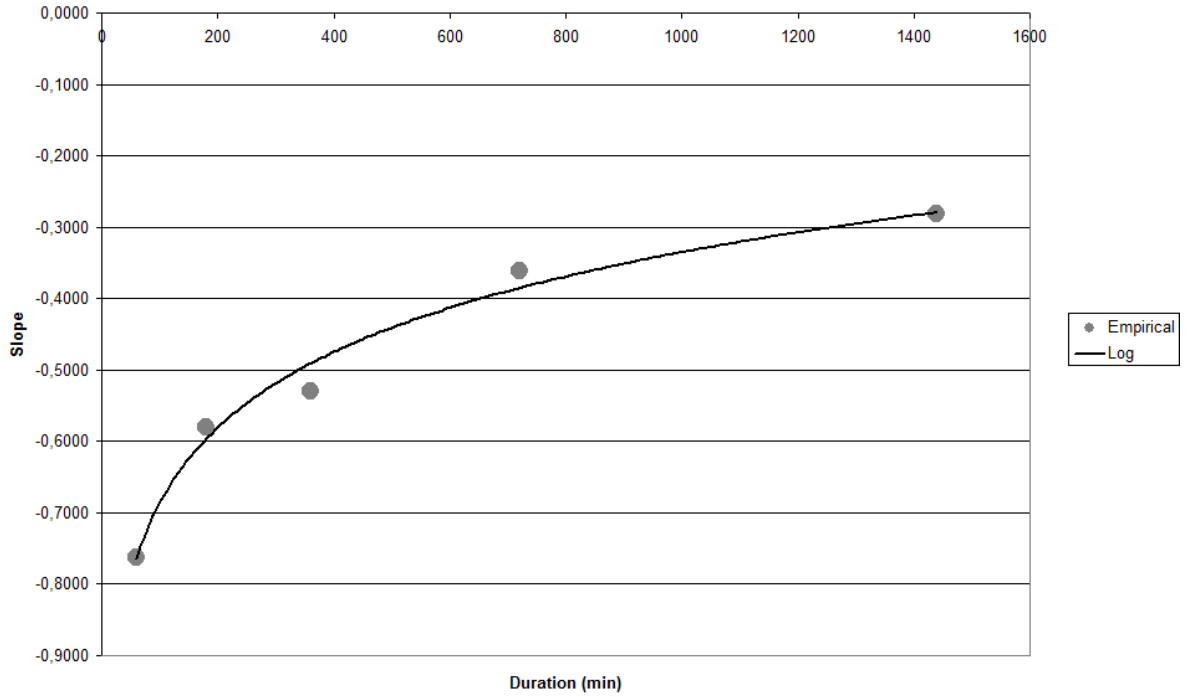


Figure 7.1.2-4 Empirical DDEC slope (grey dots) versus duration, logarithmic modelling (continuous black line).

7.2.1 Analysis and Modelling of Intersite Correlation

Regional frequency analysis usually assumes that records from different sites are statistically independent. This assumption is unlikely to be valid in practice so it is important to know how intersite dependence affects the estimates. We have computed sample estimators to modelling cross-correlation between rainfall events and the effective number of independent sites in our region.

This cross-correlation obviously is function of the storm duration, let us consider precipitation series at two nearby sites. One would expect that the time series will be cross-correlated because the sites are relatively close to each other and therefore subject to similar climatic and hydrological events. As the sites considered become farther apart their cross-correlation decrease. This tendency also decrease with slow durations, because the probability of having similar hydrological events decreases with the duration of the event itself.

Figure 7.2.1-1 shows this trend for all the analysed durations.

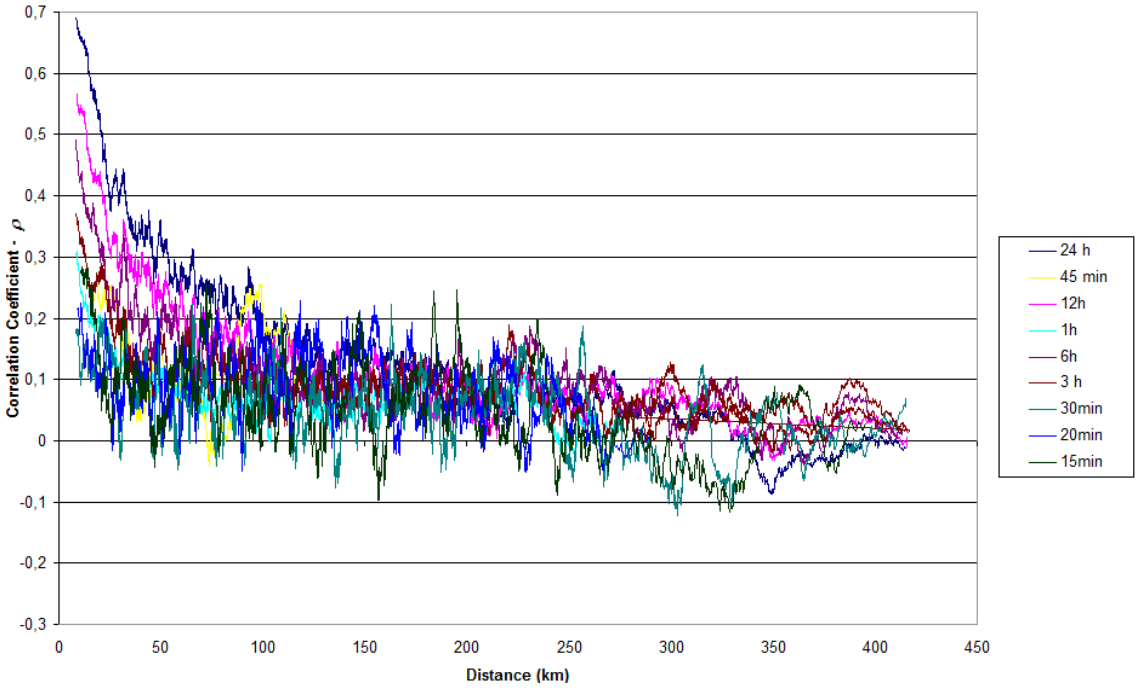


Figure 7.2.1-1 Tendency of the empirical cross-correlation coefficient with distance for all the rainfall durations, moving weighted average curve.

The figure above illustrates the tendency of the empirical correlation coefficient for all durations with distance. The trend of correlation coefficients with distance can be modelled by considering a suitable regional cross-correlation function.

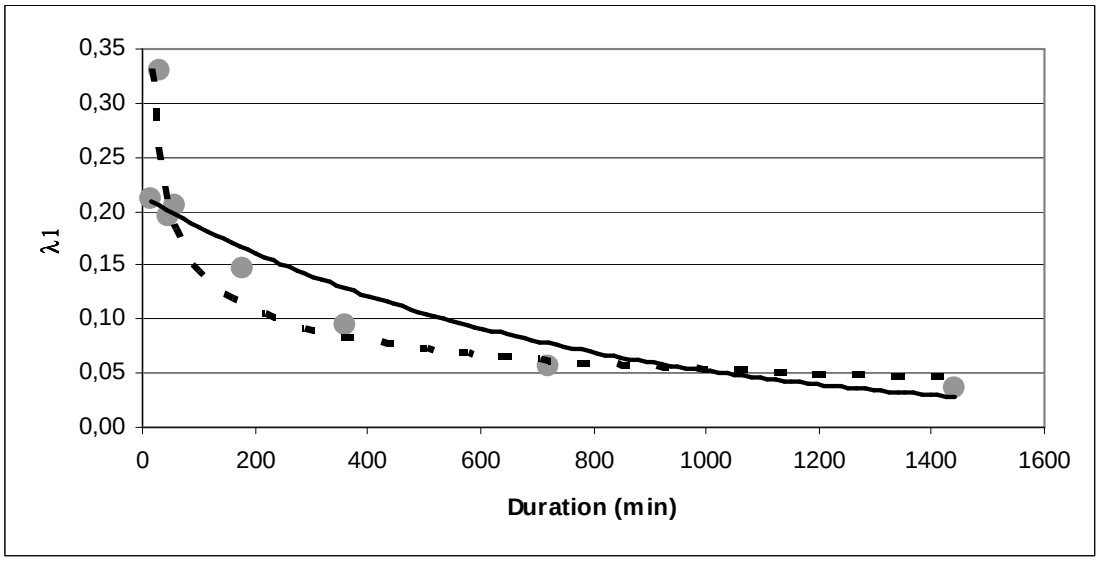
The proposed modelling (see *Tasker and Stedinger, 1989*). approximate the true correlation coefficient $\rho_{k,j}$ through empirical correlation functions of the distance $d_{k,j}$ among sites k and j . The used correlation function is as follow:

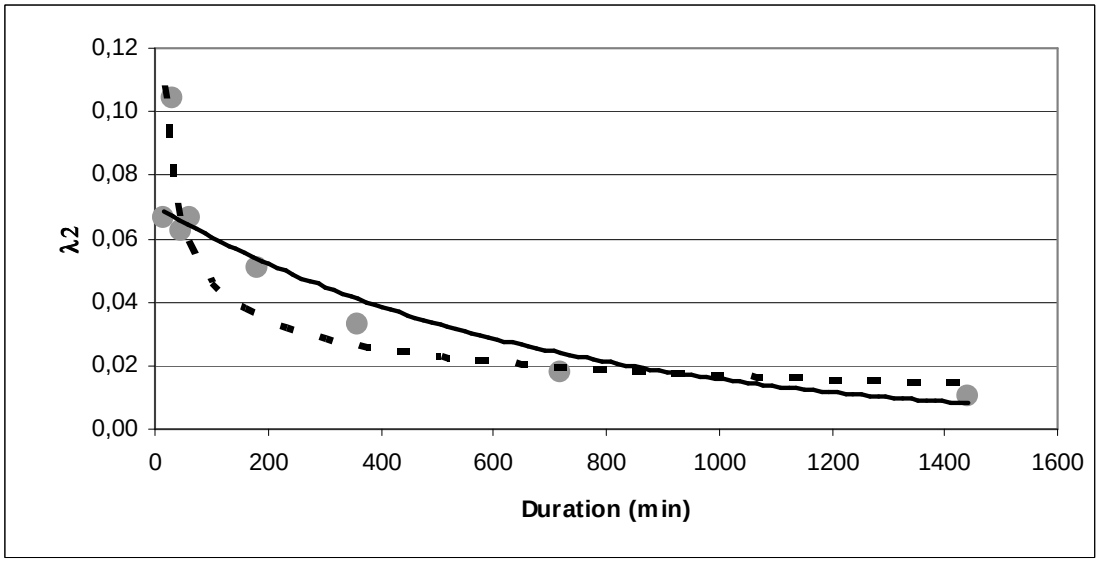
$$\rho k, j = \exp\left(-\frac{\lambda_1 d_{k,j}}{1+\lambda_2 d_{i,j}}\right)$$

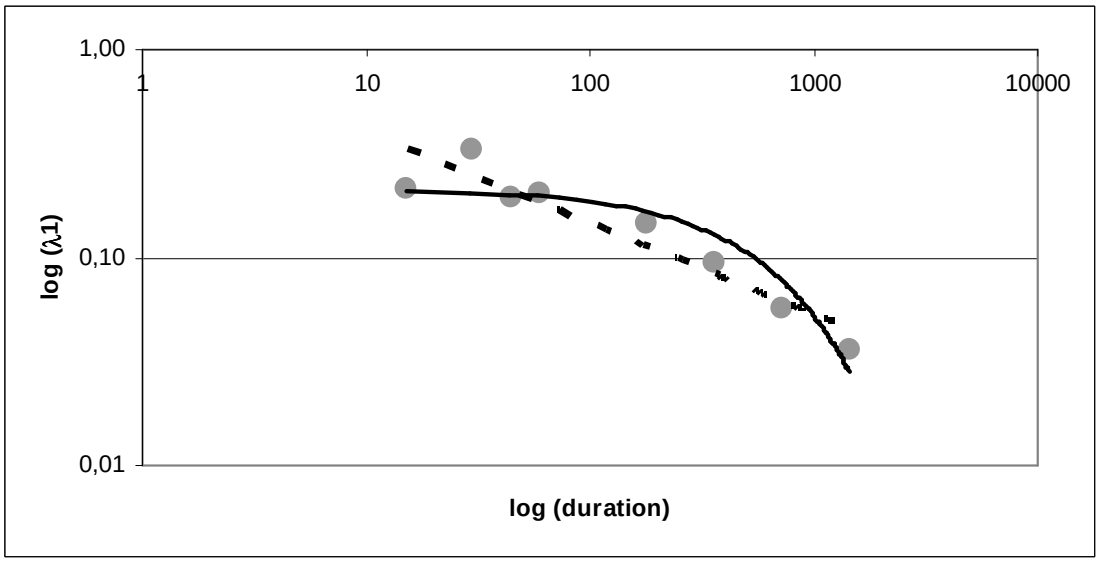
where $\lambda_1 > 0$ and $\lambda_2 \geq 0$ are the formula regional parameters of the formula.

The estimates of the parameters λ_1 and λ_2 are obtained by applying a weighted least squares regression algorithm that uses, as weight for each sample cross-correlation coefficient between two sequences (sites), for the corresponding number of concurrent rainfall events.

The estimated parameters λ_1 and λ_2 are plotted in *Figure 7.2.1-2* as function of the duration, both in natural and logarithmic scale.




0,12
0,10
0,08
0,06
0,04
0,02
0,00
λ2
0
200
400
600
800
1000
1200
1400
1600
Duration (min)


1,00
0,10
0,01
log (λ1)
1
10
100
1000
10000
log (duration)

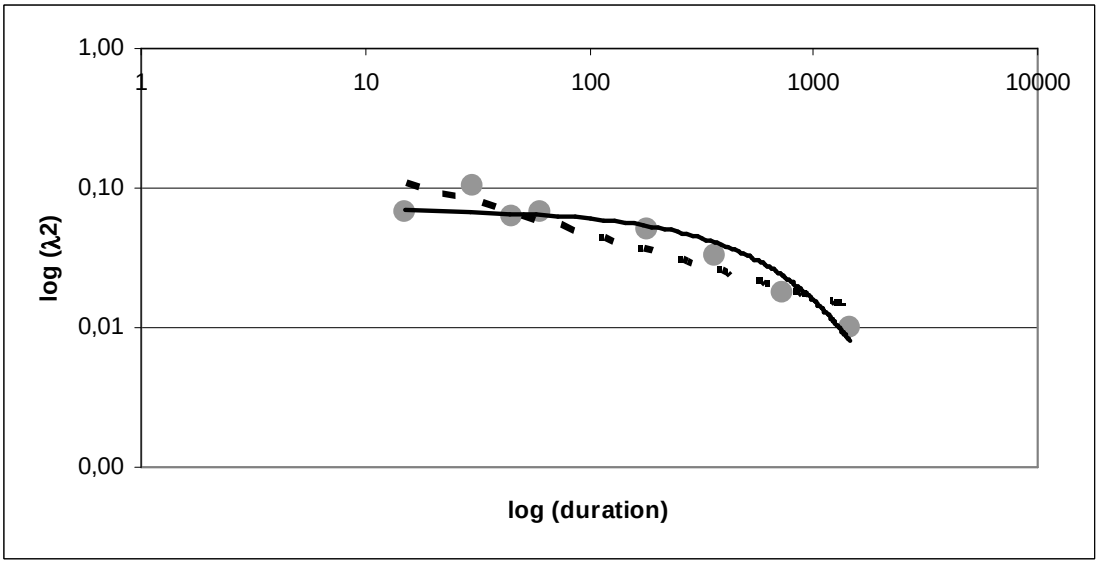


Figure 7.2.1-2 Parameters λ_1 and λ_2 to be used in equation (5.3.2-1), for couple of catchments (grey dots), exponential modelling (dashed line) and potential modelling (continuous black line) in natural and log-log scale.

Finding a trend for the two parameters of equation (5.3.2-1) would allow their extrapolation for every other duration, would be interesting to find the best estimation and at this aim two different tendencies have been analysed: the exponential and potential trends are shown in Figure 7.2.1-2.

The parameters λ_1 and λ_2 estimated from the two different modelling are used in equation (5.3.2-1) giving different values of correlation coefficients: the following Figure 7.2.1-3 presents the obtained cross-correlation coefficients together with the values of cross-correlation given by applying a weighted least squares regression algorithm and the original moving weighted average curve for the duration of 24 hours.

In conclusion, the estimation of λ_1 and λ_2, at first given by applying a weighted least squares regression algorithm, is then obtained by an exponential and potential model. It is possible to conclude that the potential model has a better approximation of the parameters, because of the higher value of the Nash-Sutcliffe efficiencies *E*.

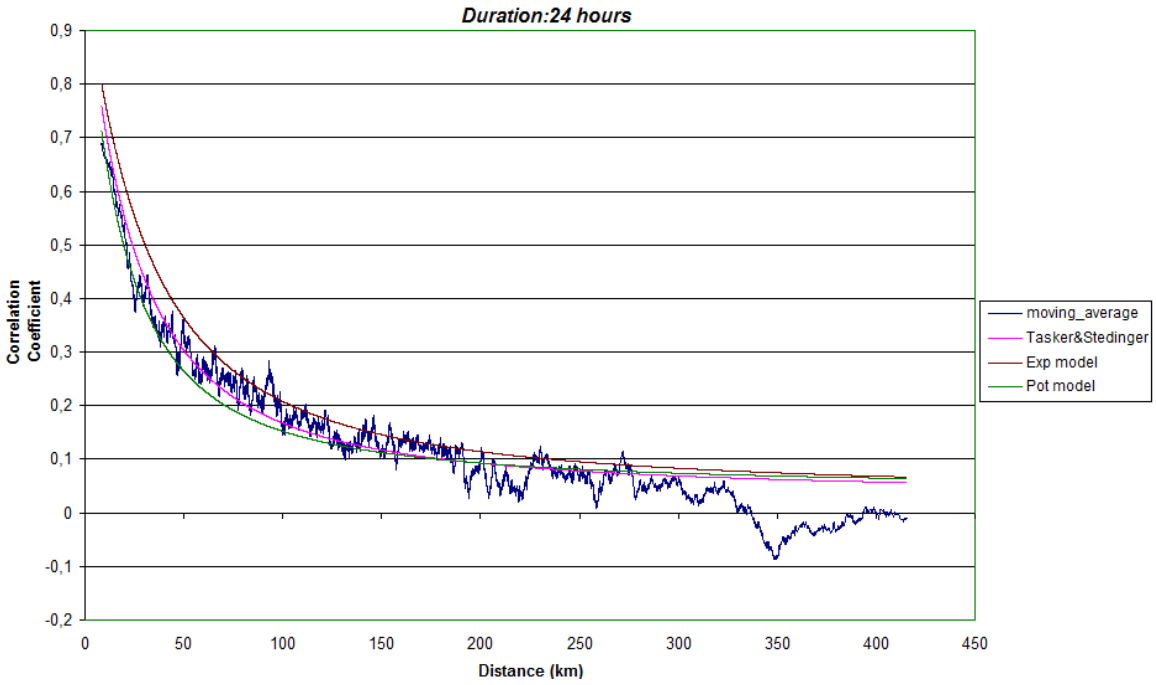


Figure 7.2.1-3 Moving weighted average curve for empirical cross-correlation coefficients (blue line), correlation formula (5.3.2-1) calibrated for the whole study area (pink line), correlation exponential formula (black line) and correlation potential formula (green line) calibrated for the whole study area.

The same Figure is present in Annexes for all other durations as (9).

The reliability of these estimates has been quantified trough the Nash-Sutcliffe formula, that defines the efficiency *E* as:

$$E = 1 - \frac{\sum_i (\rho_{i,emp} - \rho_{\mathrm{mod}})^2}{\sum_i (\rho_{i,emp} - \overline{\rho_{emp}})^2}$$

where $\rho_{i,emp}$ is the empirical estimated correlation coefficient, ρ_{mod} is the corresponding correlation coefficient estimated by the model which efficiency we want to analyse, and $\overline{\rho_{i,emp}}$ is the mean of all the empirical $\rho_{i,emp}$.

The Nash-Sutcliffe model efficiency coefficient is used to assess the predictive power of hydrological models and gives us the idea of the accuracy of the model estimates. Nash-Sutcliffe efficiencies can range from -∞ to 1. An efficiency of 1 (*E=1*) corresponds to a perfect match of modelled variable to the observed data. An efficiency of 0 (*E=0*) indicates that the model predictions are as accurate as the mean

of the observed data, whereas an efficiency less than zero ($-\infty<E<0$) occurs when the observed mean is a better predictor than the model. Essentially, the closer the model efficiency is to 1, the more accurate the model is.

This method can be used to describe the predicative accuracy of other models as long as there is observed data to compare the model results to. In other applications, the measure may be known as the Coefficient of determination, or R^2. For example, Nash-Sutcliffe efficiencies have been reported in scientific literature for model simulations of discharge, and water quality constituents such as sediment, nitrogen, and phosphorus loadings.

The values of efficiency are calculated for each duration and model, they are shown in Table 7.2.1-1 below.

	E(T&S)	*E(pot)*	*E(exp)*
d (min)			
1440	0,927	0,896	0,835
720	0,915	0,906	0,775
360	0,780	0,759	0,553
180	0,658	0,607	0,552
60	0,633	0,602	0,609
45	0,336	0,334	0,320
30	0,131	-0,058	-0,035
20	0,026	-0,315	-0,291
15	0,192	0,162	0,069
Average E	*0,511*	*0,433*	*0,376*

Table 7.2-2 Nash-Sutcliffe efficiency for different modelling of intersite correlation.

From the results put in the Table above, it is reasonable to conclude that the potential model has a better approximation of the parameters λ_1 and λ_2, actually the efficiency increases for the potential modelling and it is higher on average.

Small durations, in particularly 20 minutes, have a smaller efficiency. For this reason *Figure 7.2.1-2* does not consider the 20 minutes duration for the estimation of parameters, thus the estimates for the duration 20 minutes, because of the sample variability, will be less accurate.

7.2.2 Validation Procedure

The assessment of the accuracy of PREC quantiles needs a comparison term.

Another regional estimate of the T_{EC}-year rainfall is used as reference value for the comparison with the PREC rainfall depth. The rainfall depths obtained by the DDEC will be compared with the rainfall depths estimated using regional depth-duration-frequency equations (RDDFEs) developed for the same catchment in northern-central Itlay, see Brath, Castellarin, Montanari, 2003.

The developed RDDFEs can be used in any location of the study area for estimating the *T*-year storm for values of durations t ranging from 1 to 24 hours.

We have considered the durations of 1 hour and 24 hours and proceeded with a spot analysis in order to check in the neighbourhood of 3 locations the accuracy of the PREC estimate. The given rainfall depth is compared with the one obtained by the RDDFE. by means of linear regression analyses, regional depth-duration equations (RDDEs) of the following type:

$$R_{T,t} = A(t,T,\Theta)R_{T,24hr} + B(t,T,\Theta) \qquad (7.2.2\text{-}1)$$

where $R_{T,t}$ and $R_{T,24hr}$ are the *T*-year rainfall depths with *t*- and 24-hour storm duration respectively, while *A* and *B* are regional coefficients that may be considered as varying with the storm duration, recurrence interval, and geographical location, expressed in equation (7.2.2-1) by the vector Θ.

The slope A and intercept B of equation (7.2.2-1) were initially assumed to be independent of geographical location and we obtain the following relation (7.2.2-2):

$$R_{T,t} = 0.138t^{0.624}R_{10yr,24hr}\left[f\ln\left(\frac{T}{10}\right)+1\right] + (24-t)^{0.770}\left[0.474\ln(T)+0.951\right]$$

where RT,t and $R_{10yr,24hr}$ are expressed in mm, f is equal to f_{TR} = 0.259 within the Tyrrhenian Region and is expressed as

$$f = 0.602 - 0.055\ln(MAP)$$

The rainfall depth to be compared with the obtained one will be calculated by equation (7.2.2-2) where $R_{T,24hr}$ is known for a given value of MAP in the study region, and T is obtained for the considered duration as seen in paragraph 7.1.

The considered stations, say stations i = 1,..,3, with relative values of MAPi are random selected and 20 stations are identified in its neighbourhood. The comparison is developed by a *Jackknife Procedure*: in turn, the considered gaging stations, say, station i, is removed from the set of gages and the T-year storm is estimated by the DDEC for the given duration.

The analysed gaging stations with corresponding values of MAP are shown in *Table 7.2.2-1* together with a quantitative analysis of the overall reliability of the DDEC estimates of the design storm, which relative residuals that has been evaluated as follows:

$$\varepsilon_{T,t,i} = \frac{\hat{R}_{T,t,i} - R_{T,t,i}}{R_{T,t,i}}$$

where $\hat{R}_{T,t,i}$ is the jackknifed DDEC estimate of the T-year t-hour design storm for a given site i, and $R_{T,t,i}$ is the corresponding estimate by equation (7.2.2-2).

t (h)	MAP(mm)	f	T (years)	$R_{10yr,24h}$	$R_{T,t,i}$ (mm)	$\hat{R}_{T,t,i}$ (mm)	**Δ**	εr
1	700	0,241691	1315	95,88	77,53376	75,81375	1,720008	0,022687
1	1100	0,241691	1529	133,85	90,42481	80,88799	9,536823	0,117902
1	1500	0,241691	1247	141,41	90,69418	82,61627	8,077913	0,097776

t (h)	MAP(mm)	f	T (years)	$R_{10yr,24h}$	$R_{T,t,i}$ (mm)	$\hat{R}_{T,t,i}$ (mm)	**Δ**	εr
24	700	0,241691	978	95,88	202,6084	278,4816	-75,8733	0,272453
24	1100	0,241691	1085	133,85	286,212	246,3258	39,88619	0,161924
24	1500	0,241691	741	141,41	289,3106	556,4372	-267,127	0,480066

Table 7.2-3 Values of considered MAP in which neighbourhood 20 stations have been considerate, T and $\hat{R}_{T,t,i}$ are estimated by the empirical DDEC while $\hat{R}_{T,t,i}$ is calculated.

8 DISCUSSIONS AND CONCLUSIONS

This work deals with the definition of probabilistic envelope curves of extremes rainfall events for the Austrian national rainfall dataset and for a wide north-central Italian region.

The evaluation of the frequency regime of extreme rainfall events is a fundamental task for the evaluation of the flooding potential in natural catchments and urban areas, which, in turn, are required steps for the identification and design of flood risk mitigation measures and policies.

The present study focuses on extreme rainfall events extending in this context the idea of a probabilistic interpretation of regional envelope curves (REC) that was recently proposed for record floods by the scientific literature (*Castellarin 2007*). The study is structured in two main parts. First depth-duration envelope curves (DDECs) are constructed for Austria and an Italian Region. Then estimates of the recurrence interval associated to a particular DDEC are computed using an adaptation of the algorithm proposed by Castellarin (2007). In principle, this makes DDECs useful graphical tools that can be employed for estimating design rainstorms for a given duration, any location in the study area and high or very high recurrence intervals.

The study analysed a comprehensive national dataset of peaks over threshold (POT) series of rainfall depths recorded at roughly 700 rainfall Austrian stations for duration spanning from 30 minutes to 24 hours. It is evaluated the maximum probable rainfall event for the given duration and the return period *T* of the expected DDEC. At first *T* is determined for the entire POT series as not spatially correlated by applying the Hazen plotting position to each investigated region. The estimation of the recurrence interval of the expected DDEC requires the selection of a cross-correlation formula to model the intersite correlation among the rainfall peak series and a plotting position.

The estimation of the recurrence interval for cross-correlated series requires the quantification of the equivalent number of independent data which, in turn, is a function of the cross-correlation among sequences.

The followed approach for modelling cross-correlation between rainfall events is to approximate the true correlation coefficient $\rho_{k,j}$ through empirical correlation functions of the distance $d_{k,j}$ among sites k and j, as function of the formula regional parameter α. From the regression of the empirical values of $\rho_{k,j}$ and $d_{k,j}$ for the durations of 1 and 24 hours, we extrapolate the value for α for all other duration. It is in this way defined for each duration class and site in Austria the return period T associated to the maximum probable rainfall event.

Then a wide geographical region located north-central Italy has been investigated. For this region annual maximum series (AMS) of rainfall depths with duration spanning from 15 minutes to 24 hours are available for a rather dense gauging network (more than 200 gauges). The empirical estimator of the recurrence interval T associated with a DDEC, which, in principle, enables us to use DDECs for design purposes in ungauged sites is derived by adapting the algorithm proposed for probabilistic regional envelopes curves (PRECs) by the scientific literature. The DDEC constructed for the Italian study region relative to a particular duration t describes the upper bound of the record rainfall depth with duration t divided by the local mean annual precipitation (MAP) as a function of the MAP itself. MAP can be considered as representative of the spatial location. As a matter of fact, instead of considering the physical coordinates for points it is possible to identify the site by the value of MAP.

Intersite correlation among AMS recorder at different raingauges for different duration is modelled by using the Tasker and Stedinger regional cross-correlation formula which parameters λ_1 and λ_2 have been deeply analysed in order to make them estimable for each duration. The same analysis has been developed for DDEC intercept and slope with duration

The reliability of the rainfall quantiles (rainfall depths associated with a given duration and recurrence interval) retrieved from DDEC is assessed relative to a a different estimating method proposed by the scientific literature. The quantitative

analysis of the overall reliability of the DDEC estimates of the design storm, includes the relative residuals evaluation.

The results of the analysis seem to indicate that the proposed DDECs represent practical tools to determine plausible extreme rainfall events at gauged and ungauged sites (*deterministic interpretation*) and can be used to provide a realistic estimate of the recurrence intervals associated with such rainfall events (*probabilistic interpretation*).

CONCLUSIONI

Il presente lavoro considera eventi meteorici estremi ed estende, in tale contesto, l'idea dell'interpretazione probabilistica delle curve inviluppo regionali (REC), recentemente proposte dalla letteratura scientifica per dati di portate estreme (*Castellarin, 2007*).

Lo studio è strutturato in due parti principali. Nella prima curve inviluppo di durata-altezza di precipitazione (DDECs) sono costruite per l'Austria e l'Italia. In seguito è stata valutata la stima del tempo di ritorno associato ad una particolare DDEC attraverso l'adattamento dell'algoritmo proposto da Castellarin (2007). In linea di principio tale approccio rende le DDECs utili strumenti grafici che possono essere usati per la stima degli eventi meteorici di progetto di assegnata durata, per ogni punto della regione in analisi e per alti o molto alti tempi di ritorno.

Tale studio analizza una completa banca dati nazionale di serie di picchi al sopra di una soglia (POT) di precipitazioni raccolte da 700 stazioni di registrazione in Austria per le durate comprese tra i 30 minuti e le 24 ore. E' stata valutata la precipitazione massima probabile di assegnata durata e periodo di ritorno dalla curva inviluppo attesa. Inizialmente il tempo di ritorno T è stato stimato per tutte le POT serie come non correlate dal punto di vista spaziale, tramite l'applicazione della plotting position di Hazen in ogni regione individuata. La stima del tempo di ritorno per la DDEC richiede comunque l'identificazione di una formula di cross-correlazione, al fine di modellare la correlazione spaziale tra le serie, e la valutazione dell'ottimale plotting position.

La stima del tempo di ritorno per serie correlate implica la quantificazione del numero equivalente di dati, che è a sua volta funzione della cross-correlazione tra le serie. L'approccio seguito per la modellazione della correlazione spaziale si traduce nell'approssimare il coefficiente di correlazione empirico $\rho_{k,j}$ tra i siti k e j posti a distanza $d_{k,j}$, mediante modelli di correlazione regionali, a loro volta funzione del parametro regionale α. Dalla regressione dei valori empirici di $\rho_{k,j}$ e $d_{k,j}$ per le durate di 1 ora e 24 ore, si sono estrapolati i valori di α per tutte le altre durate. Si è in tal

modo definito, per ogni classe di durata e per ogni sito in Austria, il tempo di ritorno associato all'evento meteorico massimo probabile.

Successivamente nelle regioni italiane di Emilia Romagna e Marche sono state considerate le serie dei massimi annuali di precipitazione osservati in una rete di stazioni per le duarte comprese tra i 15 minuti e le 24 ore. Lo stimatore empirico del tempo di ritorno associato alla DDEC, che consente l'uso della DDEC stessa per il progetto in bacini non strumentati, è derivato dall' adattamento, nella regione oggetto di studio, dell'algoritmo proposto dalla letteratura scientifica per le curve inviluppo regionali probabilistiche (PRECs). La DDEC costruita per la regione di studio italiana relativa a una data durata *t*, descrive il limite superiore delle altezze di precipitazione registrate con durata *t* , divisa per la precipitazione media annua (MAP) quale funzione del MAP stesso. Il MAP può infatti essere utilizzato come rappresentativo della posizione spaziale. Infatti, invece che valutare le coordinate UTM dei siti, si sono identificati gli stessi attraverso valori di MAP.

La correlazione spaziale tra i siti viene modellata utilizzando la formula di cross-correlazione regionale di Tasker e Stedinger, i cui parametri λ_1 e λ_2 ,sono stati accuratamente analizzati così da renderli stimabili graficamente per ogni possibile durata. La stessa analisi è stata affrontata per la pendenza e l'intercetta della DDDEC, così da rendere anche tali grandezze stimabili per ogni durata.

L'affidabilità dei quantili di precipitazione (altezze di precipitazione associate a una certa durata e tempo di ritorno) ricavati dalla DDEC, è infine analizzata tramite il confronto con un diverso modello di stima proposto dalla letteratura scientifica. Le altezze di precipitazione valutate dalla DDEC sono paragonate con le altezze di precipitazione calcolate attraverso curve di possibilità climatica costruite per lo stesso bacino italiano. Il risultato dell'analisi sembra indicare che le DDECs proposte costituiscano uno strumento pratico per la valutazione di probabili eventi di pioggia, da utilizzarsi quali eventi di progetto in siti strumentati e non (*interpretazione deterministica*), e possono inoltre fornire una stima realistica del tempo di ritorno associato a tale evento di pioggia (*interpretazione probabilistica*).

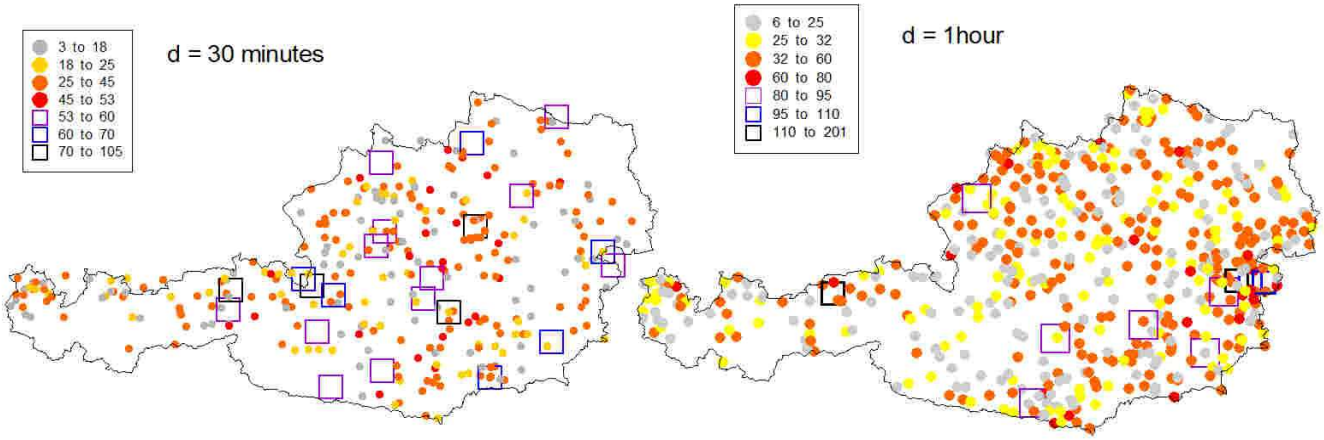
APPENDIX

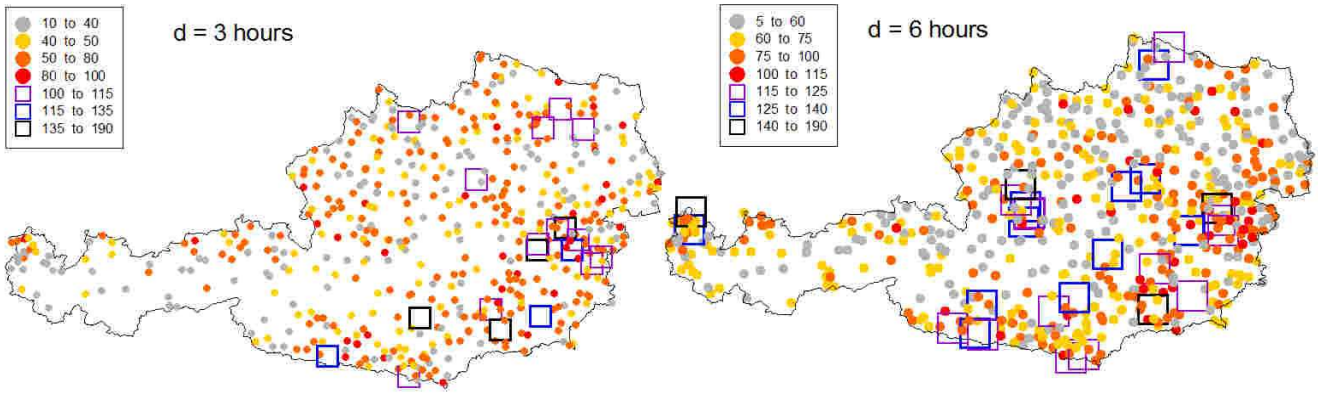
Table 7.2-1 Commonly Used Frequency Distributions in Hydrology

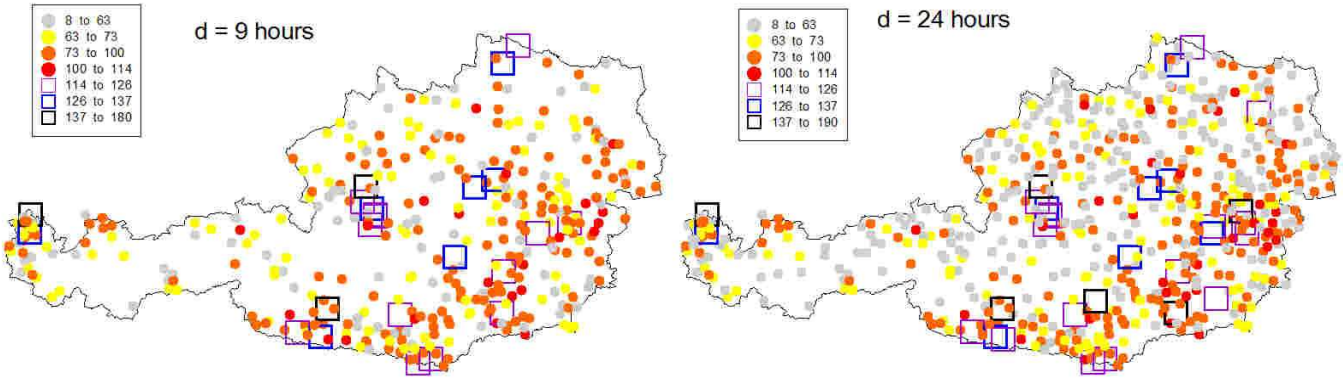
Distribution	pdf and/or cdf	Range	Moments
Normal	$fx(x) = \frac{1}{\sqrt{2\pi\delta_x^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu x}{\delta x}\right)^2\right]$	$\infty<x<\infty$	μx and σ^2; $\gamma x=0$
Lognormal	$fx(x) = \frac{1}{x\sqrt{2\pi\delta_Y^2}} \exp\left[-\frac{1}{2}\left(\frac{\ln(x)-\mu y}{\delta y}\right)^2\right]$	$x>0$	$\mu x=\exp(\mu y+\sigma y^2/2)$ $\gamma x=3CVx+CVx^3$
Pearson type 3	$fx(x) = \lvert\beta\rvert[\beta(x-\xi)]^{\alpha-1} \frac{\exp[-\beta(x-\xi)]}{\Gamma(\alpha)}$ $\Gamma(\alpha)$ is the gamma function for $\beta>0$ and $\xi=0$: $\gamma x=2CVx$	$\alpha>0$ for $\beta>0$: $x>\xi$ for $\beta<0$: $x<\xi$	$\mu x=\xi+\alpha/\beta;\sigma x^2=\alpha/\beta^2$ $\gamma x=2/\sqrt{\alpha}$ $\gamma x=-2/\sqrt{\alpha}$
Exponential	$fx(x) = \beta\exp[-\beta(x-\xi)]$ $Fx(x) = 1-\exp(-\beta(x-\xi))$	$x>\xi$ for $\beta>0$	$\mu x=\xi+\alpha/\beta;\sigma x^2=\alpha/\beta^2$ $\gamma x=2$
Gumbel	$fx(x) = \frac{1}{\alpha}\exp\left[-\frac{x-\xi}{\alpha}-\exp\left(-\frac{x-\xi}{\alpha}\right)\right]$ $Fx(x) = \exp\left[-\exp\left(-\frac{x-\xi}{\alpha}\right)\right]$	$-\infty<x<\infty$	$\mu x=\xi+0{,}5772\,\alpha$ $\sigma x^2=\pi^2\alpha^2/6\approx 1{,}645\,\alpha^2$ $\gamma x=1{,}1396$
GEV	$Fx(x) = \exp\left[-\left(1-\frac{k(x-\xi)}{\alpha}\right)^{1/k}\right]$	σx^2 exists for $k>-0{,}5$ when $k>0, x>(\xi+\alpha/\kappa); \kappa<0, \xi>(\xi+\alpha/\kappa)$	$\mu x=\xi+\alpha/k*(1-\Gamma(1+k))$ $\sigma x^2=(\alpha/k)^2\Gamma(1+2k)-(\Gamma(1+k))^2$
Weibull	$fx(x) = \left(\frac{k}{\alpha}\right)\left(\frac{x}{\alpha}\right)^{k-1}\exp\left[-\left(\frac{x}{\alpha}\right)^k\right]$ $Fx(x) = 1-\exp\left(-\left(\frac{x}{\alpha}\right)^k\right)$	$x>0$; α, $\kappa>0$	$\mu x=\alpha\,\Gamma(1+1/k)$ $\sigma x^2=\alpha^2(\Gamma(1+2/k)-(\Gamma(1+1/k))^2)$
Generalized Pareto	$fx(x) = \left(\frac{1}{\alpha}\right)\left[1-k\frac{(x-\xi)}{\alpha}\right]^{1/k-1}$ $Fx(x) = 1-\left[1-k\frac{(x-\xi)}{\alpha}\right]^{1/k}$	for $k>0$, $\xi\le x<\infty$ for $k>0$, $\xi\le x<\xi+(\alpha/\kappa)$	$\mu x=\xi+\alpha/(k+1)$ $\sigma x^2=\alpha^2/((1+k)^2(1+2k))$ $\gamma x=[2(1-k)(1+2k)^{1/2}]/(1+3k)$

ANNEXES

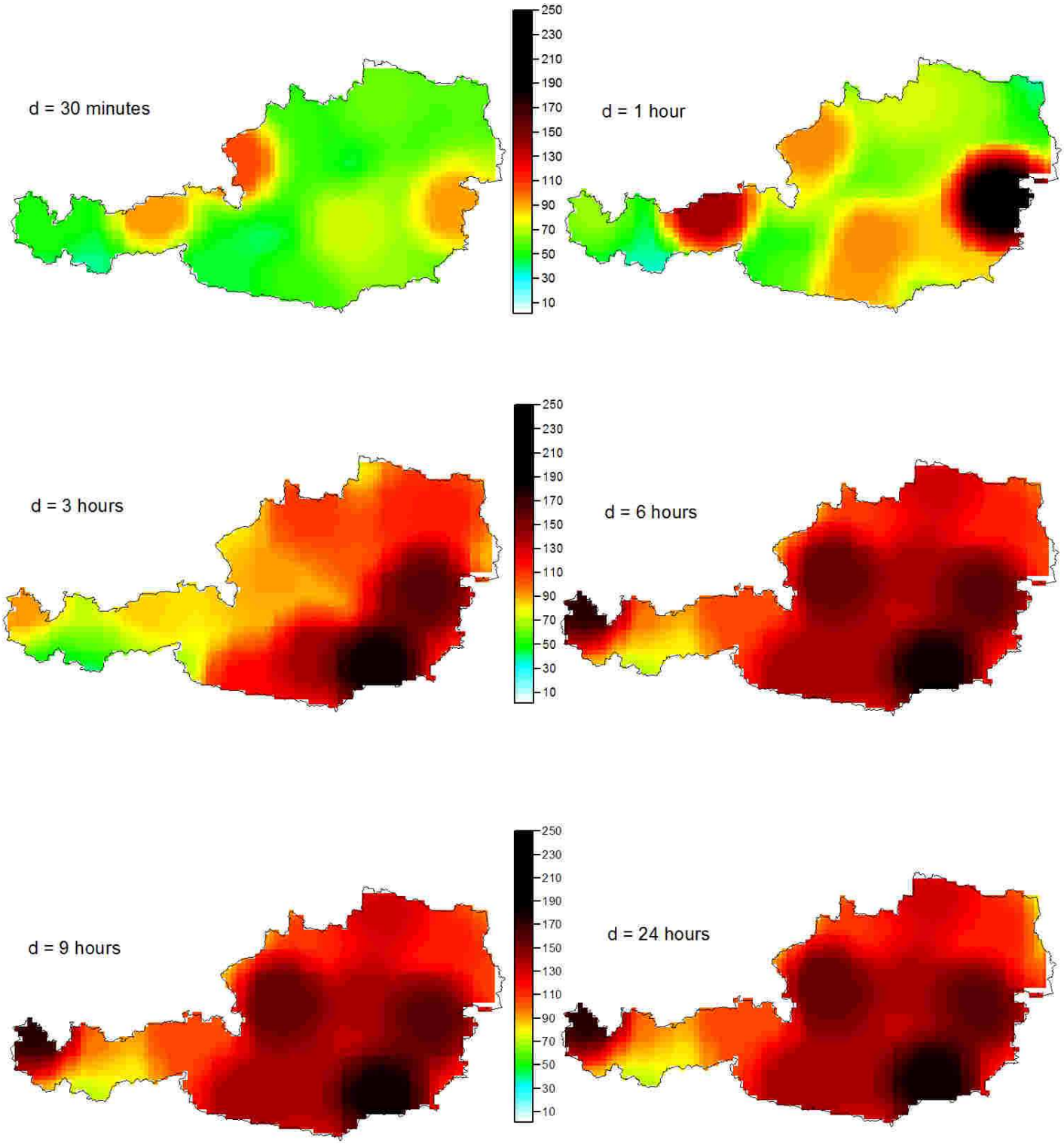
1. Collected Austrian data for each rainfall station and duration class



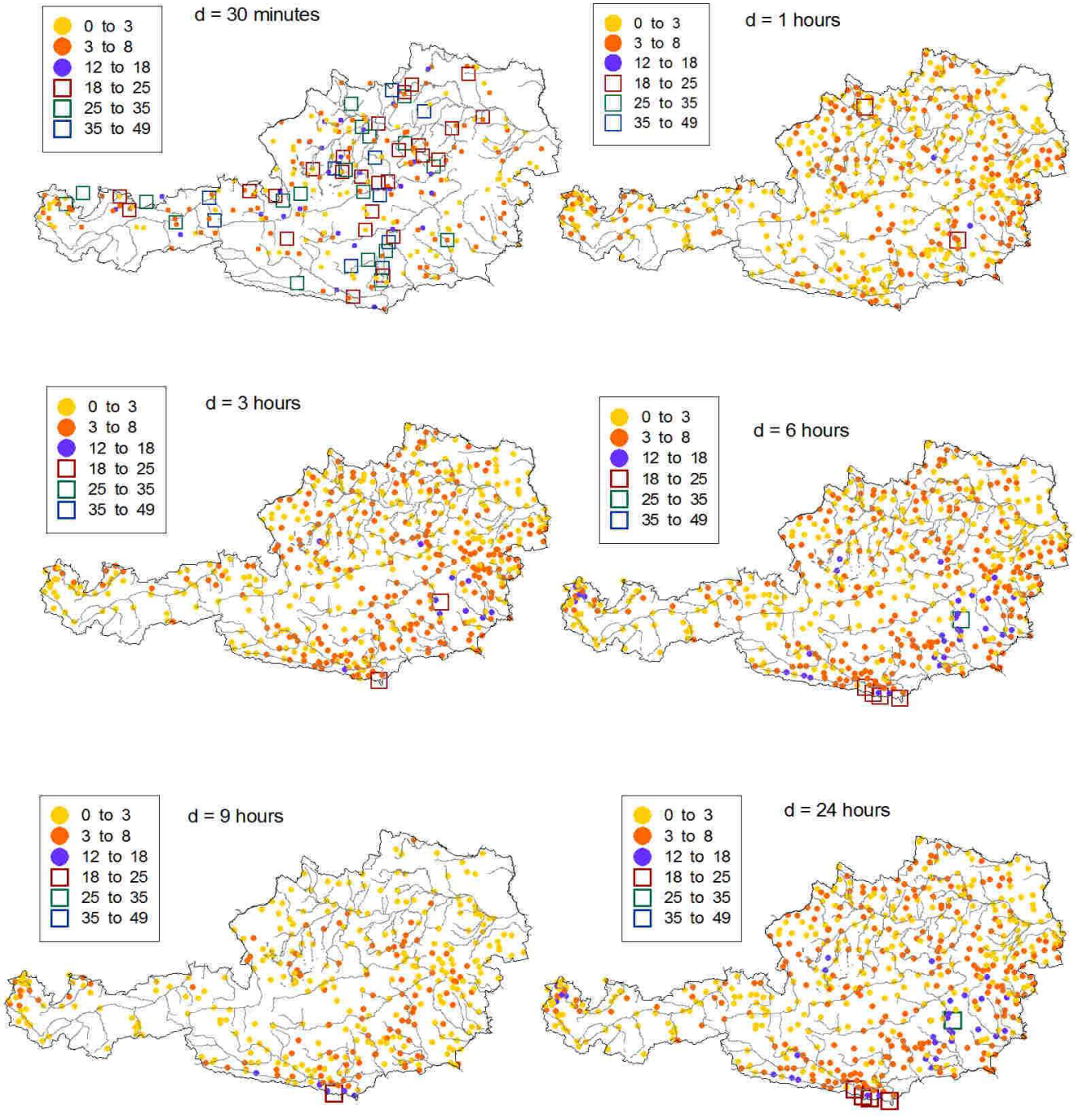




2. Maximum rainfall in distance 50 km by each gridpoint

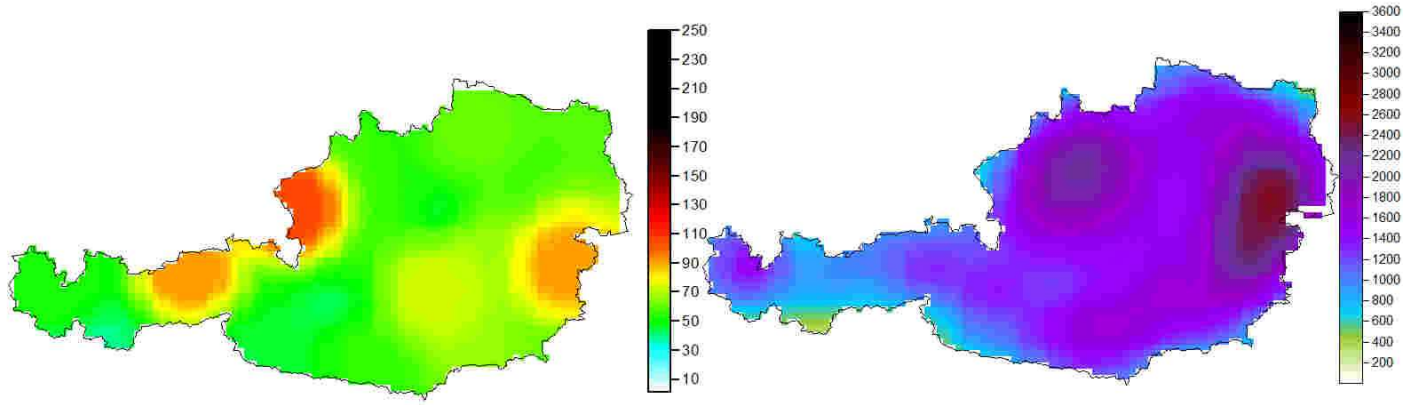


3. L*e*ngth in years of the series of Austrian data

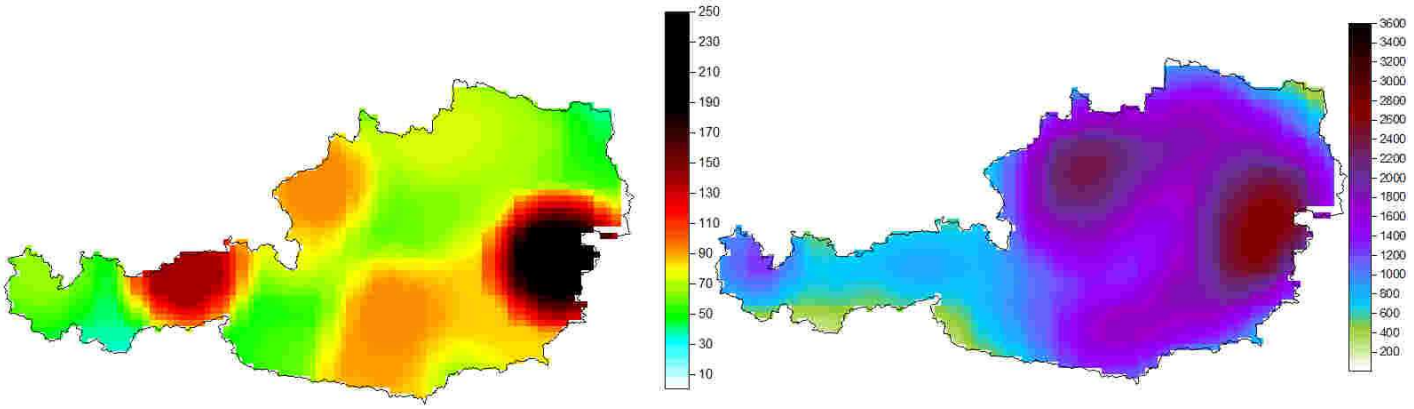


4. Return period in years for the maximum rainfall event in mm, it is not considerate spatial correlation among sites.

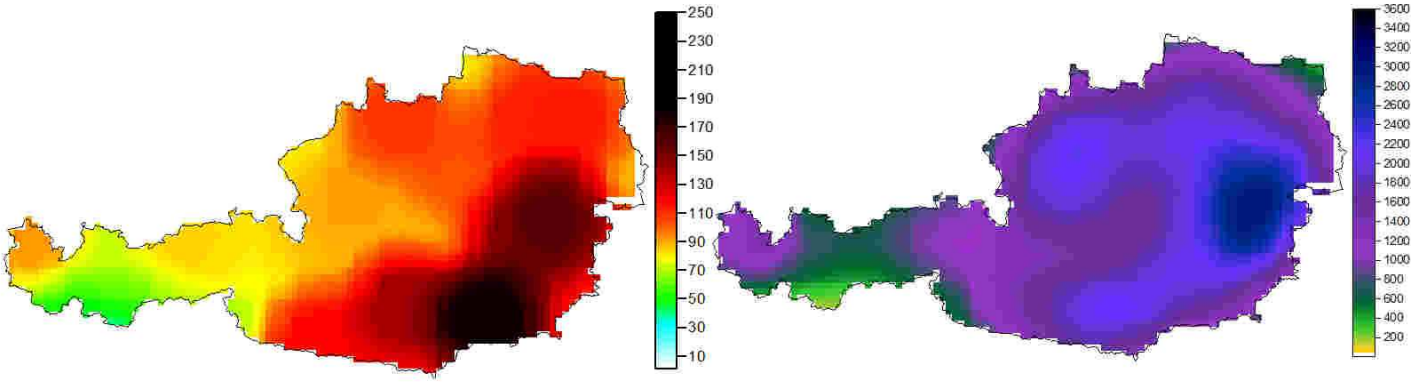
d = 30 minutes



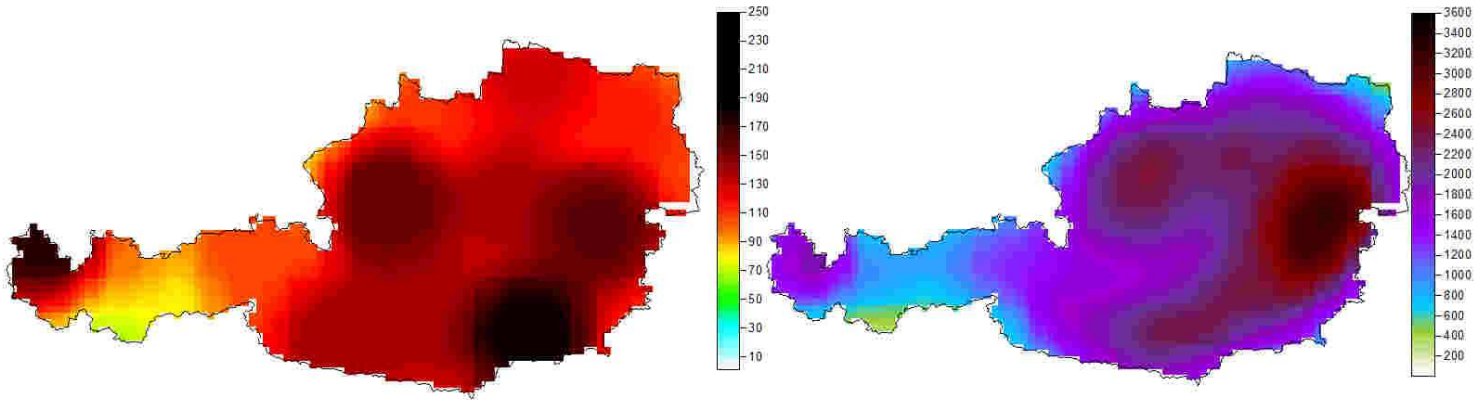
d = 1hour



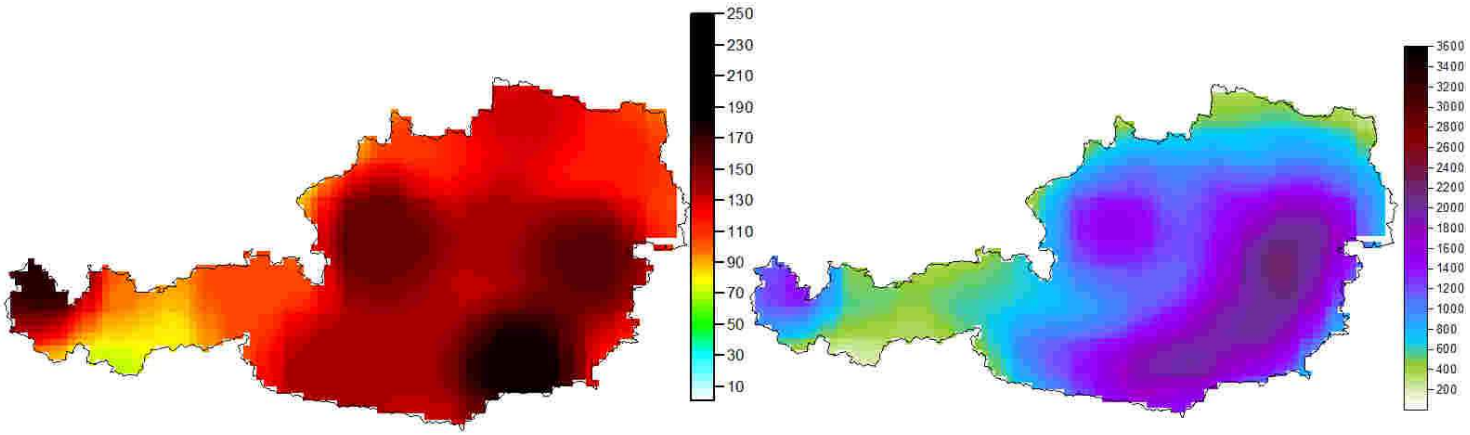
d = 3hours



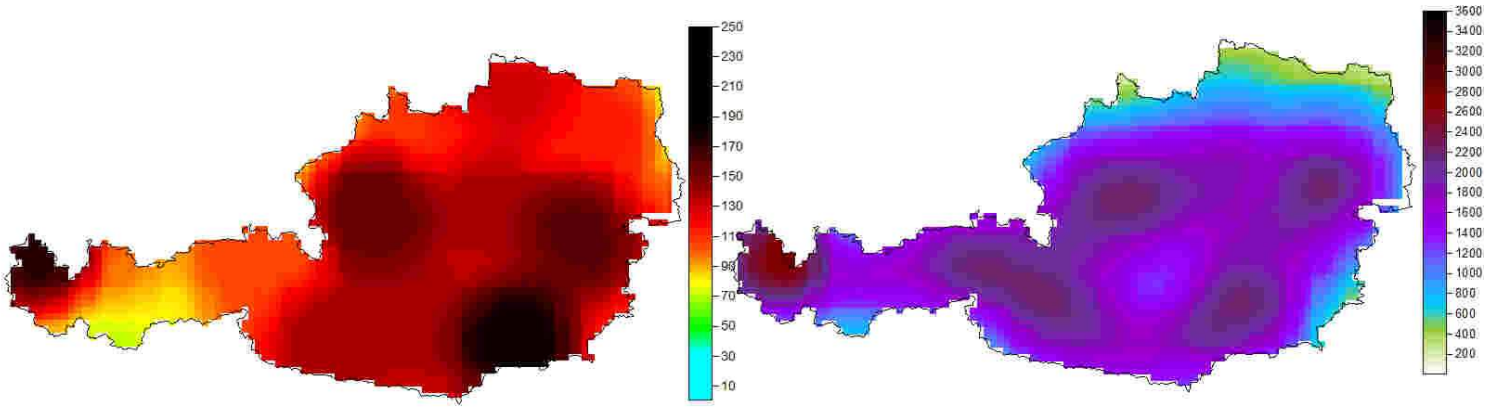
d = 6 hours



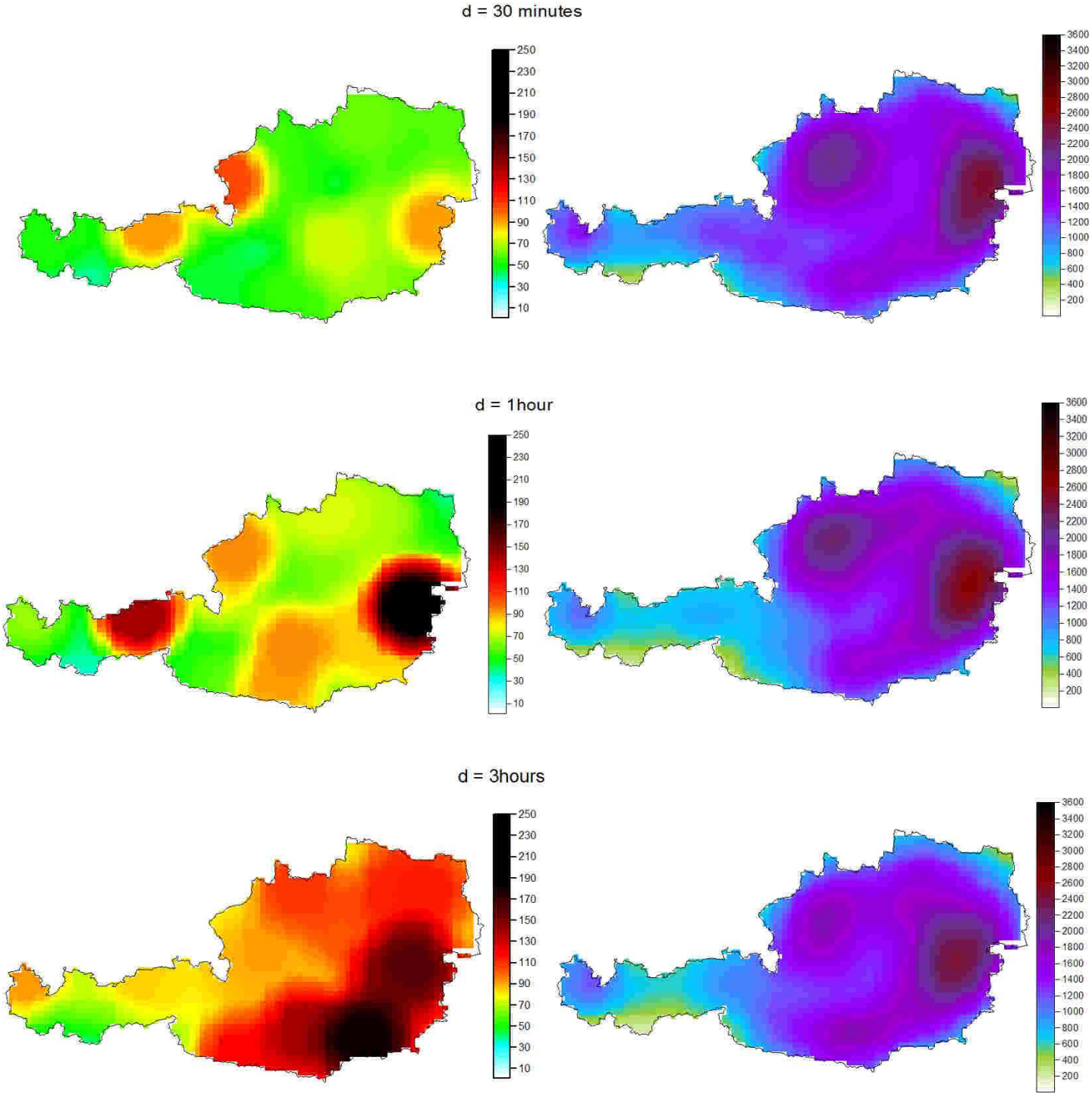
d = 9 hours



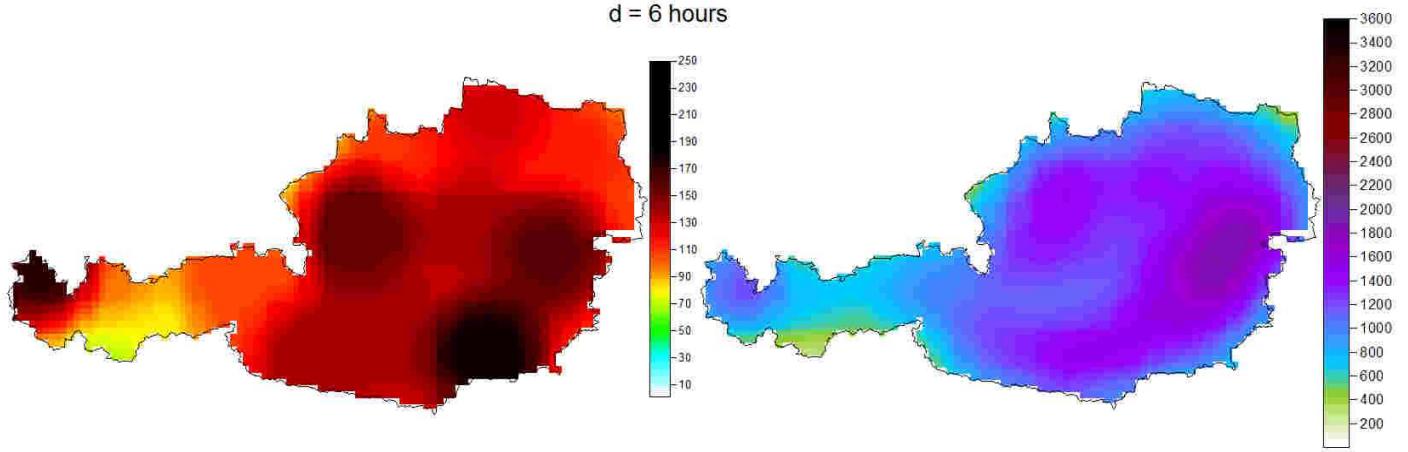
d = 24 hours



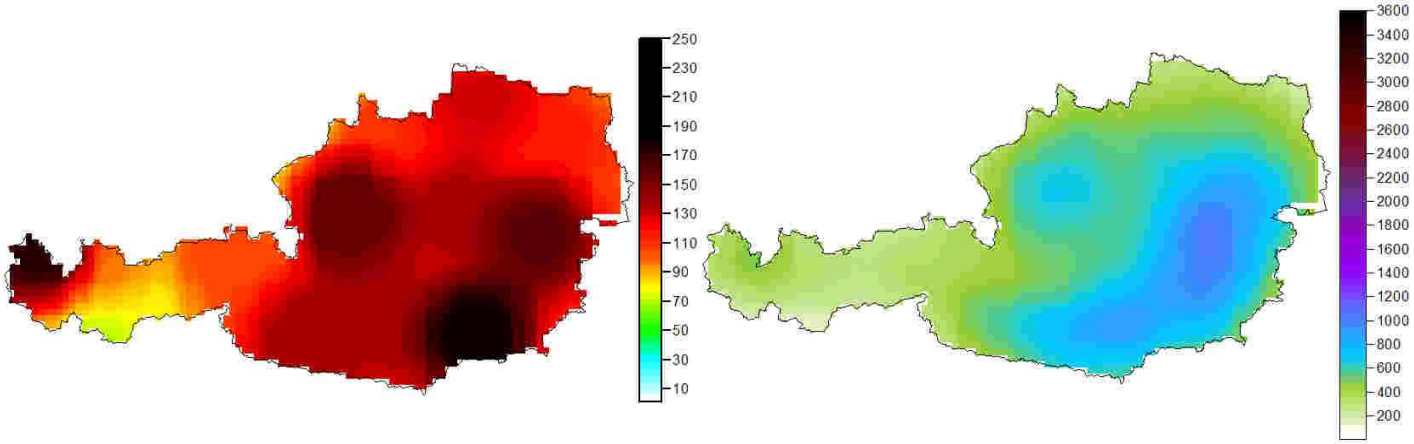
5. Return period in years for the maximum rainfall event in mm, it is now considerate spatial correlation among sites.



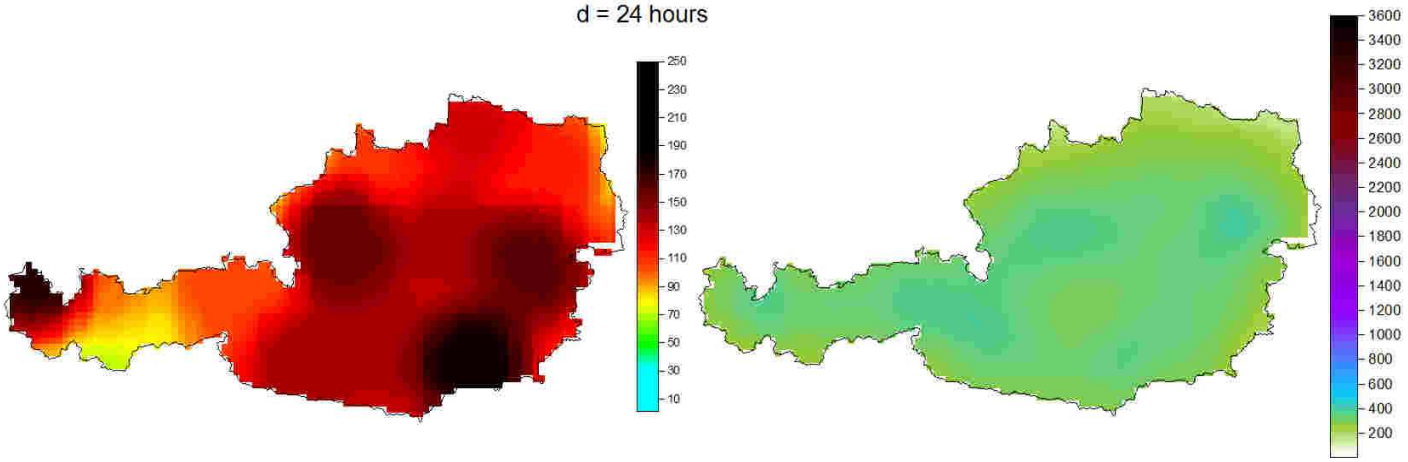
d = 6 hours



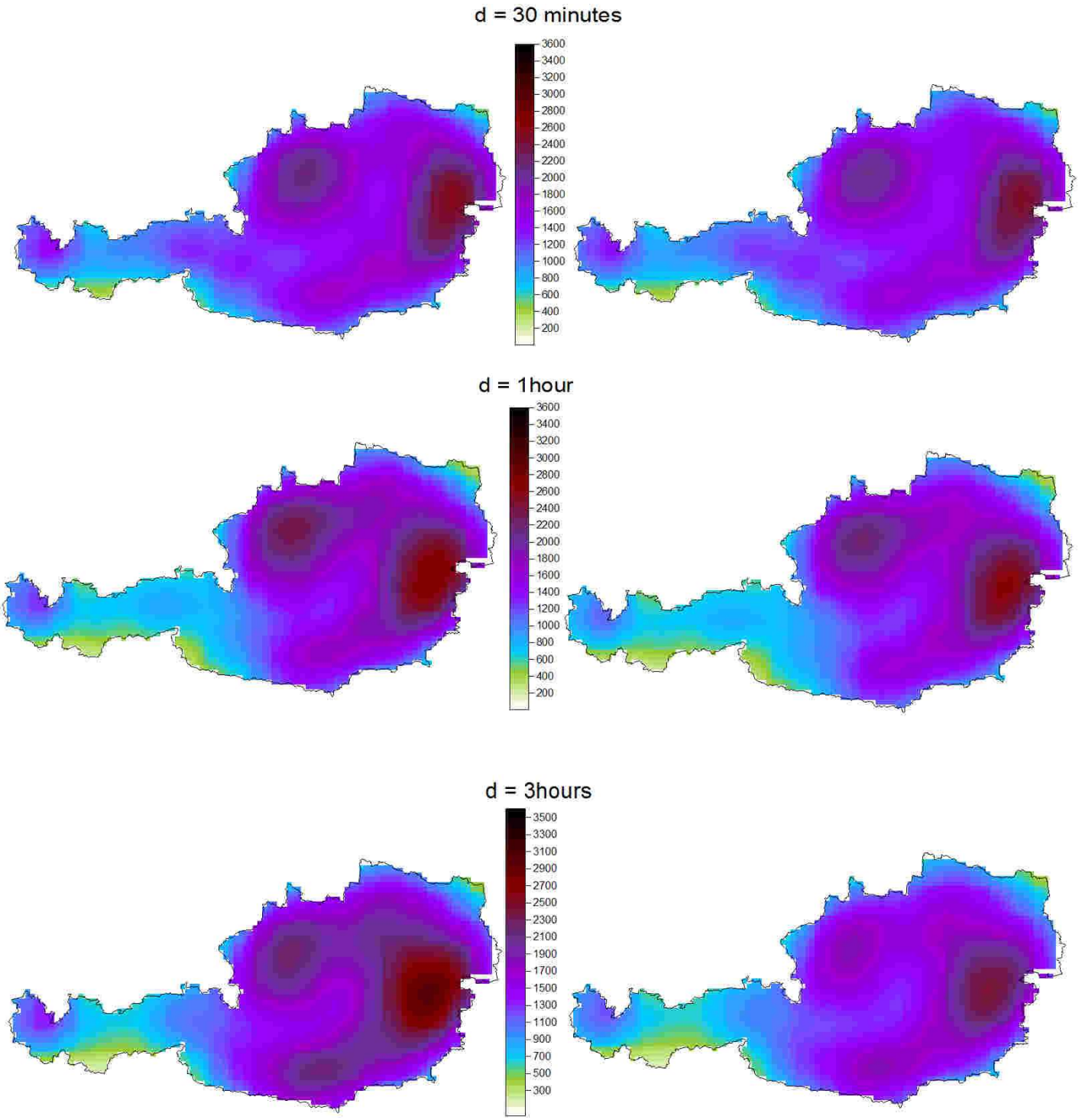
d = 9 hours



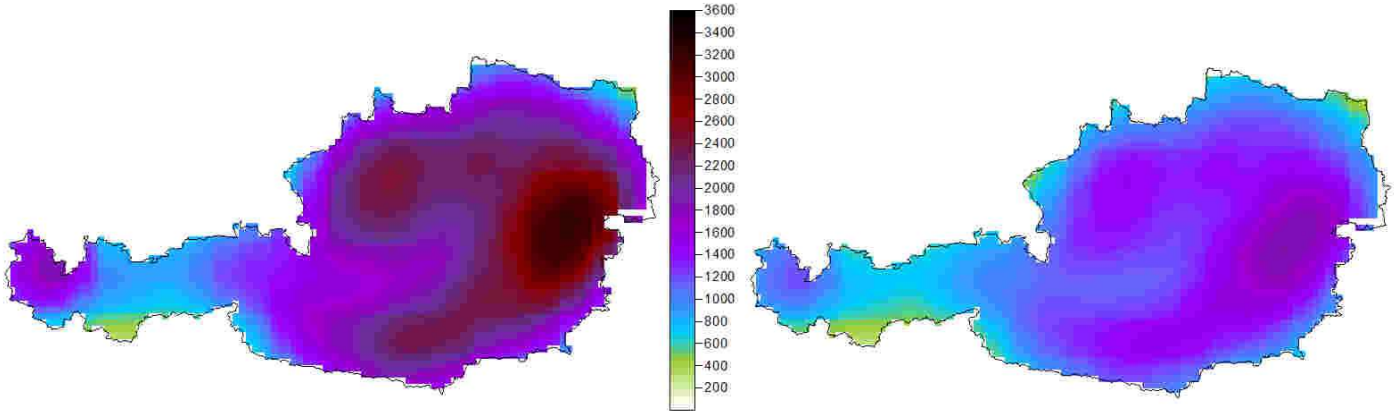
d = 24 hours



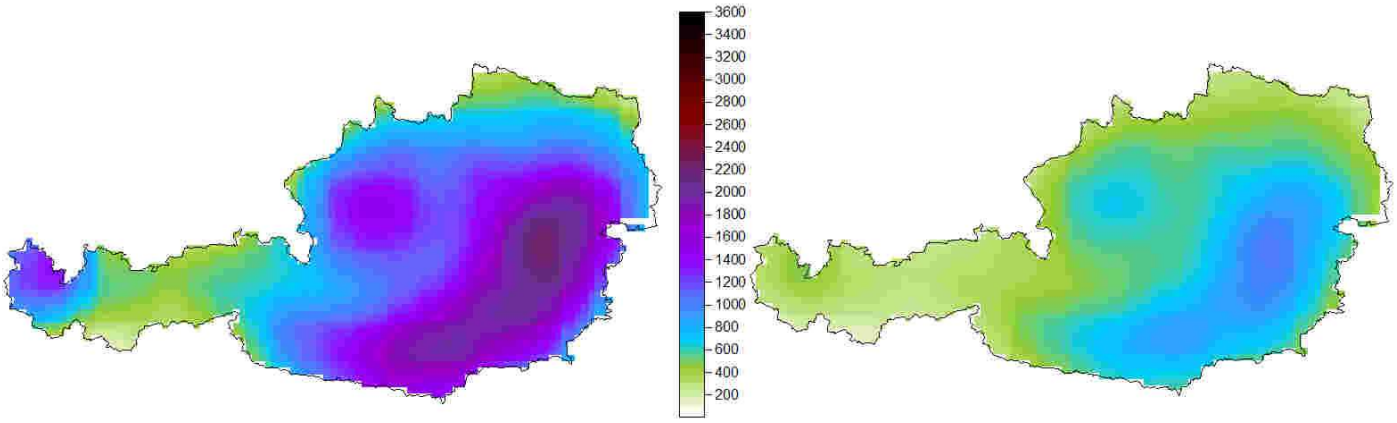
6. Comparison between the return period (in years) for the correlated case and the uncorrelated one.



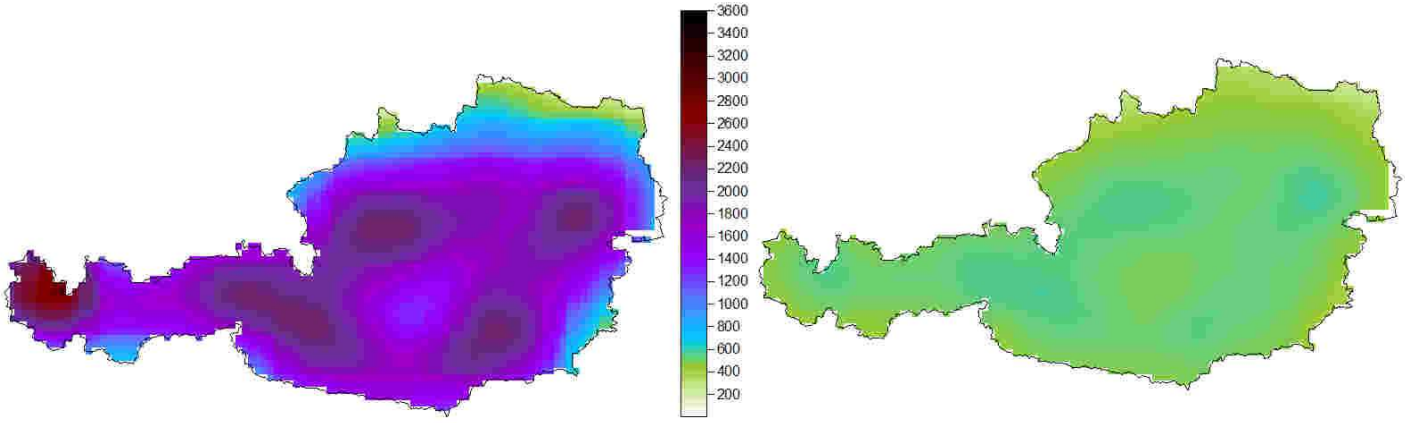
d = 6 hours



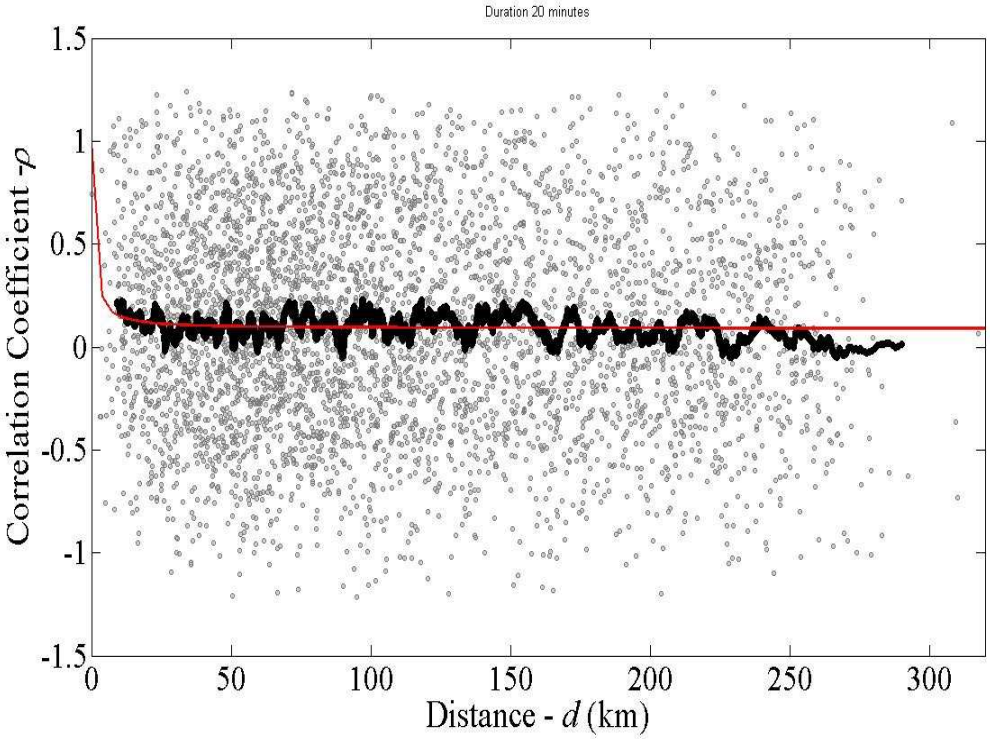
d = 9 hours

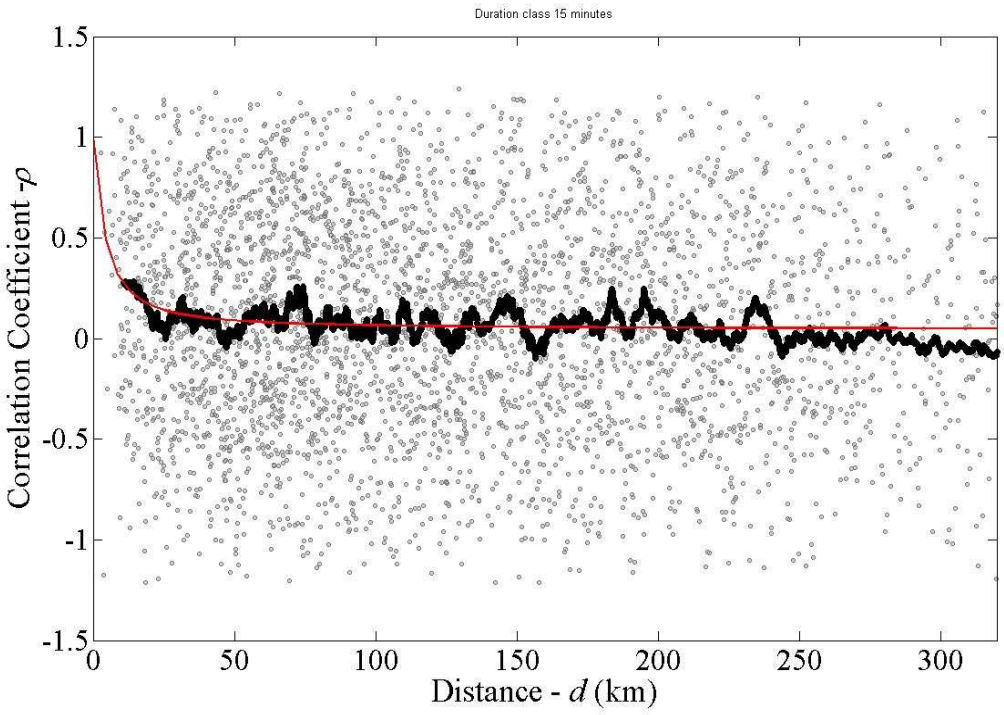


d = 24 hours

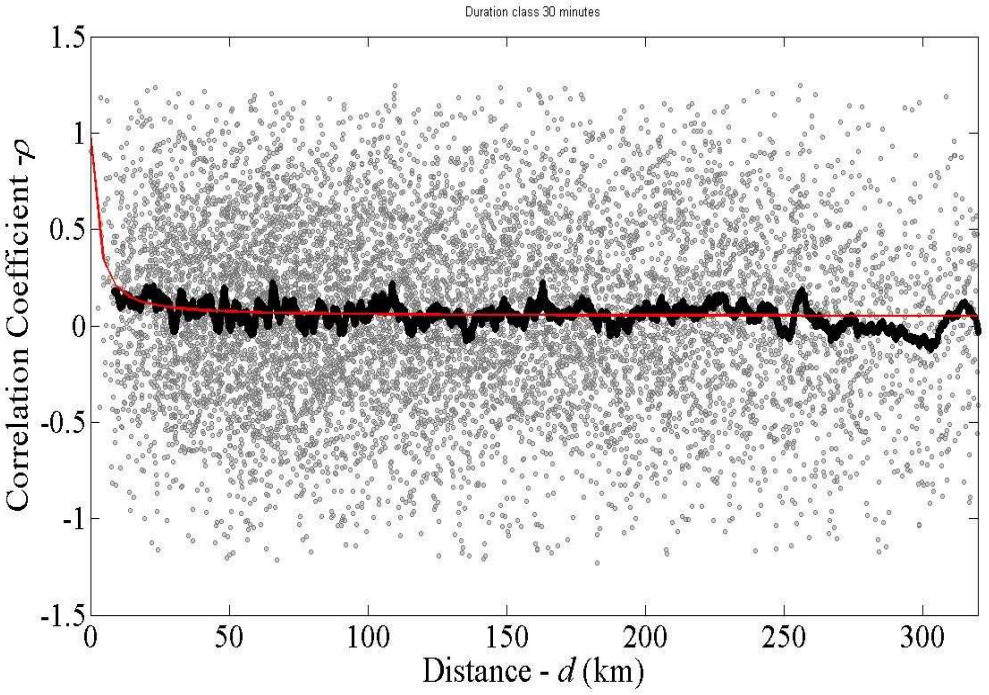


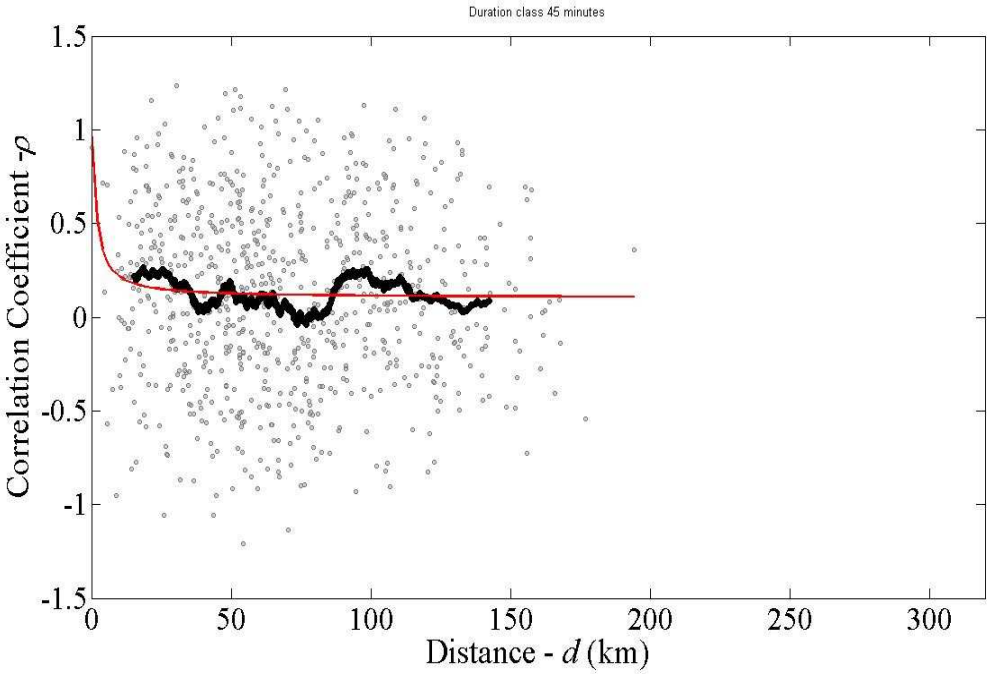
7. Empirical cross-correlation coefficient for couple of stations in the hole study area, for all durations



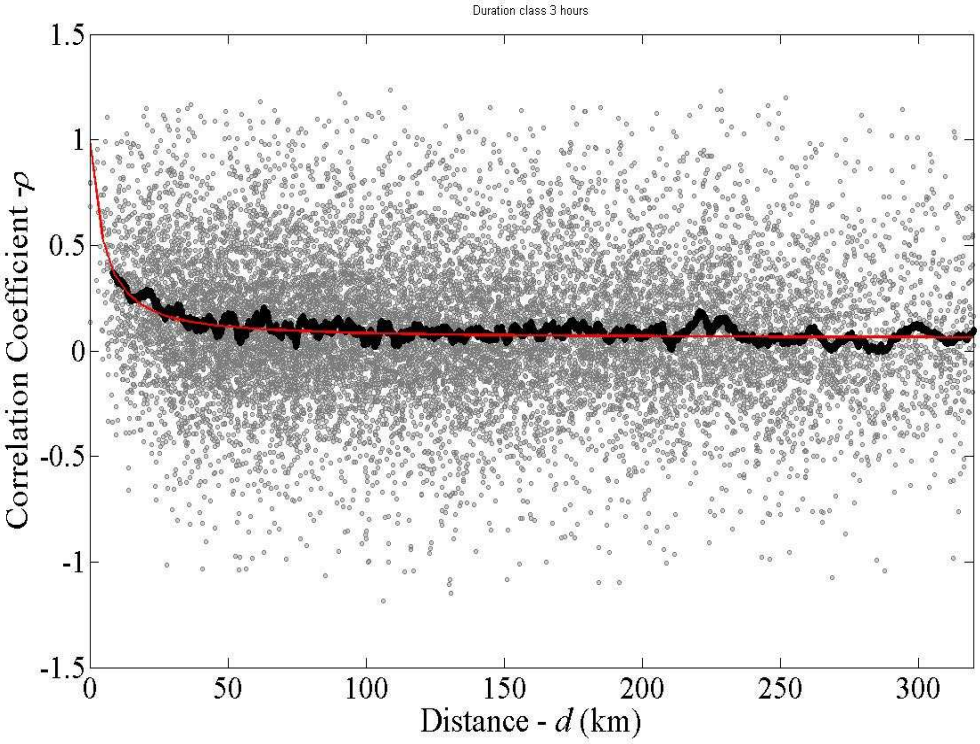


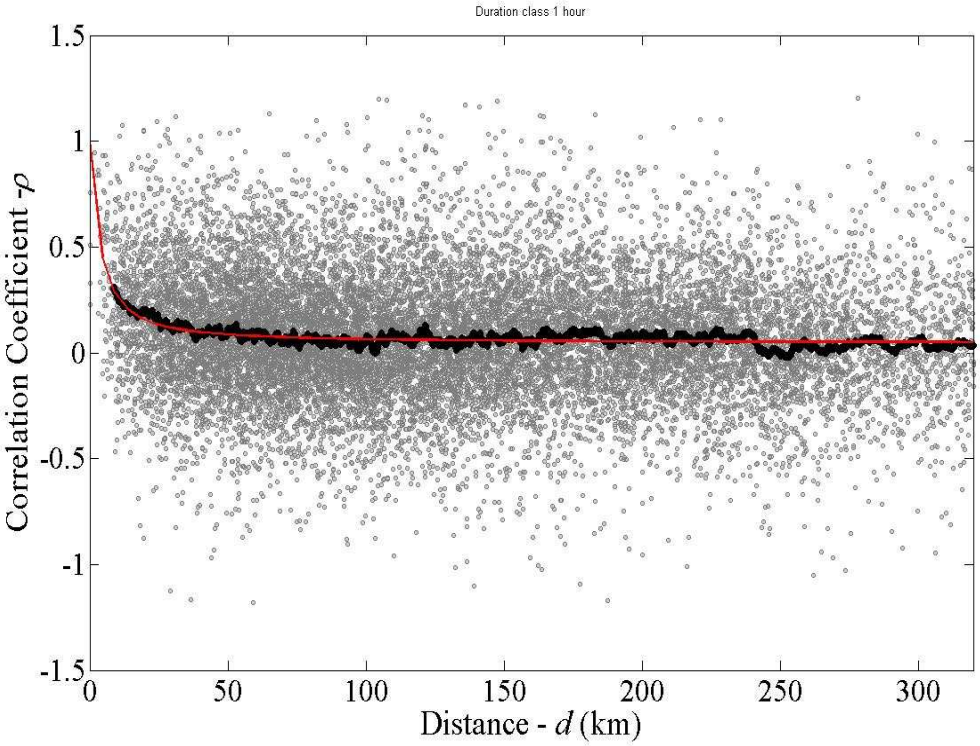
Continue: Empirical cross-correlation coefficient for couple of stations



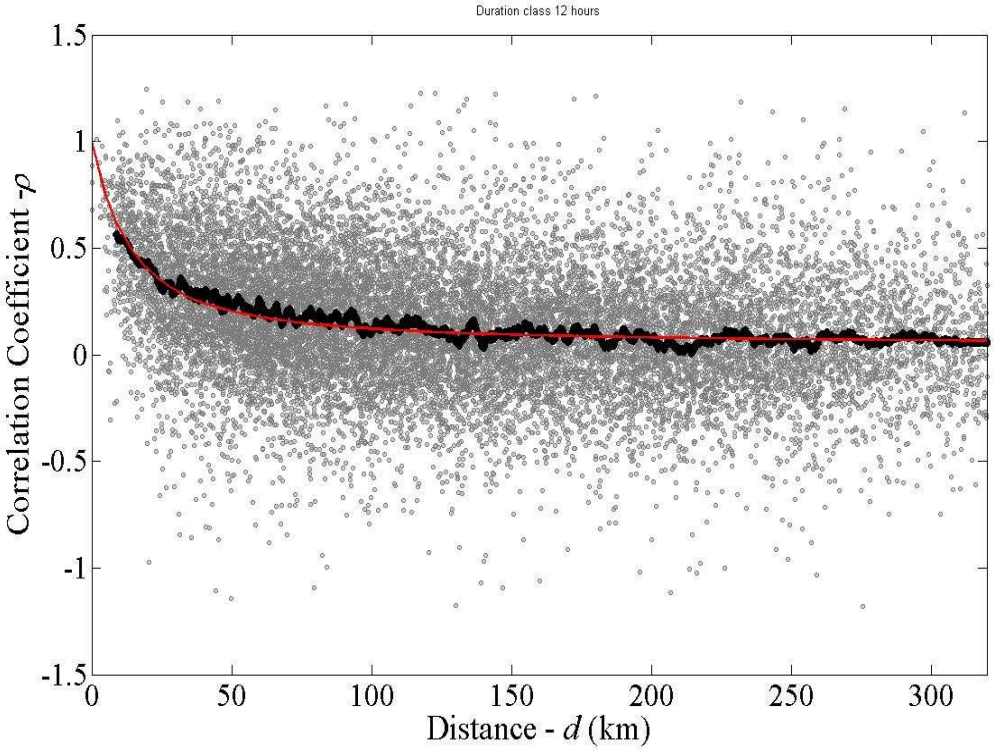


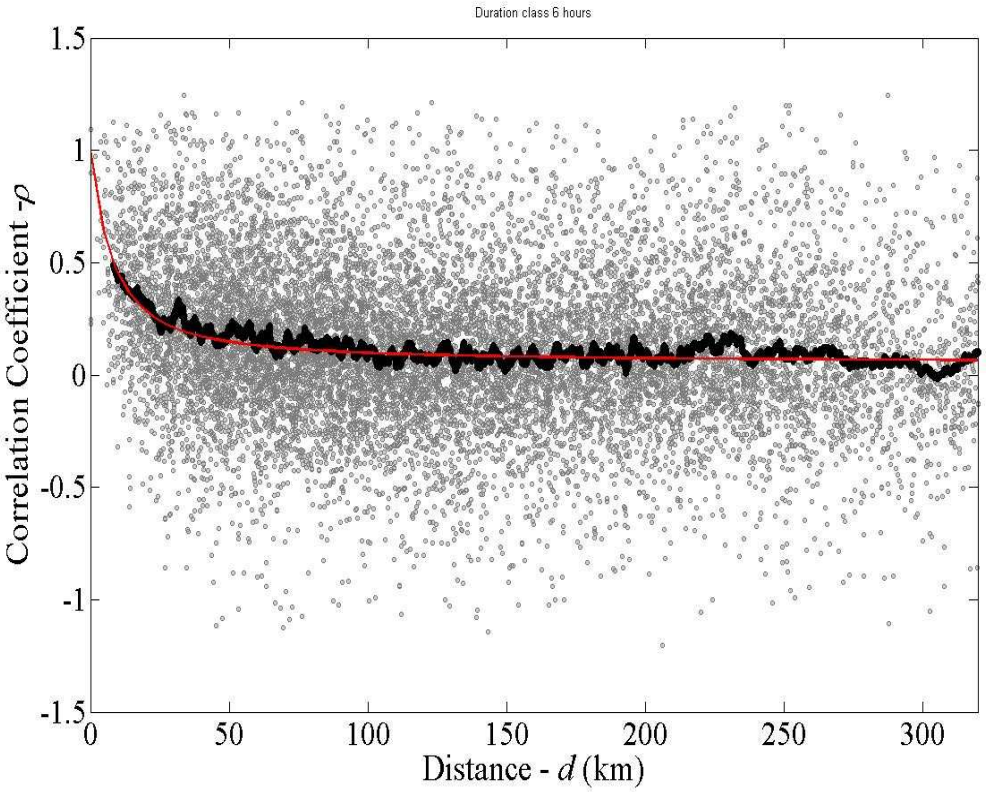
Continue: Empirical cross-correlation coefficient for couple of stations



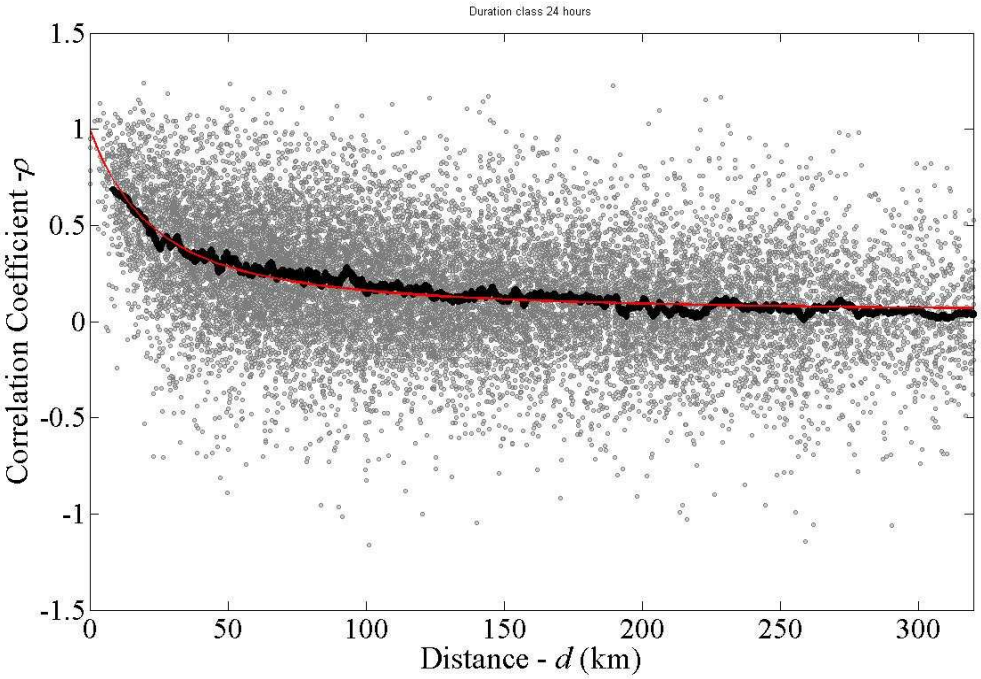


Continue: Empirical cross-correlation coefficient for couple of stations

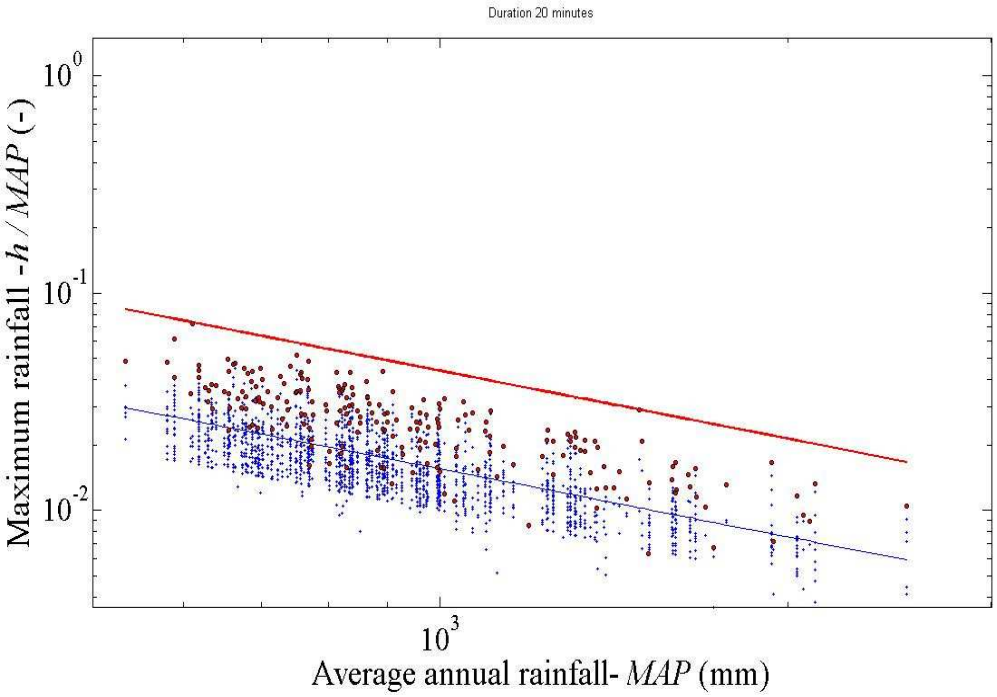


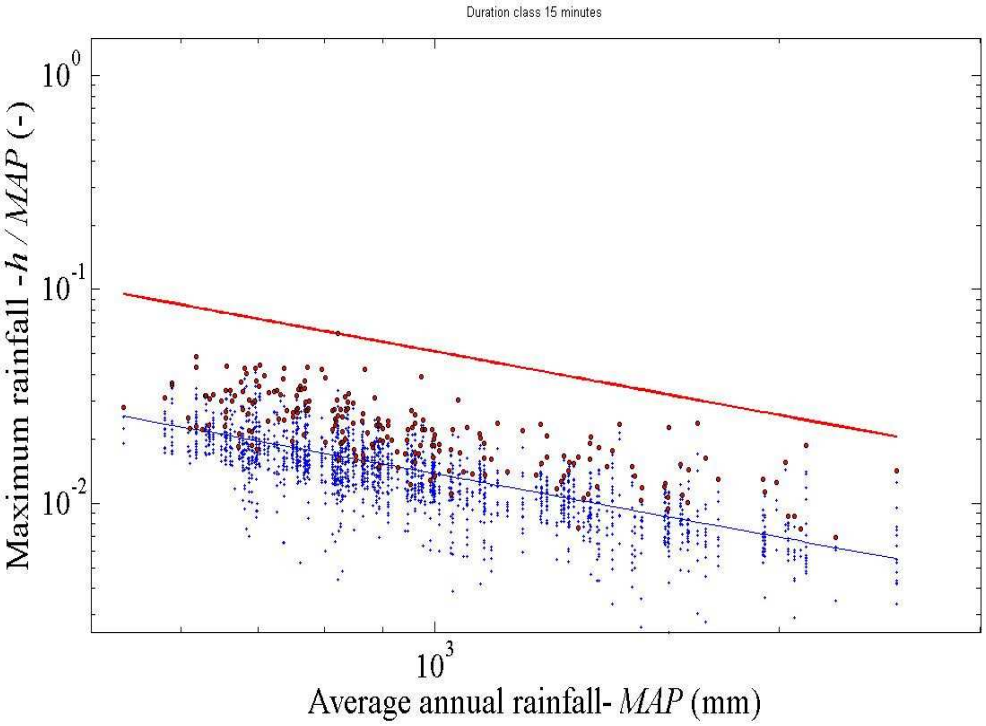


Continue: Empirical cross-correlation coefficient for couple of stations

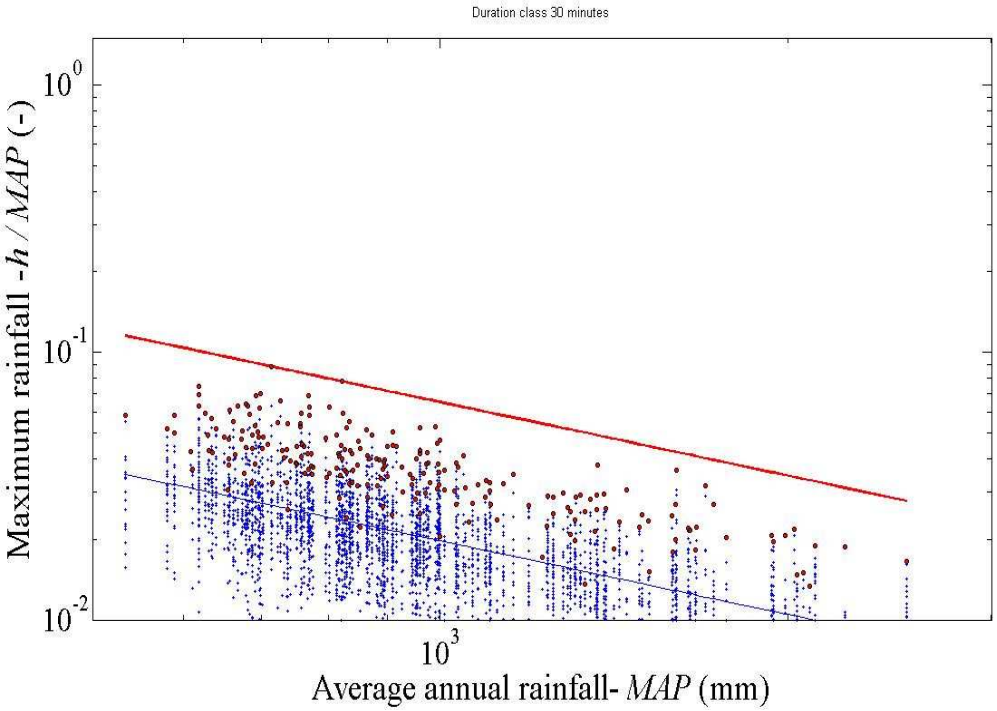


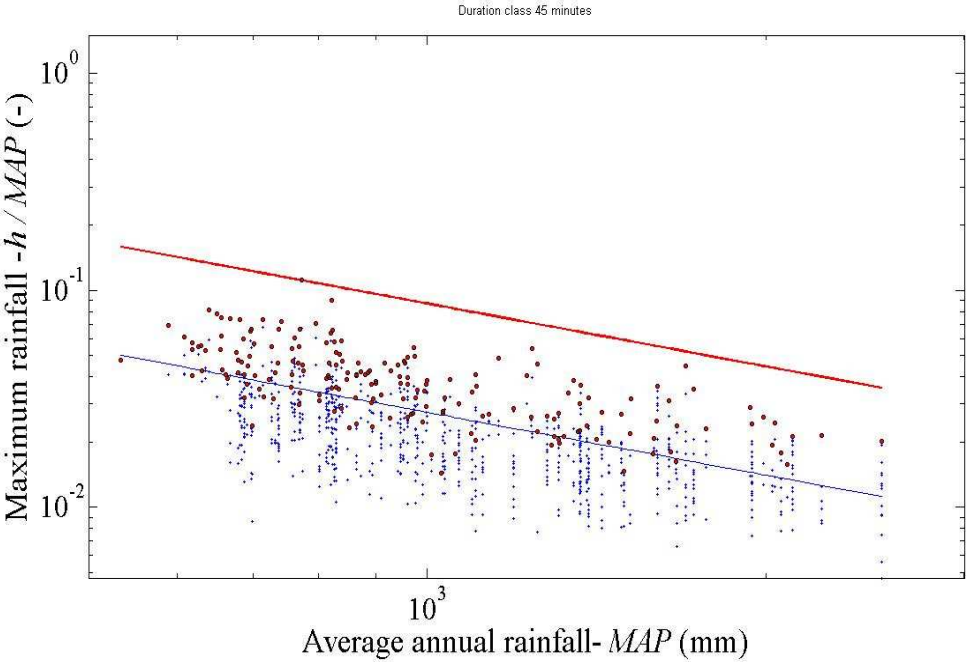
8. Figure 7.2-4 Empirical envelope curves. Depth-Duration Envelope Curve (red line); annual maximum rainfall depths (blue +); rainfall depths of record for different durations (red dots); trend-line for log(duration)-log(average rainfall depth).



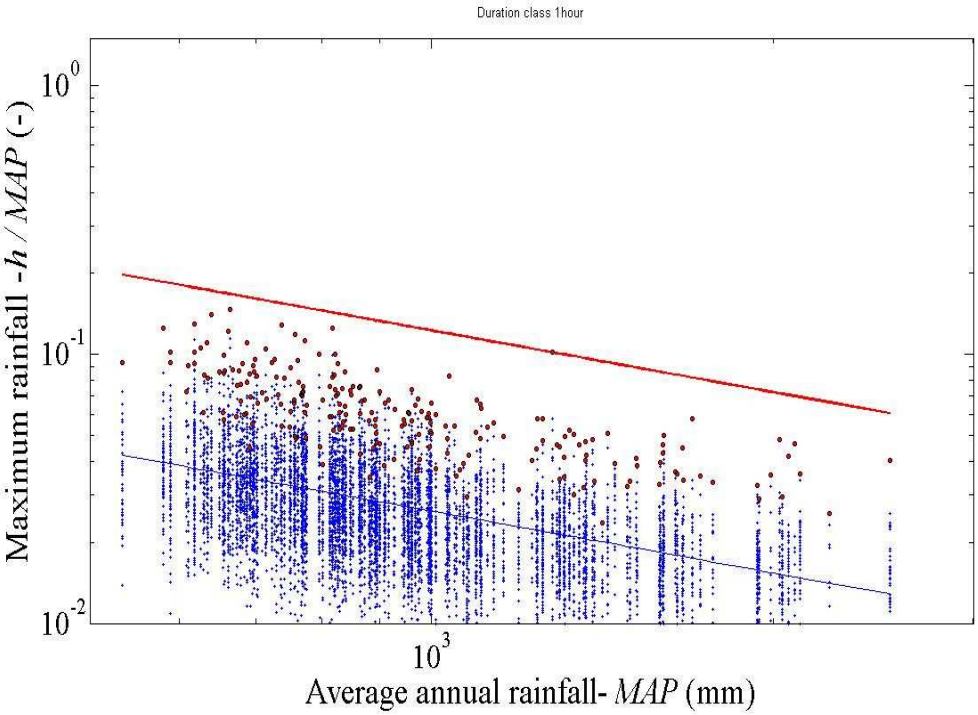


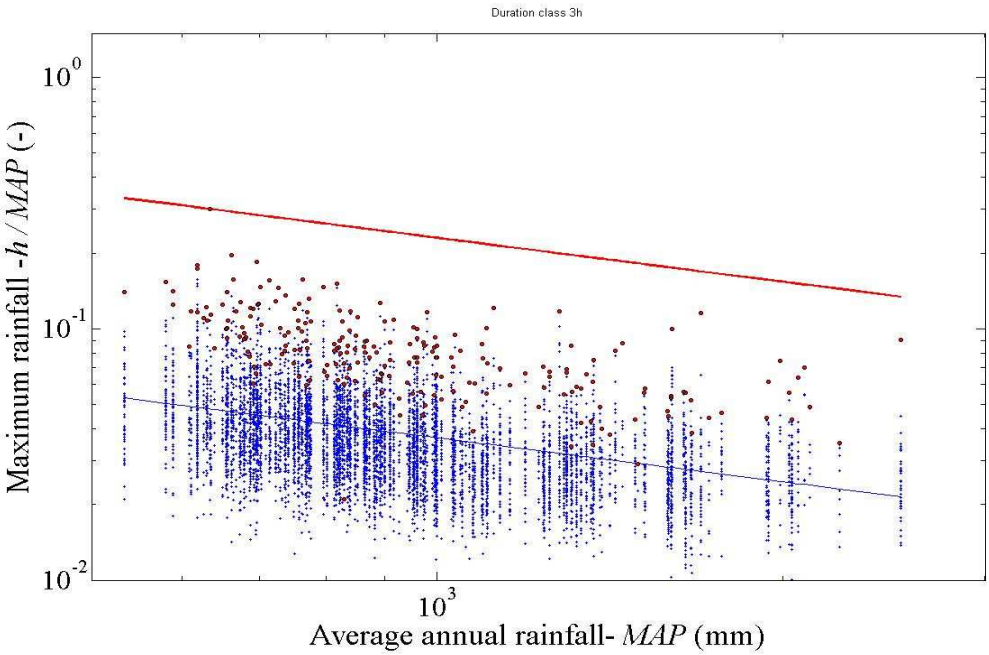
Continue: empirical envelope curves.



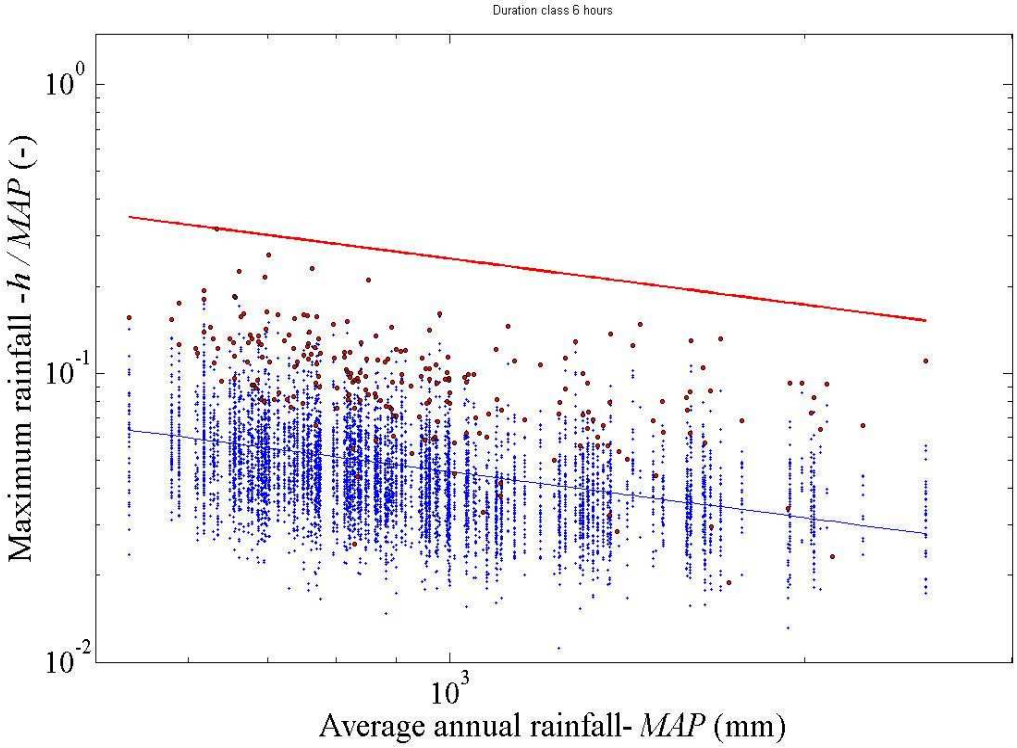


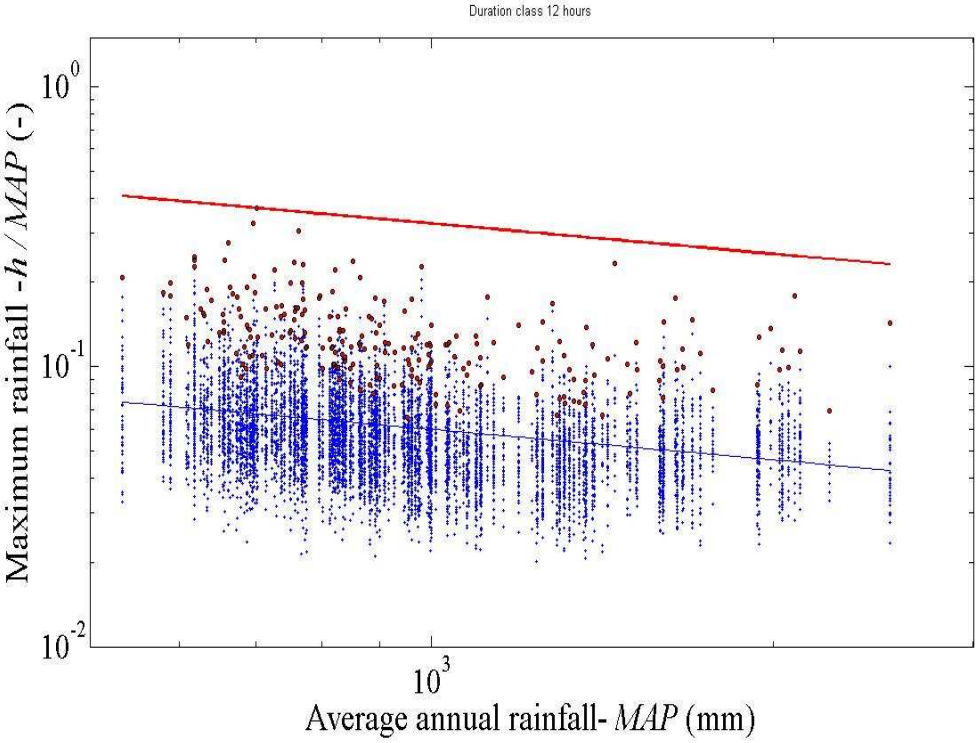
Continue: empirical envelope curves.



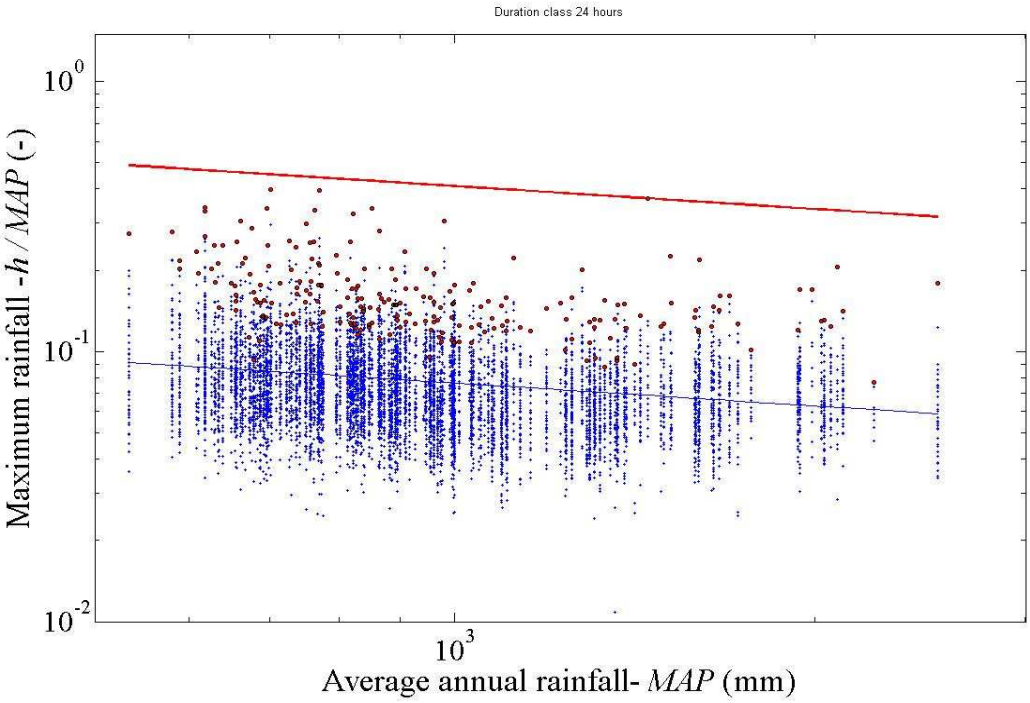


Continue: empirical envelope curves.

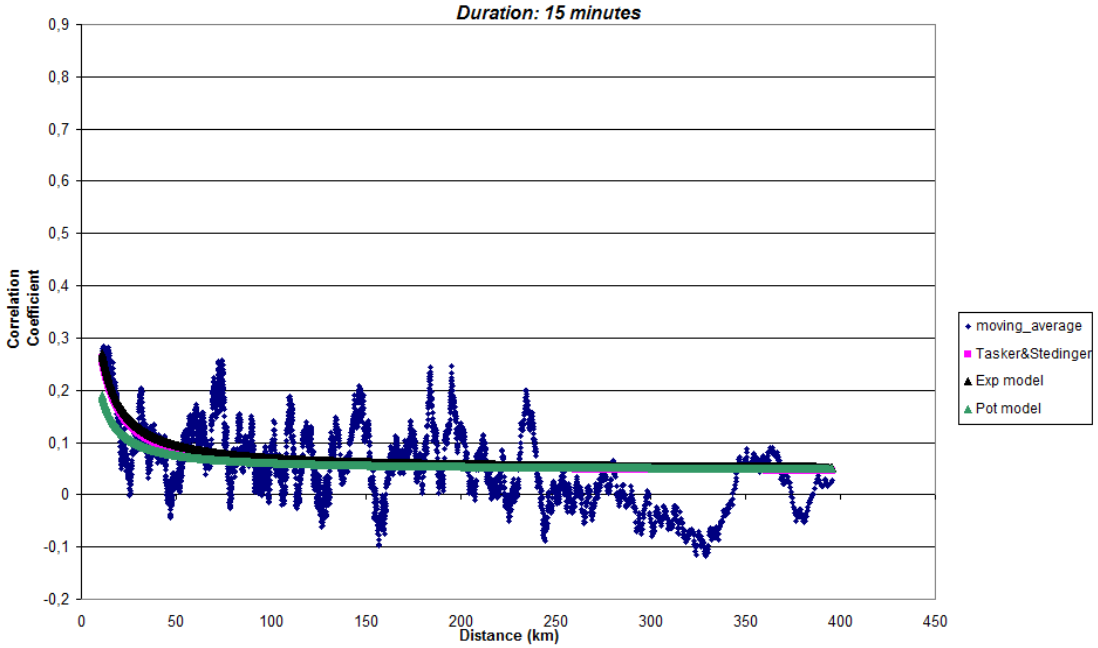


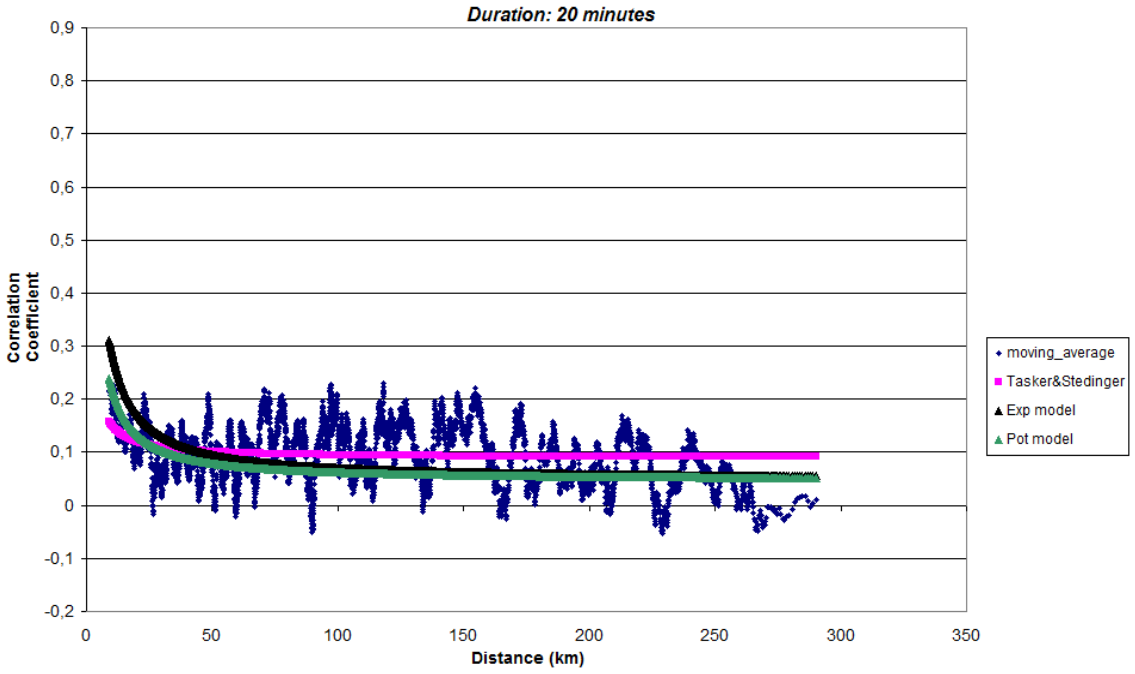


Continue: empirical envelope curves.

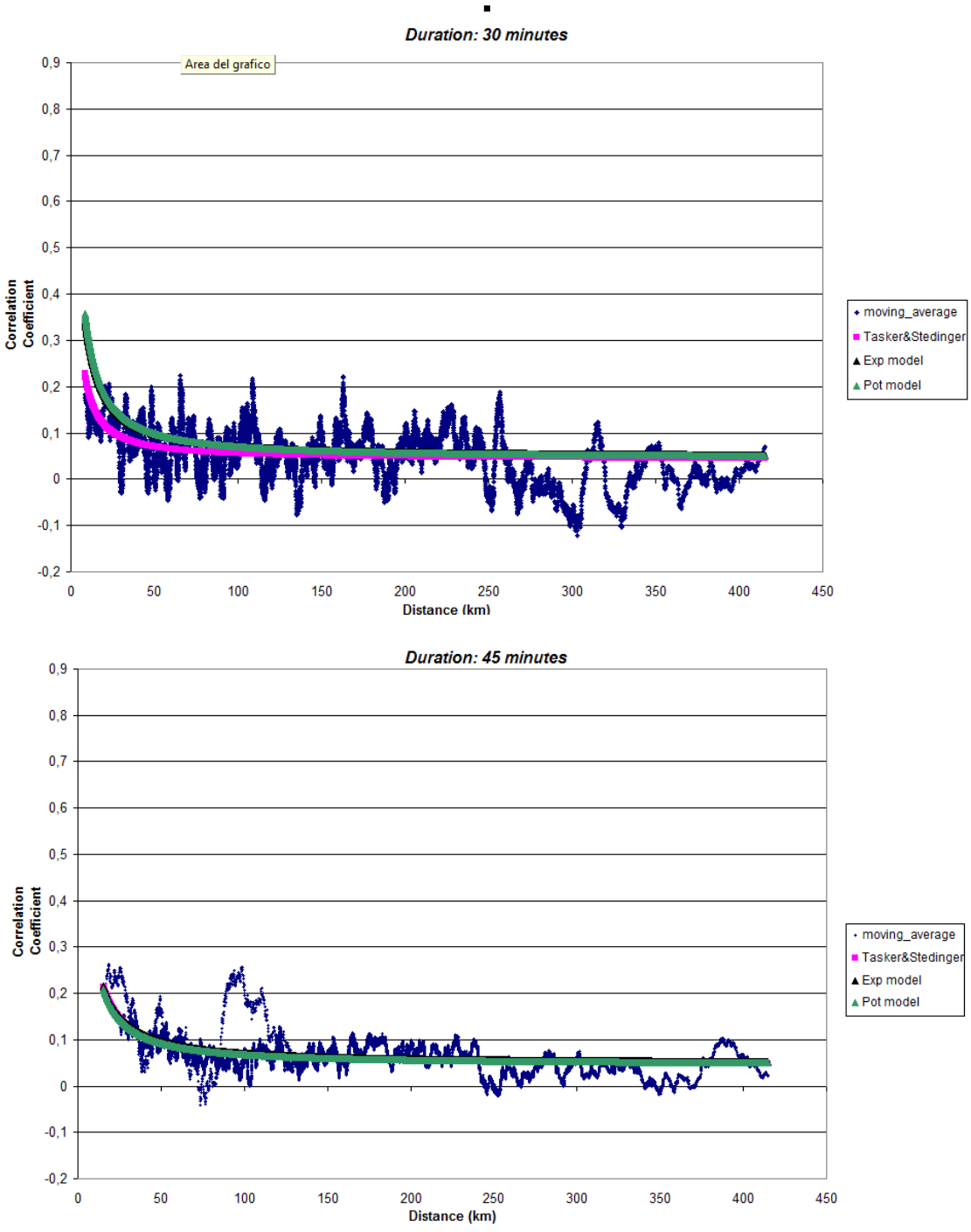


9. Moving weighted average curve for empirical cross-correlation coefficients (blue line), correlation formula (5.3.2-1) calibrated for the whole study area (pink line), correlation exponential formula (black line) and correlation potential formula (green line) calibrated for the whole study area.

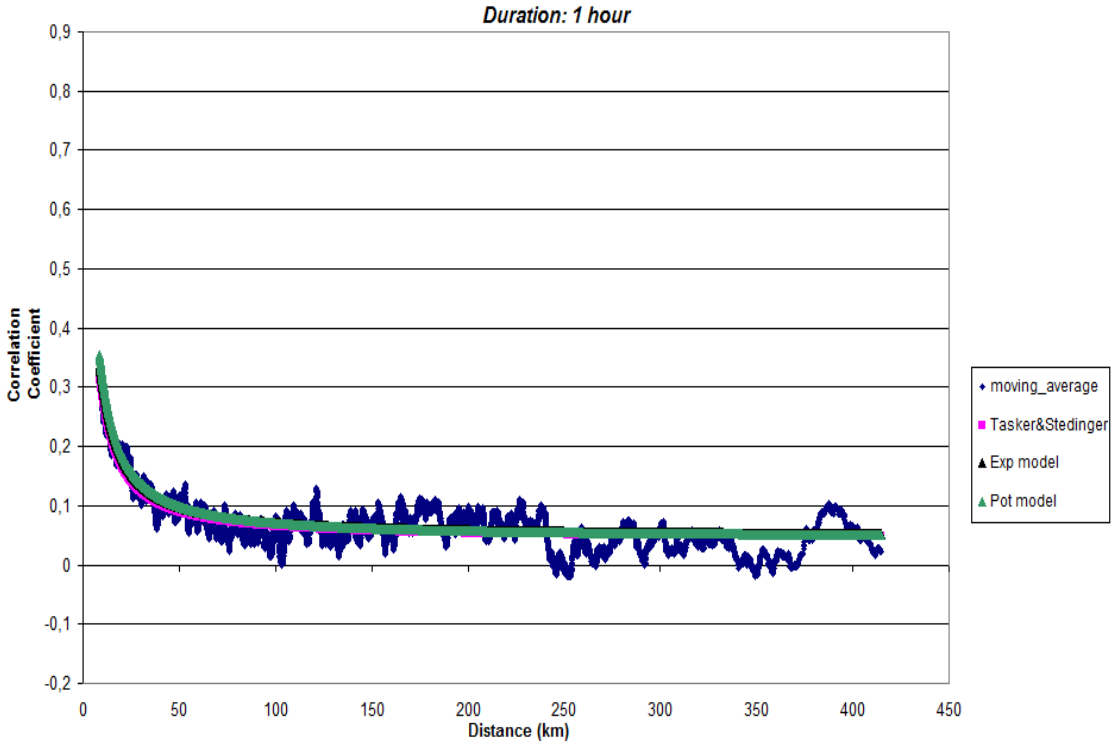


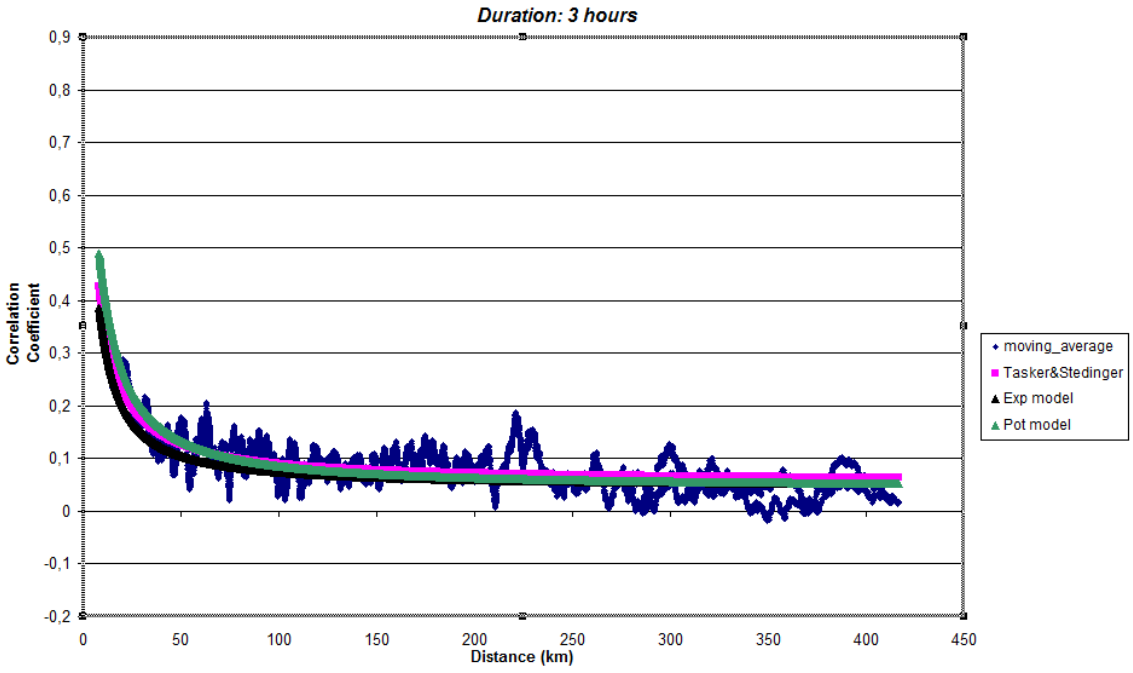


Continue: Moving weighted average curve for empirical cross-correlation coefficients (blue line), correlation formula (5.3.2-1) calibrated for the whole study area (pink line), correlation exponential formula (black line) and correlation potential formula (green line) calibrated for the whole study area

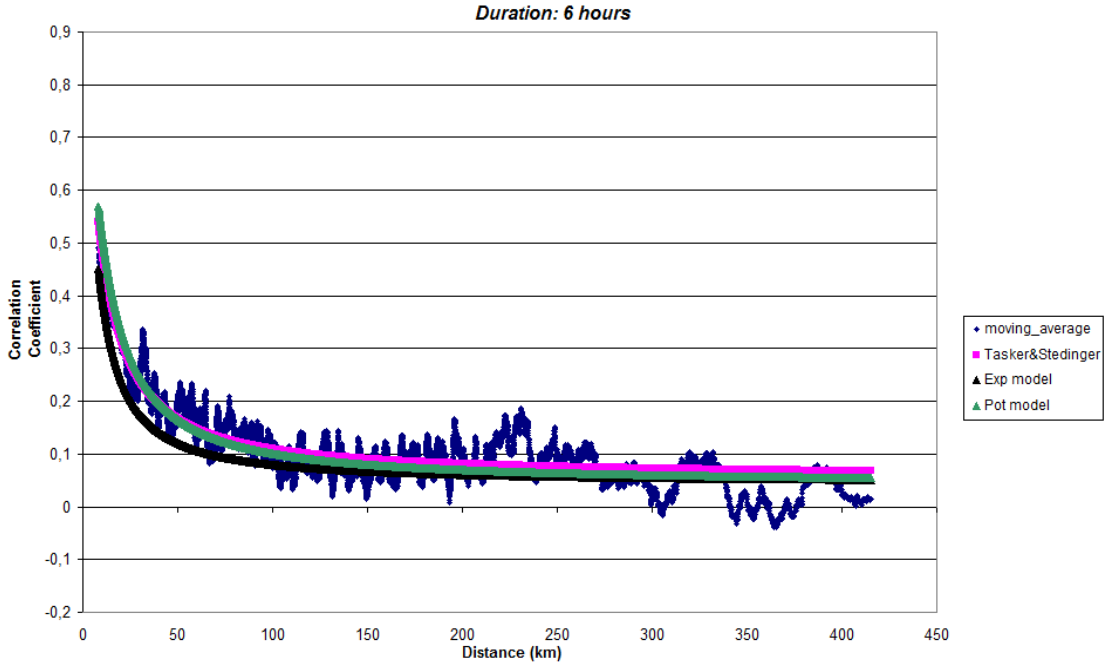


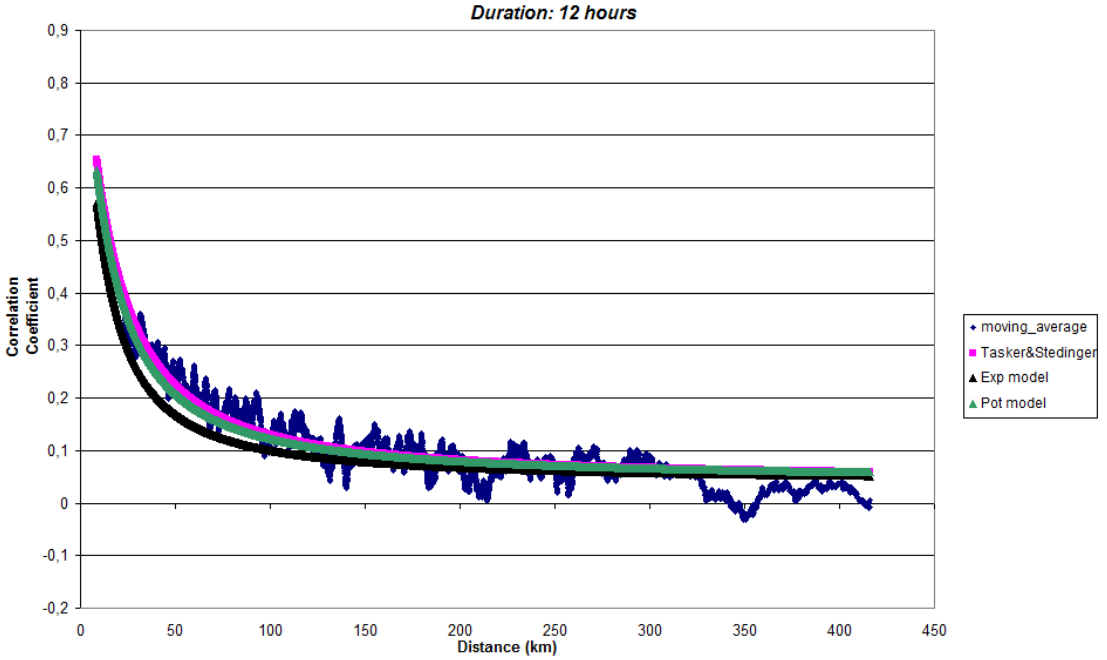
Moving weighted average curve for empirical cross-correlation coefficients (blue line), correlation formula (5.3.2-1) calibrated for the whole study area (pink line), correlation exponential formula (black line) and correlation potential formula (green line) calibrated for the whole study area



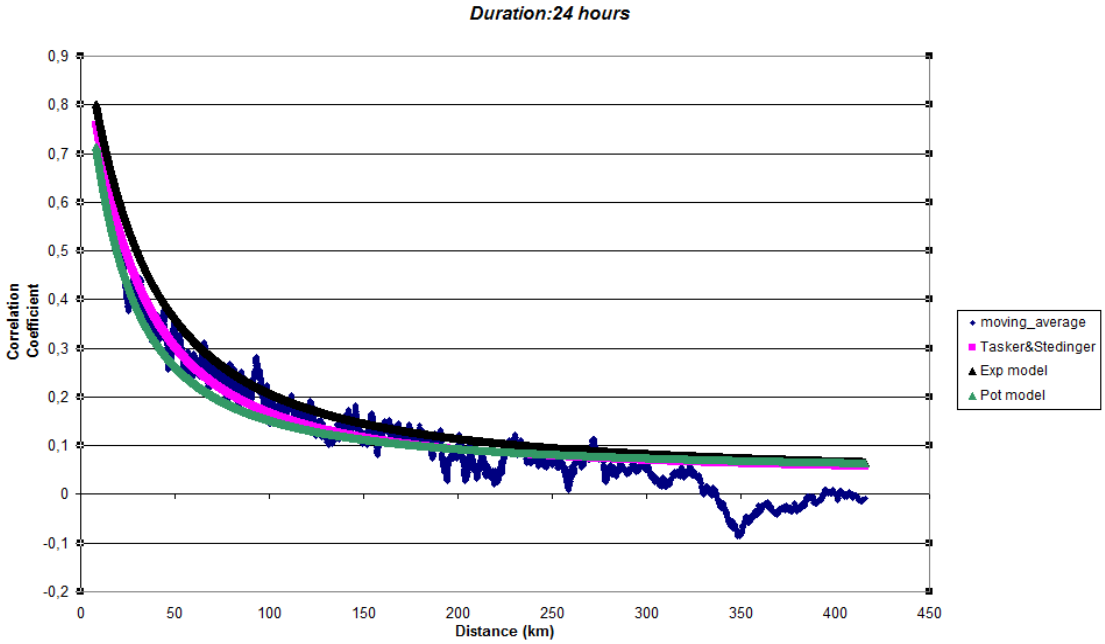


Moving weighted average curve for empirical cross-correlation coefficients (blue line), correlation formula (5.3.2-1) calibrated for the whole study area (pink line), correlation exponential formula (black line) and correlation potential formula (green line) calibrated for the whole study area





Moving weighted average curve for empirical cross-correlation coefficients (blue line), correlation formula (5.3.2-1) calibrated for the whole study area (pink line), correlation exponential formula (black line) and correlation potential formula (green line) calibrated for the whole study area



REFERENCES

1. U. Moisello

   "Idrologia Tecnica", La Goliardicapavese.

2. A. Castellarin, Richard M. Vogel, Nicholas C. Matalas,

   "Probabilistic behaviour of a regional envelope curve", Water Resources Research, Vol. 41, W06018, doi:10.1029/2004WR003042, 2005.

3. A. Castellarin,

   "Probabilistic envelop curve for design flood estimation at ungauged sites", *Water Resources Research*, Vol. 43, W04406, doi:10.1029/2005WR003484, 2007.

4. , Richard M. Vogel, Nicholas C. Matalas, John F.England, Attilio. Castellarin,

   "An assessment of exceedance probability of envelope curves", *Water Resources Research*, Vol. 43, W07403, doi:10.1029/2006WR005586, 2007.

5. Ven Te Chow, Professor of Hydraulic Engineering University of Illinois

   "Handbook of applied hydrology" *a Compendium of Water-resources Technology*, McGraw-Hill Book Company, 1971.

6. David R. Maidment, Professor of Civil Engineering University of Texas at Austin

   "Handbook of Hydrology", McGraw-Hill Book Company, 1993.

7. Edward H. Isaaks, R. Mohan Srivastava,

   "An introduction to Applied Geostatistics", Oxford university Press, 1989.

8. P. Claps, F. Laio

"Can continuous streamflow data support flood frequency analysis? An alternative to the partial duration series approach", *Water Resources Research*, Vol. 39, NO.8, 1216, doi:10.1029/2002WR001868, 2003.

9. J. R. M. Hosking, J. R. Wallis

   "The effect of intersite dependence on regional flood frequency analysis", *Water Resources Research*, Vol. 24, NO.4, pages 588-600, April 1988.

10. A. Brath, A. Castellarin, A. Montanari

    "Assessing the reliability of regional depth-duration-frequency equations for gaged and ungaged sites", *Water Resources Research* Vol. 39, NO. 12, 1367, doi:10.1029/2003WR002399, 2003.

11. G. Di Baldassarre, A. Castellarin, and A. Brath

    "Relationships between statistics of rainfall extremes and mean annual precipitation: an application for design-storm estimation in northern central Italy", Hydrology and Earth System Science, 10, 589–601, 2006.

12. R. Merz, G. Blöschl

    "A process typology of regional floods" *Water Resources Research*, Vol.39, NO.12, 1034, doi:10.1029/2002WR001952,2003.

13. P. Kahlig

    "On deterministic criteria for heavy rainfall at a poin" Theoretical and Applied Climatology, Vol. 46, N.4, p. 203-208, 1993.

14. Davide Grosso

    "Tecniche di valutazione della precipitazione massima probabile (PMP)", 2007

Acknowledgments

First of all a sincere acknowledgment to Professor Blöschl and Ing. Merz for the valuable help and the time spent over this work. A thanksgiving to Professor Brath for giving me the possibility of making this exchange experience, and a particular thanksgiving to Ing. Castellarin for having made my interest on hydrology grow, for the constant concern demonstrated and for the fundamental help, from a both personal and professional point of view.

Torno alla lingua madre per ringraziare i miei genitori e la mia famiglia. Grazie, per essere arrivati fin qui con me dal primo all'ultimo esame, da una città all'altra.

Un ringraziamento che non trova parole per Alessandro, per il suo amore paziente e gratuito, che mi ha sorretta in questi anni. Ai miei amici tutti,in particolare a Silvia un sincero ringraziamento.

A Silvana e Vanni, che oggi mi danno il supporto informatico per stampare questa Tesi.

A tutti voi…

Grazie,

Lorenza.