EXTREME VALUE ANALYSIS OF RAINFALL MAXIMA UNDER CLIMATE CHANGE, FOR FLOOD MANAGEMENT IN CYPRUS

BY

KOULLA VRACHNOU

RESEARCH PROJECT

Submitted in partial fulfillment of the requirements
for the degree of Environmental Sciences
at The Cyprus Institute
Nicosia, Cyprus, 2019

Examination Committee:

Associate Professor Adriana Bruggeman (Supervisor)
Associate Professor Panos Hadjinicolaou
Associate Professor George Biskos
Associate Research Scientist George Zittis

Abstract

Throughout the years, around the world people suffer from extreme events such as floods. Loss of lives, properties, diseases, and so much more negative impacts are affecting human lives. Climate change is responsible for more extreme rainfall events and contributes to the increasing trend of floods in the future and thus as a result more people will be affected. The overall aim of this project is to analyze climate data to study changes in rainfall extremes in Cyprus. First, we do the evaluation of modeled annual daily rainfall extremes using observed rainfall data and then we use modeled data for several time periods in order to see changes in extremes in the future. We use thirteen 12-km grid cells for Limassol, from a regional Weather Research and Forecasting model simulation for the RCP8.5 scenario, and we apply extreme value analysis using Hydrognomon. The Generalized Extreme Value distribution (GEV-Max) is used to fit our data and L-moments with and without specified shape parameter and maximum likelihood methods are studied.

From the analysis, we found that the modeled rainfall extremes of 1981-2010 are underestimating the observed rainfall extremes. Comparing the data of 1981-2010 with the data of 2021-2050 we concluded that we get more extremes in the near future. The fitted distributions for 1981-2010 versus 2021-2100 showed that extreme values for return periods of 20 and 100 years are increasing for 11 grid cells out of 13.

Acknowledgments

I would like to thank my supervisor Adriana Bruggeman for the useful comments and the help that she provided me through all the way of the thesis. I would also like to thank her for the precious time that she spent with me and all the guidelines that she gave me. My accomplishment would not be possible without her.

In addition, I would like to thank George Zittis for the provision of the data set for the thesis. The data were prepared in the framework of the Interreg V-A Greece-Cyprus 2014-2020 project ERMIS-F.

I would like to express my huge gratitude to my whole family and my loved ones, for the support and the encouragement through my studies and through my effort of writing this thesis. Mum, dad and siblings I am grateful for your love.

Last, I want to thank Panos Hadjinicolaou, George Biskos and George Zittis, the internal examiners, for their time and review of the thesis and of the presentation of my thesis.

Table of Contents

List of tables

List of figures

1. Introduction

1.1. Problem Statement

In recent years, climate change has become one of the main major issues that humanity has to deal with. The alteration in the surface temperature, precipitation, wind speed, chemical composition of the atmosphere and other climate parameters, has led to unexpected, hazardous changes in the magnitude and frequency of climate extremes, such as floods (Kharin and Zwiers, 2000). Figure 1 shows the annual reported number of natural disasters, separated in categories by type. Throughout the years, floods are the most common natural disaster in many countries around the world and they have become an increasingly significant issue in the Mediterranean region (Llasat et al. 2010).

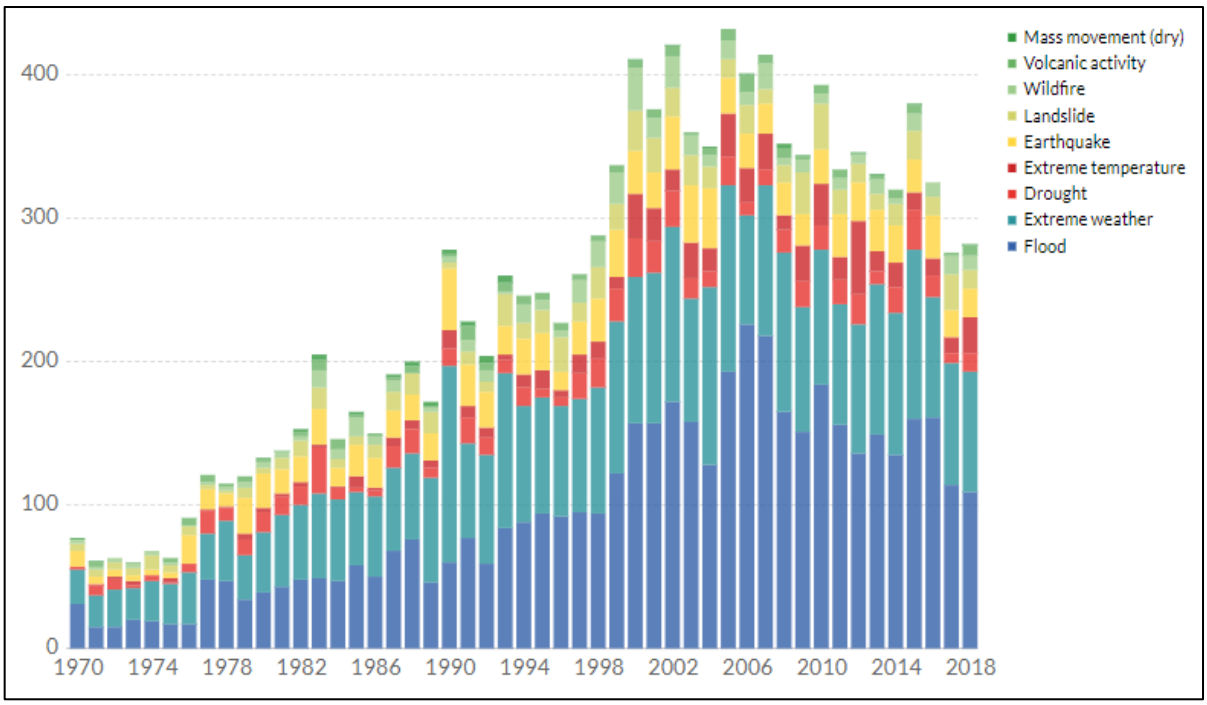

Figure 1: Annual reported number of natural disasters (source: https://www.emdat.be/).

Every year, floods are causing huge damage around the world, and deaths are the most important impact of them. In figure 2 and table 1 flood events and the deaths that they caused are presented. The data for both table and figure are derived from the OFDA/CRED International Disaster Database maintained by the Centre for Research on the Epidemiology of Disasters in Brussels, CRED, in cooperation with United States Office for Foreign Disaster Assistance, OFDA (Jonkman, 2005). In the 20th century, floods caused death to about 100,000 persons and put in danger over 1.4 billion people. Of course, the impacts of a flood depend on the characteristics of the flood and the characteristics of the affected area. As a result, floods have socioeconomical impacts and human health effects, damage eco-systems, destroy historical and cultural values, cause death and affect agricultural production. Table 1 shows the five freshwater flood events with the most people killed in the recent past. Over the period January 1975–June 2002, 1883 freshwater flood events recorded in the OFDA/CRED database are reported to have killed 176,864 people and affected 2.27 billion (Jonkman, 2005).

Table 1: Overview of the five freshwater flood events with most people killed (Jonkman, 2005).

Country	Year	Month	Day	Killed	Total affected	Description
Venezuela	1999	12	19	30.000	483.635	Flash and river floods and landslides around Caracas and other areas.
Afghanistan	1988	6	-	6345	166.831	Floods in Badakhshan, Baghlan, Heart, Kabul, Jouzjan,Samangan, Takhar provinces.
China. P. Rep.	1980	6	-	6200	67.000	Floods in Sichuan, Anhui, Hubei.
India	1978	7	-	3800	32.000.000	Floods in north and northeast India.
China. P. Rep.	1998	8	6	3656	238.973.000	River floods combined with storms and landslides in Hubei, Hunan, Sichuan, Jiangxi, Fujian, Guanxi Prov.

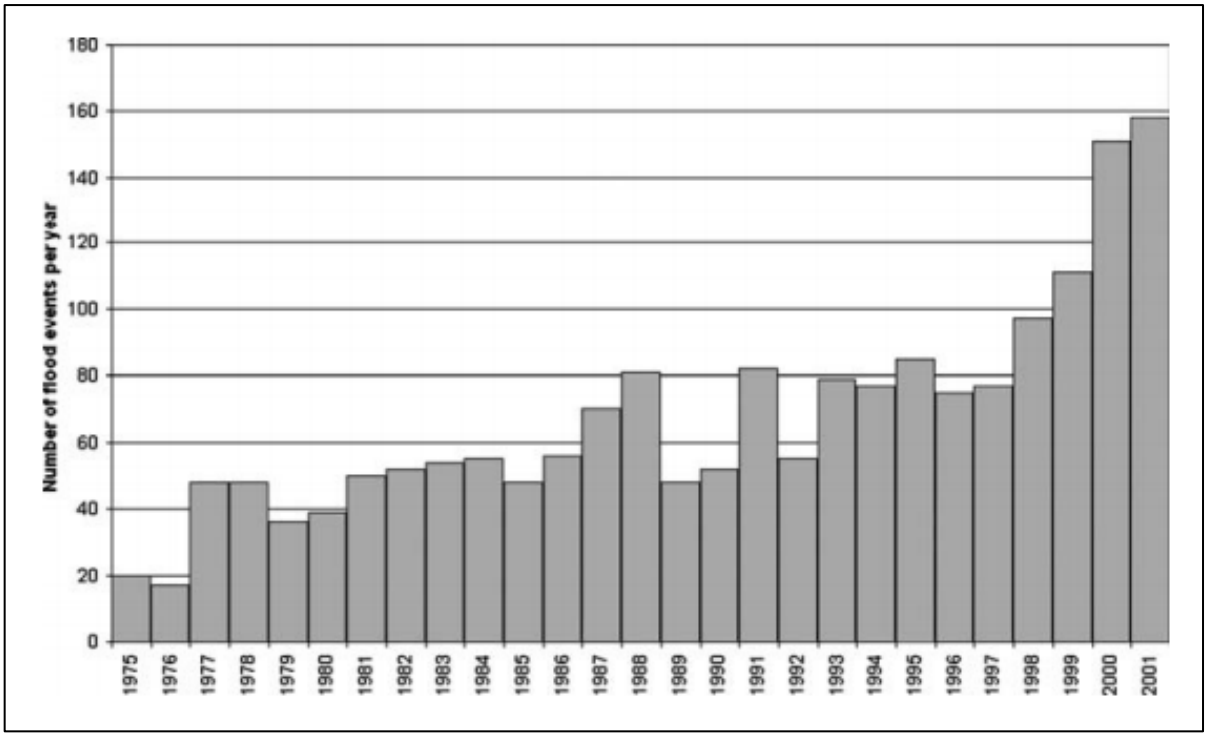

Figure 2: Number of flood events per year included in the OFDA/CRED dataset for the period 1975-2001 (Jonkman, 2005).

The intense weather change of the past few years makes the use of reliable flood models necessary, in order to simulate and predict the effect of extreme events. When it comes to flood events, there is not enough time to apply any kind of emergency procedures. So, converting the precipitation forecasts into runoff forecasts can improve the runoff forecast lead time (Anderson et al. 2002). Modeling has been

used to deal with major insurance issues from emergency managers, city planners, meteorologists and scientists all over the world (Knebl et al. 2005).

In order to reduce flood risk in Europe, the EU Member States are required, according to EU Directive 2007/60 to create flood hazard and flood risk maps and develop flood risk management plans for the rivers within their boundaries. Through the years, several changes have been made to the EU Flood Directive. The new Directive 2007/6/EC now requires Member States to recognize areas that are in danger from flooding, to map the flood area and property and humans at risk within these areas and to take sufficient and coordinated actions to decrease flood risk (Tsakiris et al. 2009). Cyprus has made the last reporting for the flood directive in 2016 in order to follow the EU Directive 2007/60. Flood risk and flood hazard maps have been made, however, they did not consider climate change (WDD, 2016).

1.2. Research Objectives

The overall goal of this project is to analyze the effect of climate change on extreme rainfall for Cyprus. The first objective is to evaluate the modeled data by comparing them with the observed annual daily rainfall extremes of a 30-year period, from 1980 to 2010. Secondly, modeled rainfall data of the past (1980-2010) will be compared with modeled data of the future (2021-2100) to find changes in extremes. The study area is Limassol, where the Water Development Department has identified three flood sensitive areas. We compare the modeled rainfall extremes with the extremes used for the preparation of the flood maps for the Garyllis Basin in Limassol, in the framework of the EU Flood Directive.

2. Literature Review

2.1. Statistics of climate data

2.1.1. General information

Extreme events are very important when it comes to climate. Environmental and socioeconomical impacts result from changes in the magnitude and frequency of climatic extremes. The current literature review focuses on the frequency analysis of extremes and the use of different methods according to several sources. A comparison between the methods is described with the application fields of each one. Also a description of the differences in time scale and modeling period is made in the following sections in order to understand the time period that each method was applied.

2.1.2. Definitions

Plenty of methods have been used in the past to estimate return values of climatic extreme events. Return period (T) value is the threshold that is exceeded by an annual extreme in any given year with the probability p=1/T, where T is expressed in years. Return values are the terms of a distribution of annual extremes and are found by using a generalized extreme value (GEV) distribution that is fitted at every grid point to samples of annual temperature and precipitation extremes (Kharin et al. 2007).

The L-moments method, also known as the method of probability-weighted moments, can be used to estimate the three GEV distribution parameters of location, scale and shape (Kharin et al. 2007). Maximum likelihood estimation is another method that is used to estimate the values of the parameters of a model. The parameter values are found such that they maximize the likelihood that the process described by the model produced the data that were actually observed (Brooks-Bartlett, 2018).

2.1.3. Extreme value distributions, maximum likelihood, moments and L-moments methods

Kharin and Zwiers (2000) described the L-moments method and compared it with the maximum likelihood method to analyze extremes in three 21-year time periods centered at years 1985, 2050 and 2090. For each time period they estimated return values at every grid point for 10-, 20-, 50-, and 100-year return periods.

The return values are found in two steps. First, the Generalized Extreme Value (GEV) distribution is fitted to a sample of annual extremes by the method of L moments, and then they are obtained by inverting the fitted GEV distribution. The advantages of the L moment method over the asymptotically optimal maximum likelihood method are that it is simpler to compute and its parameter estimates have better sampling properties for short samples (Kharin et al. 2000). The time period was 21 years and they used an ensemble with three members, so the total sample size for each time window was 63 (Kharin et al. 2000).

L-moments method has also many disadvantages which make it not suitable for any kind of sample. This method can produce non feasible parameter estimates. As a result, the estimated distribution may not contain all of the data from which distribution parameters are estimated (Kharin et al. 2000).

According to Kharin and Zwiers (2005) the L moments are only defined for identically distributed random samples, so in the case of time-dependent GEV distribution parameters, we are not able to use this method for their estimation. They found that the maximum likelihood method is more feasible than the L-moment

method, and in most cases was preferred for estimating extreme values in transient climate change simulations. Last, the L-moment method could produce biased return value estimates for the climate change rates that are typical in the considered simulations by the end of the twenty-first century, and this biases get larger when the sampling windows get bigger.

On the other hand, Kharin et al. (2007) choose the L-moment method over the more used method of maximum likelihood to estimate the GEV distribution parameters, due to the short 20-year samples that are available for analysis. For the purpose of the study, they analyzed annual extremes of daily maximum and minimum surface air temperature and of 24-h precipitation amounts for the time period 1981-2000 from simulations of the twentieth-century climate, and for two 20-yr time periods 2046-65 and 2081-2100. This annual extremes are drawn from samples of size 365.

The reason why maximum likelihood method is less efficient than the L-moment method in short samples, is that the estimates of the shape parameter are unreliable and lead to poor performance for return values (Coles and Dixon, 1999). Overall, the L-moment method is an appropriate and viable technique for the task of their paper and the benefits of the powerful maximum likelihood method do not override the simplicity of the L-moment method (Kharin et al. 2007).

In Kharin et al. (2013) estimates of the parameters are made for each year from overlapping 51-year time window. An advantage of the maximum likelihood method is that time covariates can be included, which allows extreme value statistics to be estimated more accurately for each individual year.

Martins et al. (2000) examine the behavior of Maximum Likelihood estimators in small samples and demonstrate that absurd values of the GEV shape parameter κ can be generated in small samples. This explains the unstable behavior of small-sample maximum likelihood quantile estimators. An advantage of maximum likelihood estimators is that they can employ censored information without difficulty.

Salas et al. (2014) studied the nonstationarity in hydrologic extremes. To tackle this case, they propose several approaches such as frequency analysis, in which the parameters of a given model vary in accordance with time. Through the paper, he shows that some basic concepts and methods used in designing flood-related hydraulic structures assuming a stationary world can be extended into a nonstationary framework. Also, the authors present a simple and unified framework to estimate the return period and risk for nonstationary hydrologic events.

To evaluate the performance of hydraulic structures under nonstationary conditions, using basic metrics such as return period and risk, the authors consider three conditions of nonstationarity, increasing events, decreasing events and random shifting events.

In conclusion, Salas et al (2014) have shown that the extension of the well-known geometric distribution to include time varying exceedance probabilities has allowed determining the return period and risk for nonstationary conditions of extreme events. The parameters of the nonstationary GEV distribution have been generally estimated based on the method of maximum likelihood but for the stationary GEV other methods (such as those based on the L-moments) may be more efficient, especially for small sample sizes. Last but not least, based on the GEV family, it is possible to determine the standard errors of parameters under nonstationary conditions.

Gao et al. (2016) studied the nonstationary modeling of extreme precipitation in China. Through the study the authors found out that the nonstationary GEV distributions performed better than their stationary equivalents.

Moreover, they use nonstationary GEV distribution with time and climate indices as the potential covariates to model the nonstationary precipitation extremes in China. Maximum likelihood method is used for parameter estimation in both stationary and nonstationary GEV distributions.

Generalized extreme value distribution (GEV) combines the three possible asymptotic extreme value distributions that are obtained as the limits taken over samples of increasing size. This three types are Gumbel distribution (EV-1) for κ=0, Fréchet (EV-2) for κ<0 and Weibull (EV-3) for κ>0 (Kharin et al. 2005; Koutsogiannis 2004). The GEV distributions are:

$$H(x) = \begin{cases} exp\left\{-\exp[-\dfrac{x-\psi}{\lambda}]\right\}, & \kappa = 0 \\ exp\left\{-[1-\dfrac{\kappa(x-\psi)}{\lambda}]^{1/\kappa}\right\}, & \kappa < 0, \quad x > \psi + \lambda/\kappa \\ exp\left\{-\left[1-\dfrac{\kappa(x-\psi)}{\lambda}\right]^{1/\kappa}\right\}, & \kappa > 0, \quad x < \psi + \lambda/\kappa \end{cases}$$

where

H(x) is the GEV distribution function,
x is the extreme value,
κ (kappa) is the shape parameter,
ψ (psi) is the location parameter
λ (lambda) is the scale parameter.

Koutsogiannis (2004) reports that for the study of hydrological extremes, Gumbel distribution might give the wrong conclusions for the risk because it underestimates the higher extreme rainfall values. For that reason, it is generally not used. On the other hand, the extreme value distribution of type II (EV2) is more stable. Thus, Gumbel distribution is not the right distribution for rainfall extremes and EV2 distribution is a choice that represents the reality largely and it is easy to use even with short rainfall data. Although EV3 distribution is the most usually found in nature, it is not used for rainfall because significant intense rainfall values are uncontrolled and high falls may be recorded. As for the EV1 distribution, it has one potential disadvantage, which is very important from the engineering point of view. Comparing with EV2, for large return periods it gives the smallest possible quantiles for any positive value of the shape parameter κ. Thus, EV1 has the highest possible risk for engineering structures.

The shape parameter κ is the most significant parameter of the GEV distribution because it determines the kind of the distribution of maxima (EV1 or EV2) and therefore the behavior of the distribution in its tail. A set form of shape parameter, κ, simplifies the fitting and therefore the general mathematical handling of the distribution. The fixed value of κ=0.15 reduces the weighted average of square errors between empirical and EV2 quantiles. Also, this value of κ could have a negative result in predictions of the EV2 distribution for short return periods and underestimate the precipitation extremes, because of the fault of this distribution. So, if only short return periods are of interest, then the value κ=0.10 should be used (Koutsogiannis, 2004).

By analyzing the estimated GEV shape parameters, we can clearly notice a strong correlation between the shape parameter value and the data length, showing that only very large samples can give its true distribution or the true behavior of the

extreme rainfall. So we cannot rely only on data, as small samples may distort the true result. Also, a shape parameter with negative value is completely inappropriate for rainfall. So, if the data suggest a GEV distribution with negative κ, this should not be used. Instead, it is preferred to use a GEV distribution with a shape parameter value of 0.114 for more safety (Papalexiou et al., 2013).

2.2. Floods study and use of extreme value analysis for Garyllis river, Cyprus

2.2.1. Basic information

Garyllis river is located in Limassol district. In cases of extreme events, people face many problems with flooding in several spots along the river, which affect the economy, society and the quality of life, as these put in danger human lives. For the requirements of the flood directive, the Water Development Department (WDD, 2014) has done a modeling study of Garyllis river along with main actions that are planned to be applied about the protection from floods. For the scope of this project, we used this report as an example for the extreme value analysis.

The northeast end of the Garyllis basin is adjacent to the Louvaras community boundary. The upstream area of the basin is located in the areas of Fasoula and Palodia villages. From the Macedonia Avenue, the river has been diverted to a new river bed leading to Karnagio area. The old river bed of Garyllis passes through the area of Saint Antonios and ends up at the sea, to the old port of Limassol. Figures 3 and 4 show the river basin and the area at which the diversion of the river into the new river bed is made (before the urban area) (WDD,2016).

Figure 3: Garyllis basin, the orange outline shows the basin of the old river bed (source: WDD, 2014).

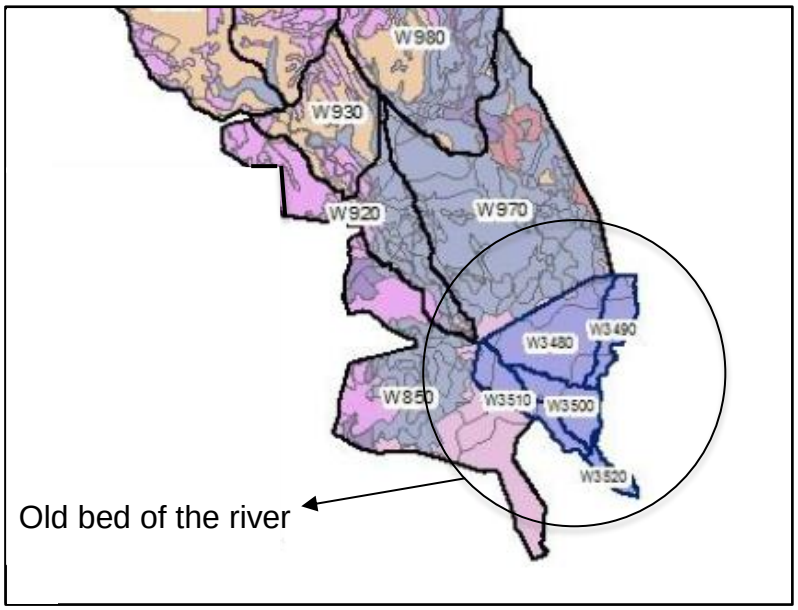

Figure 4: Area of the old bed of the river (source: WDD, 2014).

The flood-affected areas are in areas with prevailing residential use. There are also limited areas with commercial and other activities and very few agricultural land.

2.2.2. Analysis of extremes for Garyllis river and development of flood maps

Water Development Department (WDD, 2014) is using the HEC-HMS model in order to create flood and risk maps for the Garyllis River.

For the model calibration, first they used the observed flow data of streamflow station 9-4-3-80 for two extreme events, 9/01/1989 and 21/11/1994, and compared them with the modeled flow data before and after the optimization of the hydrological parameters of the basin. For the extreme flood event of 9/01/1989, they used the total rainfall data from 8/01/1989 until the end of the event, of the stations 311, 347 and 400. Stations 338, 428 and 325 were out of order, so these could not be used. The total rain over the watershed area was 181 mm. For the extreme flood event of 21/11/1994, they used the same rainfall stations with a total rain of 234 mm. They used a 5-min time step for two days before the event, the day of the peak and one day after the event, but the duration of the events is 24 hours as almost all rain fell within a 24-hour interval. For the calibration and optimization, they changed the curve number (CN), the initial losses, the antecedent soil moisture condition (AMC), the lag time and the Manning coefficient.

For the design flood events, they estimated the precipitation over the watershed area from the weighted point rainfall. They used a reduction factor of 0.95 for the watershed up to the flow station 9-4-3-80 and 0.93 for the downstream area of the new bed of the river. For the old bed of the river they didn't use any reduction factor. The time of concentration (t_c) of the watershed was 2.86 hr for flow station 9-4-3-80, 5.07 hr up to the outlet of the new bed of the river and 1.16 hr for the watershed of the old bed of the river.

To choose the design floods as input to the model, they compared the modeled flows with the analysis of the annual extreme flows of station 9-4-3-80. They used Hydrognomon to fit a GEV to the observed annual streamflow maxima and got the extreme flow for T=20, 100 and 500 years. They concluded that the two streamflow values have a better match if for return period T=20 years they multiply the time of concentration for the basin with a factor 2 and for T=100 and 500 years with a factor

1.5. For the new and the old beds of the river they chose durations as shown in table 2 for all the rain stations of the area.

Table 2: Duration (hr) of the design flood events and rain values (mm) for the new and old beds of the river for return periods T for the available rain stations (WDD, 2014).

New bed of the river						
Rain station	T=20 yr **Rainfall (mm)**	T=20 yr **Duration (hr)**	T=100 yr **Rainfall (mm)**	T=100 yr **Duration (hr)**	T=500 yr **Rainfall (mm)**	T=500 yr **Duration (hr)**
313	69	10.14	83	7.61	-*	-
325	83	10.14	114	7.61	163	7.61
338	67	10.14	75	7.61	85	7.61
347	76	10.14	99	7.61	133	7.61
388	61	10.14	73	7.61	90	7.61
391	59	10.14	78	7.61	107	7.61
394	64	10.14	90	7.61	131	7.61
400	99	10.14	148	7.61	242	7.61
428	67	10.14	76	7.61	85	7.61
Old bed of the river						
Rain station	T=20 yr **Rainfall (mm)**	T=20 yr **Duration (hr)**	T=100 yr **Rainfall (mm)**	T=100 yr **Duration (hr)**	T=500 yr **Rainfall (mm)**	T=500 yr **Duration (hr)**
311	-	-	-	-	80	7.61
391	43	2.32	57	1.74	78	1.74
394	45	2.32	62	1.74	91	1.74

-*no data were reported

In the report, the source of the rainfall data for the design events is not mentioned. Based on a comparison of the values with the data from the report of the Meteorology Department (Pashiardis, 2009), we assume that they used those values and adjusted them through interpolation for the duration of the design events. These data are presented in table 3.

Although the WDD followed the requirements of the Flood Directive, some important information is missing in the report. The exact duration of the events is not given and the rainfall data cannot be easily found and the order of the followed steps is vague.

Table 3: Code, location, coordinates and elevation for the rainfall stations, maximum rainfall for return periods of 20, 100 and 500 years for durations of 6 and 24 hours (Pasiardis, 2009).

Rain station					Rainfall (mm)					
					T=20 yr		T=100 yr		T=500 yr	
Code	Location	Coordinates (°)		Elevation (m)	6 hr	24 hr	6 hr	24 hr	6 hr	24 hr
		Longitude	Latitude							
311	Alassa (E.S.)	32.93	34.77	340	54.0	72.0	66.0	88.8	76.2	103.2
313	Kouris (Dam)	32.92	34.72	220	62.4	81.6	78.6	103.2	93.0	120.0
394	Limassol (Public garden)	33.06	33.67	8	57.0	79.2	85.2	117.6	123.0	170.4
400	Kalo Chorio (Limassol) (P.S.)	33.02	34.85	730	83.4	132	137.4	218.4	224.4	357.6

2.2.3. Future plans

Currently, there is no protection zone along the river, except from the estuary area and the construction of the Polemidia dam and the diversion of the bed of Garyllis that reduce the flood risk. Until now, historical floods records of very low to medium and high risk before the river bed diversion exist but still it is necessary to ensure proper management of flood events. Based on the current conditions, WDD's future plans include the investigation of the design of required structures for the settlement and increase of the water supply of the old bed of Garyllis in order to supply with safety at least the 20 years floods. Also, they are planning to upgrade and check the adequacy of the road crossings that are located along the old and new bed of the river, design widening and configuration structures and/or replacement of existing techniques and interventions.

3. Methods

3.1. Analysis of observed and modeled rainfall extremes of 1981-2010

Data

The annual daily rainfall extremes used for the analysis were extracted from the ERMIS-F climate simulations (https://ermis-f.eu). The period 1981-2100, was simulated with the Weather Research and Forecasting (WRF) model, using initial and boundary conditions from the bias-adjusted version of the CESM1 global earth system model (Bruyère et al., 2014; 2015). Observed greenhouse gas concentrations are used for the recent past (1980-2005) and the RCP8.5 (business-as-usual) scenario is used for the future climate (2006-2100). We used the same domain and model configuration for the 12-km resolution simulations as Zittis et al. (2017). For Cyprus, the 12-km grid covers 78 land grid cells. At this resolution the model elevation ranges from 0 to 850 m above sea level (asl).

For the purpose of the analysis, 13 grid cells that cover the Limassol area were selected (see Figure 5 and Table 4). We separated the cells in three main groups according to their elevation: coast (0 – 200m), foothills (300 – 600m) and mountains (>600m). The model data are evaluated against the 1-km gridded daily dataset of 1980-2010 (Camera et al., 2014). The 1-km data were interpolated to the 12-km grid cells of WRF and then the annual extremes of the daily rainfall were extracted.

Figure 5: Map of grid point's location (source: Google Earth Pro).

Table 4: Coordinates and elevation of grid points.

Grid point	Coordinates		Elevation (m)
	Longitude (decimal °)	Latitude (decimal °)	
1	32.67	34.68	173
2	32.78	34.68	213
3	32.89	34.68	205
4	33.00	34.68	174
8	32.78	34.79	532
9	32.89	34.79	572
10	33.00	34.79	528
11	33.11	34.79	440
12	33.22	34.79	335
19	32.78	34.90	760
20	32.89	34.90	850
21	33.00	34.90	813
22	33.11	34.90	693

Generalized extreme value distribution

For the extreme rainfall analysis we use the methods of L-moments and maximum likelihood estimation (MLE). The equations of these two methods are presented in Appendix A. For both methods we choose the Generalized Extreme Value distribution for maxima (GEV-max) to fit the data because this is the best for rainfall extremes (Koutsogiannis, 2004). We use Hydrognomon, a software tool for the management and analysis of hydrological data, to derive the parameters of the GEV distribution. Hydrognomon is operationally used by the largest water organization as well as technical corporations in Greece (Kozanis et al. 2005). In total we have 26, 13 modeled and 13 observed data sets of annual extremes of daily rainfall. The basic equation of GEV distribution is shown in section 2.1.3.

3.1.1. Estimation of the parameters of the distribution

L-moments method will be used for the GEV distribution and will be compared with the maximum likelihood method as all described by Koutsogiannis (2004) and Papalexiou et al. (2013), in order to choose which distribution fits the data the best. L-moments method is simpler to compute and its parameter estimates have better sampling properties for short samples (Kharin et al. 2007).

A fixed shape parameter, κ, simplifies the fitting and the general mathematical handling of the distribution (Koutsogiannis, 2004). A κ-value of 0.15 was recommended by Koutsogiannis (2004) based on the weighted average of square errors between empirical and EV2 quantiles. So for that reason GEV distribution with a specified κ of 0.15 was selected. Papalexiou et al. (2013) found a κ-value of 0.10 can be selected if only short return periods are of interest. This value will also be tested.

Predicted extreme values of return periods 20, 50 and 100 years will be produced and according to the length of the data, other values will be compared with empirical ones. Finally, we use goodness-of-fit tests, mean absolute error and bias to decide which distribution fits our data in the best way.

Empirical distribution and Weibull plotting position

Plotting order-ranked data is a standard technique that is used in estimating the probability of extreme weather events. In this project the Weibull plotting position is used for the fitting of the data. First, the data have been sorted in descending order and then through the rank we calculated the probability of exceedance and the return periods. In estimating the return periods there is only one correct plotting position (Makkonen, 2005):

$$p = i/(n+1)$$
$$T = 1/p$$

where

p is the probability of exceedance,
i is the corresponding rank
n is the number of years with data
T is the return period.

We use this equations to find the empirical extreme values for the return periods of, T=10.33, T=15.50 and T=31 years for the 13 grid points of annual extremes for the years 1981 to 2010. We also use it for the return periods of the same grids for 80 years' time period, 2021 – 2100 (T=81, T=40.50 and T=27 years).

3.1.2. Evaluation criteria

Goodness of fit describes how good the model fits a set of data. In this project, in order to conclude whether a distribution fits sufficiently a data set and provides a reasonable description of the behavior of a sample, the Kolmogorov-Smirnov (KS) and Chi-square (x^2) tests along with bias and mean absolute error (MAE) are used.

The KS statistic gives the quantity of a distance between the empirical distribution function of the sample and the cumulative distribution function of the reference (theoretical) distribution. When this maximum distance exceeds a certain value, the null hypothesis that the data are drawn from the fitted model is rejected. The KS test tends to be more sensitive near the center of a distribution than at the tails (Beranova et al. 2018).

The Chi-square test is the oldest of all goodness of fit tests and is not very subjective comparison of frequency histograms with fitted distributions. In this procedure, the range of the sample data is divided into a selected number of intervals and the number of data points falling in each interval is compared with the expected number predicted by the fitted distribution (Cooke et al. 1993).

Mean absolute error, MAE, is computed by taking empirical values and using prediction values for return periods equal to the whole time period, to the 1/2 and to the 1/3 of the time period (T=31, T=15.50 and T=10.33 years). We also compute the Bias for the three prediction extremes. The following equations for MAE (mm/d) and Bias for 3 extremes (mm) are used:

$$\text{Bias}_i = (\text{Predicted value for } T_i) - (\text{Empirical value for } T_i)$$
$$\text{Bias} = \sum_{i=1}^{3} \text{Bias}_i$$
$$\text{MAE} = \frac{\sum|\text{Bias}_i|}{3}$$

3.2. Modeled extreme rainfall analysis for extremes alteration considering climate change for 2021-2100

Data

In this section, we use the same 13 grid cells that we use in objective 1. First, we plot the data and make an exploratory trend analysis with linear regression. Then we analyze three different sets of future time periods, as described below. We will use Hydrognomon to fit extreme value distributions, make forecasts of return periods T=20 and 100 years and compare modeled empirical extremes to the corresponding predicted extreme values.

3.2.1. Modeled extreme rainfall analysis for the future 30 year time period (2021-2050)

First, we analyze the future modeled data of the 13 grid cells for 30 year time period, from 2021 to 2050, in order to compare them with the 30 year time period results of the first objective. For this reason, because we already chosen the L-moments with specified κ=0.15 as the best estimator for data of 30 year time period, we are using only this method to analyze the data of this objective.

3.2.2. Modeled extreme rainfall analysis for 60 year moving time periods for coastal and mountain grid cells

In this part of the project, we select 2 out of the 13 grid cells, the 3rd and the 20th to make an analysis of 5 moving time periods of 60 years. We choose this specific grid cells because we wanted to analyze the highest grid point (the 20th) and the one with the same longitude and a smaller elevation (the 3rd) (Table 4). L-moments without specified κ method, maximum likelihood method and the 5 moving time periods of 2021 – 2080, 2026 – 2085, 2031 – 2090, 2036 – 2095 and 2041 – 2100 are selected for the frequency analysis.

3.2.3. Modeled extreme rainfall analysis for future 80 year time period (2021-2100)

Finally, we apply the analysis of 80 year data, from 2021 to 2100, using all three methods, L-moments with and without specified κ and maximum likelihood method, for the estimation of the GEV-Max distribution parameters. We compare the predicted extreme values to the values of the 30 year time period of the past (1981-2010).

4. Results and discussion

4.1. Analysis of observed and modeled rainfall extremes of 1981-2010

First, by working with the 30 years modeled and observed data, we found the empirical extreme values of the maximum return period of T=31 years. A comparison between the values is presented in figure 6. We note that most of the values of the modeled extremes (10 out of 13) are underestimated, which means that they are lower than the observed extremes. The largest difference between observed and modeled was 54.3 mm for grid point 22.

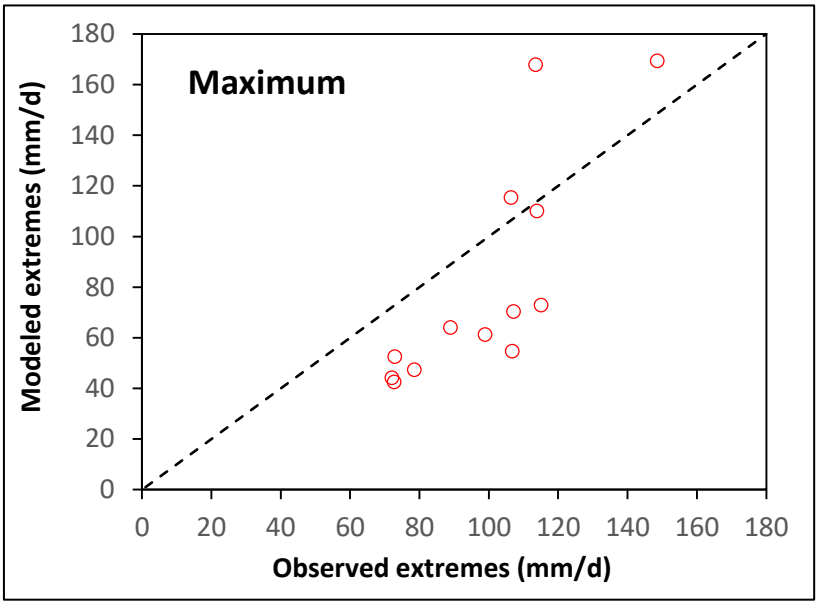

Figure 6: Maximum values of observed and modeled data of 30-year time series for 13 grid points with return period T=31yr and the 1:1 line (dashed).

4.1.1. Estimation of the parameters of the distribution

In table 5 we present the results of the 30-year analysis using the method of L-moments with specified κ=0.15. Tables A1 and A2 in Appendix B, represent the same data as table 5, for L-moments method without specified κ and for the maximum likelihood method, respectively.

As we notice in table A1, for the fitted data series of 30 years, by using the L-moments method (without κ specified), we have the freedom of fit but most of the cells get a negative value of κ. Specifically, 9 out of 13 modeled and 9 out of 13 observed have a negative κ and thus they have a left-hand tail that is bound. Although we can fit those data as statistically acceptable, negative values of κ are not what we would expect for extreme values of rainfall (Papalexiou et al. 2013). For that reason L-moments method for three parameters is not so good for the 30-year data.

Table 5: First and second moment (lambda, psi) of GEV Max distribution using L-moments method with specified κ (0.15) for the 30 years annual daily rainfall extremes for modeled and observed data of 13 grid cells; p-values of Kolmogorov-Smirnov (KS) and Chi-square (x^2) tests; bias and mean absolute error (MAE) of predicted versus empirical extremes (T=10.33, 15.50, 31 years); and predicted rainfall extremes for return periods (T) of 20, 50 and 100 year.

Grids	Data	lambda	psi	KS-p (%)	x^2-p (%)	Bias (mm)	MAE (mm/d)	Predicted extremes (mm/d)		
								T=20yr	T=50yr	T=100yr
1	Mod	5.47	4.94	9.1**	1.8*	14.2	4.7	47.5	56.0	63.2
1	Obs	7.25	4.80	98.6	43.5	-6.3	2.1	61.9	73.2	82.8
2	Mod	6.70	3.59	34.9	13.5	0.3	1.0	49.1	59.5	68.4
2	Obs	7.65	4.84	99.9	71.7	-14.4	4.8	65.6	77.6	87.7
3	Mod	6.47	3.45	33.9	31.1	-1.0	2.9	46.5	56.6	65.2
3	Obs	8.76	4.24	87.8	22.3	-8.2	2.7	70.0	83.7	95.2
4	Mod	5.81	3.70	19.3	1.3*	3.4	2.3	43.3	52.3	60.0
4	Obs	9.78	3.79	94.2	51.3	0.2	5.3	73.7	88.9	101.8
8	Mod	8.01	4.20	79.4	60.7	-3.0	1.1	63.6	76.1	86.7
8	Obs	14.33	3.32	87.9	22.3	11.4	3.8	101.3	123.6	142.6
9	Mod	6.61	5.32	65.5	31.1	3.4	1.9	59.9	70.2	79.0
9	Obs	13.37	3.47	91.4	18.9	1.4	4.3	96.4	117.3	135.0
10	Mod	6.91	4.55	10.0	0.6*	7.4	3.0	57.3	68.1	77.2
10	Obs	15.14	2.94	99.9	43.5	-17.6	9.4	101.2	124.9	144.9
11	Mod	8.20	3.37	87.7	36.8	3.6	1.2	58.4	71.2	82.0
11	Obs	13.98	3.02	72.0	43.5	12.3	8.0	94.6	116.4	134.9
12	Mod	9.34	2.82	100.0	71.7	-13.2	4.4	61.3	75.9	88.2
12	Obs	16.87	2.64	73.5	9.7	-6.60	6.0	107.7	134.0	156.3
19	Mod	11.5	4.00	69.3	2.6*	-47.6	15.9	89.0	107.0	122.2
19	Obs	13.25	3.91	51.0	31.1	3.70	4.9	101.4	122.1	139.6
20	Mod	19.82	2.77	90.8	18.9	-51.6	17.2	129.0	159.9	186.1
20	Obs	20.57	3.59	73.6	60.7	12.0	9.0	150.7	182.9	210.0
21	Mod	15.75	3.23	75.5	16.0	12.3	5.5	109.9	134.5	155.3
21	Obs	16.84	3.46	82.9	60.7	18.9	8.9	121.3	147.6	169.8
22	Mod	17.16	2.38	91.2	43.5	-85.9	28.6	105.0	131.8	154.5
22	Obs	17.65	3.31	77.1	22.3	29.6	11.0	124.5	152.1	175.4

*Not accepted at alpha 0.05;**Not accepted at alpha 0.10

The maximum likelihood method, gives “floating point overflow” error for the modeled data of the grids 2, 3, 4, 10 and for the observed data for the grids 8, 20, 22 when trying to fit our data. For three cells hydrognomon discarded several values out of the 30 data points in order to find the parameters of the distribution. Also, this method gives a large value for the bias in most of the grid points (Table A2).

4.1.2. Evaluation criteria for the fitting methods of the GEV-Max distribution

Comparing the p values of the goodness of fit tests between the methods, we note that the method with the most acceptable values is L-moments with specified κ=0.15. By using this method, for the x^2 test, 4 modeled and 1 observed data series are not accepted for alpha 0.05 and alpha 0.10 respectively (Table 5). On the other hand, by using the L-moments method with three parameters (without specified κ), 5 modeled and 3 observed data series are not accepted for these alpha values for the x^2 test (Table A1).

Observing the tables 5 and A1, we notice that the bias and MAE of L-moments method with and without specified κ are similar. For the L-moments method with κ specified, bias is -47.6 mm, -51.6 mm and -85.9 mm for the modeled data of grid points 19, 20 and 22, respectively. Without κ specified, bias is -48 mm, -47.8 mm and -82.6 mm, respectively, for the modeled data of these grids.

In table 6, we present the maximum, minimum and mean value of MAE for both the modeled and observed data series of the L-moments method with and without specified κ. By using the method with three free parameters, we can fit the distribution better than with two free parameters. For that reason, all values of MAE (max, min, mean) are larger for the L-moments with κ specified than for L-moments without κ specified. Nonetheless, the values are very close to each other so we can't distinguish one method above the other. Considering the negative κ values, which diverge from the theory, we identify the L-moments method with specified κ as the better method for our 30-year data.

Table 6: Mean, minimum and maximum values of mean absolute error (MAE) of predicted versus empirical extremes of 13 grid cells for modeled and observed annual daily rainfall extremes for L-moments method with specified κ (Table 5) and without specified κ (Table A1).

Method	Data	MAE (mm/d)		
		Mean	Min	Max
L-mom (κ=0.15)	Observed	6.2	2.1	11.0
L-mom	Observed	4.5	2.1	9.0
L-mom (κ=0.15)	Modeled	6.9	1.0	28.6
L-mom	Modeled	6.3	1.2	27.5

4.1.3. Effect of shape parameter (κ-value)

We also tested κ=0.10 instead of 0.15, as presented in table A3, for the modeled data of all 13 grid points. Specifically, for the grid points 1, 4, 10 and 19, most p values for the KS and x^2 tests are small compared to the rest of the cells, which means the data are not well fitted. In general, by comparing the results for this two values of κ (Table 5 and A3) we conclude that there is a difference between the values but not a substantial one, so we keep 0.15 as the best value for κ (Koutsogiannis, 2004). For the extreme values of T=100 years the largest difference between the two methods corresponds to grid point 3 with values 83.4 mm/d (κ=0.10) and 65.2 mm/d (κ=0.15).

Figure 7 shows the comparison between the 30-year modeled rainfall extremes of time period 1981 – 2010 for return period T=100 years, using L-moments method with κ=0.15 and κ=0.10. From this figure, we note that almost all grid cells' extreme values, except from grid cell 3, for the return period of T=100 years for the GEV-Max distribution function with κ=0.15 are higher than the values for the distribution function

with κ=0.10. However, the difference between all values is negligible, with all the points almost on the 1:1 line. In conclusion, L-moments with specified κ=0.15 is the best method to fit our data of rainfall extremes with the most acceptable values of p and the smallest errors.

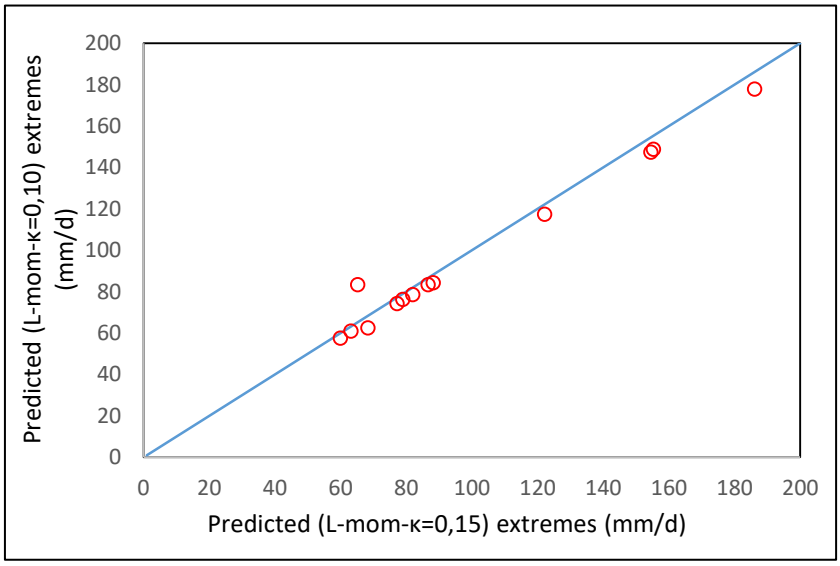

Figure 7: Extreme values of modeled data series of the 13 grid cells for return period T=100 years using L-moments with κ=0.15 and κ=0.10.

4.1.4. Observed versus modeled data for the maximum generalized extreme value distribution (GEV-Max)

By comparing the modeled data with the observed data of the 30-year time period for the chosen best method of L-moments with specified κ (0.15), we notice difference between the distribution parameters' values. The location parameter of the GEV-Max distribution, psi, for the observed data is 7 times smaller and 6 times larger than the psi for the modeled data and the scale parameter, lambda, is always larger for observed data, for all 13 grid cells.

With the extreme value analysis, for return period of T=100 years, always the observed extremes are higher than the modeled ones. By fitting the distributions, especially when we specify the tail of them with a specific value of the shape parameter (0.15) we expect that the modeled and observed values will get closer. But, although L-moments method with specified κ (0.15) is a suitable method to fit the rainfall extremes, we still notice that the modeled data cannot represent the observations for the highest extremes.

4.2. Modeled rainfall data analysis for extremes alteration considering climate change for 2021-2100

Figure 8 shows the modeled annual daily extremes of the 80-year time period, from 2021 to 2100, for the 13 grid cells. A linear regression line is indicatively added to show possible trends. The slope of the trend lines of all grid cells is very small and all correlation coefficients are less than 0.1. This indicates that there are no linear trends of annual extremes with time. Most of the figures have a negative slope, so 11 out of 13 grid cells have negative slope with the exception of only two cases, 19 and 20. Negative slopes show that the annual extremes are decreasing in future years, so only for grid cells 19 and 20 we notice a small increase in extremes, with grid 19 having the largest positive slope.

Also, even if the 13 grid points are next to each other, with similar coordinates and elevation groups, we still get differences between them. We notice an extreme in 2045 in all the grid points, but the mountain cells, 19, 20, 21 and 22, get higher extremes in the future around 2080.

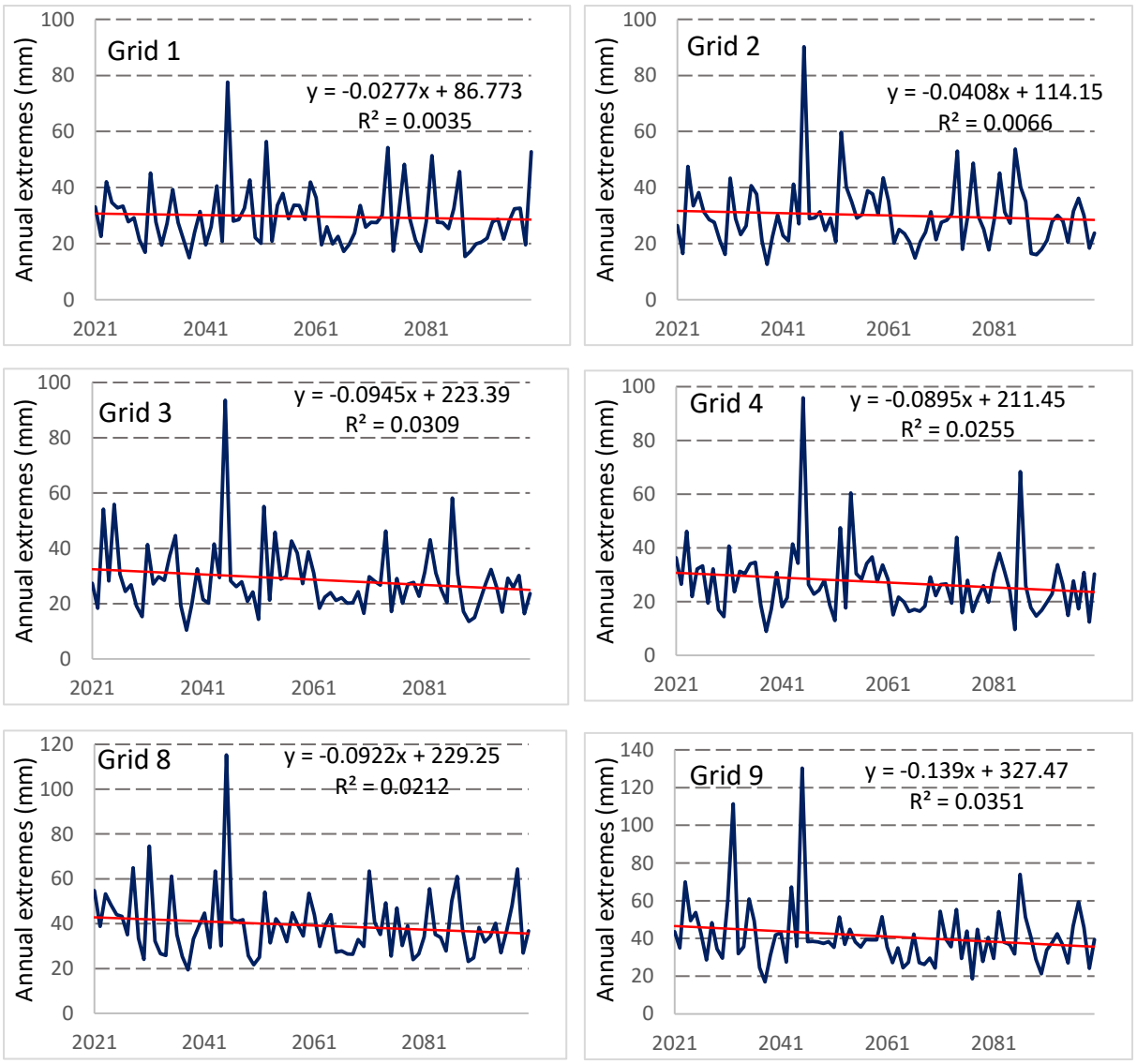

Figure 8: Time series of the 13 grids with the linear regression lines for the annual extremes of 80 year-time period of 2021-2100.

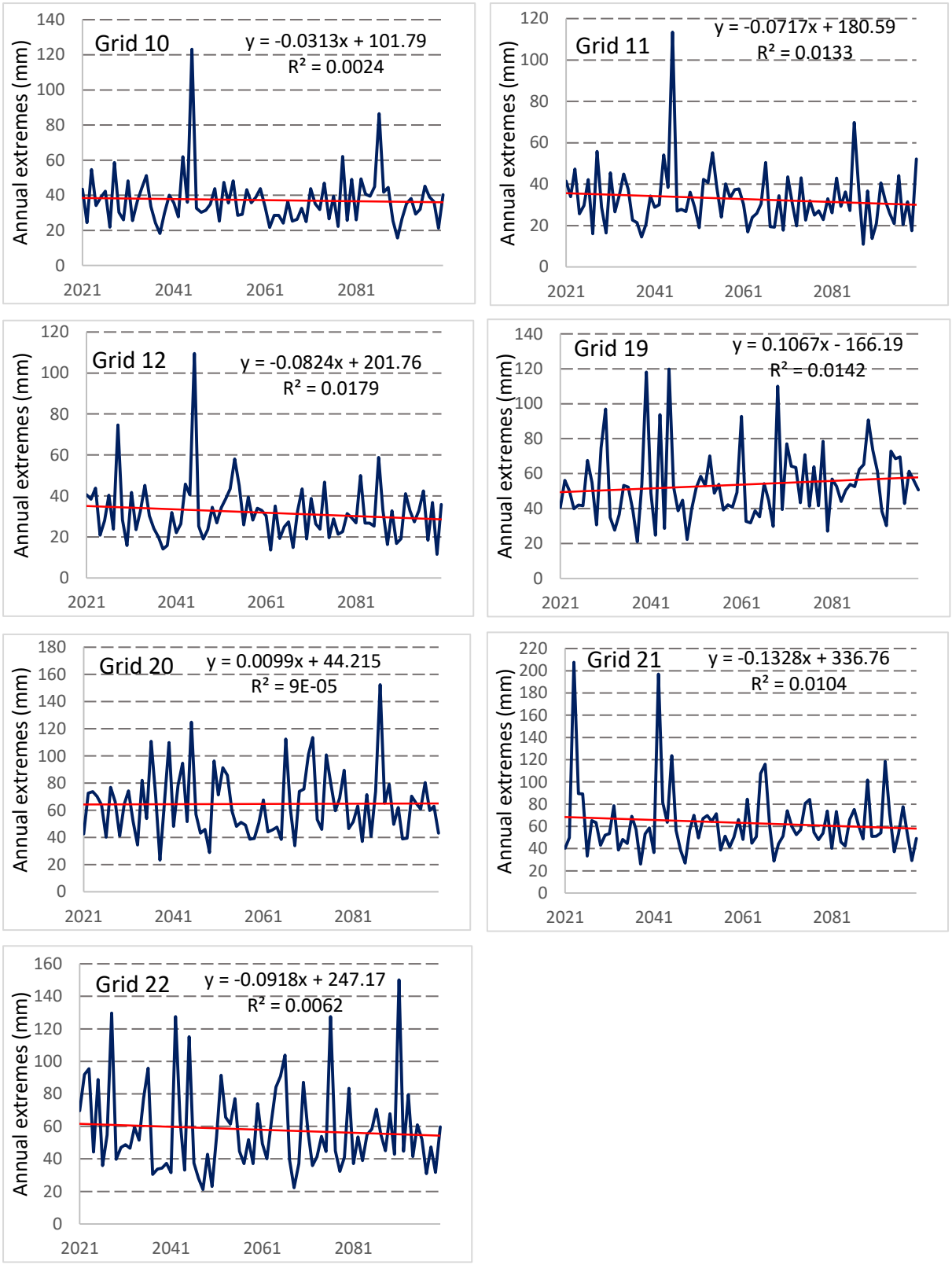

Figure 8 continued: Time series of the 13 grids with the linear regression lines for the annual extremes of 80 year-time period of 2021-2100.

Figure 9 shows the moving average of 30 years for the modeled annual extremes of the 13 grid cells. From this figure we also get same results as figure 8, with a small decrease of 3 mm around 2045, but still not a significant difference in extremes.

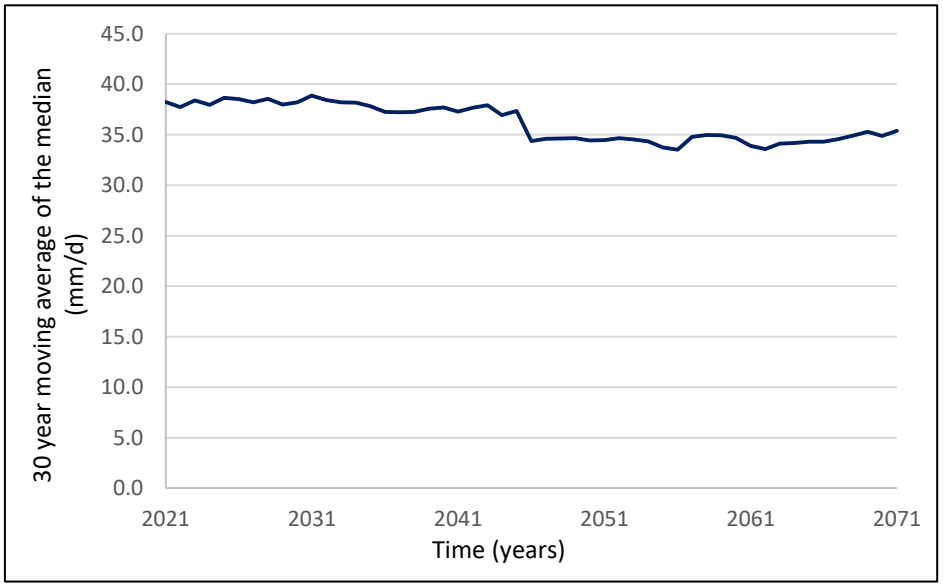

Figure 9: Moving average of 30 years for modeled data of 13 grid cells

4.2.1. Modeled extreme rainfall analysis for the future 30 year time period (2021-2050)

Table 7 includes the results of the analysis using L-moments method with specified κ=0.15 for the 30 years (2021-2050) annual daily rainfall extremes for modeled data of 13 grid cells. All values are accepted at alphas 0.10 and x^2 and KS tests also show that all the data are well fitted. Bias is negative for all the grid points, which means that all empirical extremes of the modeled data are higher than the predicted extremes.

To understand the changes in extremes through time, we compare the modeled extreme values through the empirical method for return period T=31 years of 13 grid cells for the years 1980-2010 with the years 2021-2050 in figure 10 and predicted extreme values in figure 11. In both figures, for the return periods of 31 and 100 years, empirical and predicted extremes for the 30-year time period of 2021-2050 are more intense than the past, 1981-2010.

Table 7: First and second moment (lambda, psi) of GEV Max distribution using L-moments method with specified $\kappa = 0.15$ for the 30 years (2021-2050) annual daily rainfall extremes for modeled data of 13 grid cells; p-values of Kolmogorov-Smirnov (KS) and Chi-square (x^2) tests; mean absolute error (MAE) and bias of predicted versus empirical extremes (T=10.33, 15.50, 31 years); and predicted rainfall extremes for return periods (T) of 20, 50 and 100 year.

Grids	Data	lambda	psi	KS-p (%)	x^2-p (%)	Bias (mm)	MAE (mm/d)	Predicted extremes (mm/d)		
								T=20yr	T=50yr	T=100yr
1	Mod	7.30	3.41	91.6	31.1	-14.0	8.8	52.2	63.6	73.2
2	Mod	7.97	3.09	94.9	13.5	-23.2	12.3	54.5	66.9	77.4
3	Mod	9.31	2.56	86.9	11.5	-34.1	11.4	58.7	73.2	85.5
4	Mod	8.70	2.62	92.4	11.5	-23.4	15.0	55.3	68.9	80.4
8	Mod	12.25	2.70	100.0	84.7	-26.5	9.9	78.9	98.0	114.2
9	Mod	13.94	2.58	93.0	22.3	-56.9	21.6	88.2	109.9	128.3
10	Mod	10.72	2.93	99.9	60.7	-36.8	17.1	71.5	88.2	102.4
11	Mod	10.41	2.62	95.5	84.7	-32.0	15.8	66.2	82.5	96.2
12	Mod	10.94	2.32	97.9	51.3	-38.5	19.1	66.3	83.4	97.9
19	Mod	16.56	2.36	92.6	100	-43.2	14.4	101.1	127.0	148.9
20	Mod	17.29	3.03	55.4	22.3	-6.30	5.20	117.0	144.0	166.9
21	Mod	24.07	2.03	88.7	11.5	-127.9	42.6	139.1	176.7	208.5
22	Mod	20.91	2.02	94.0	22.3	-25.6	12.4	120.4	153.1	180.7

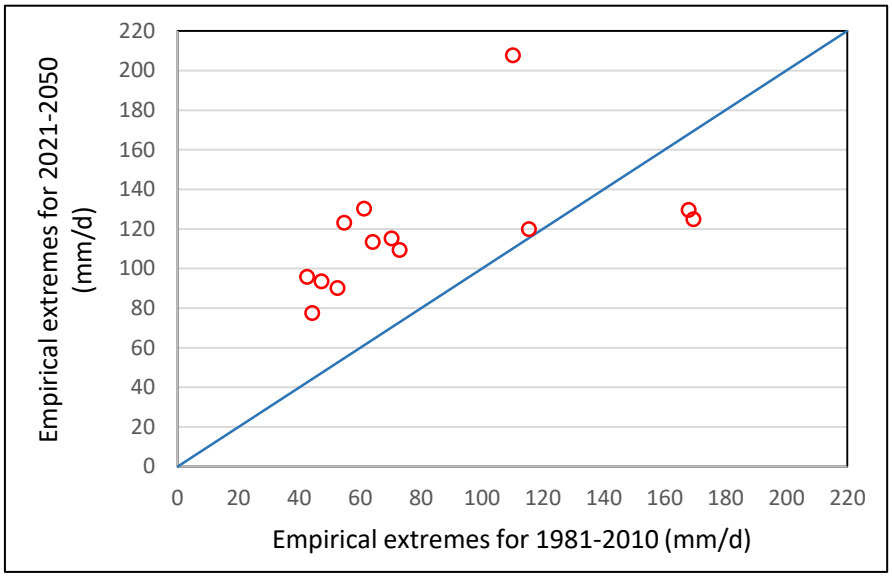

Figure 10: Empirical modeled extreme values for return period T=31 years of 13 grid cells for the years 1980-2010 vs 2021-2050

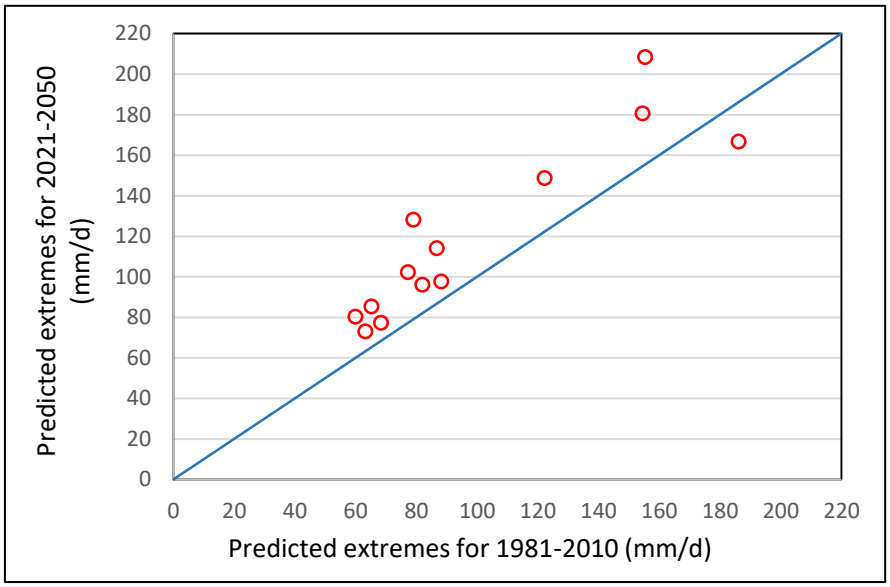

Figure 11: Predicted extreme values for return period T=100 years of 13 modeled grid cells for the years 1980-2010 vs 2021-2050

4.2.2. Modeled extreme rainfall analysis for 60 year moving time periods for coastal and mountain grid cells

In table 8, we present the GEV-Max parameters by using only the L-moments method without specified κ and the maximum likelihood method (MLE) for grid cells 3 and 20, for 60-year shifting time periods, the goodness of fit tests' values for a=5% and 10%, the predicted values for different return periods and the empirical value for return period T=61 years.

All values for both grid cells are accepted for a=5% and for a=10%. The predicted extremes for all the return periods are similar for the two methods and for T=100 years the largest difference between the methods is 14.14 mm and appears in grid cell 20 at the time period of 2041-2100.

For the coastal grid cell (grid cell 3), the empirical extreme value for return period T=61 years remains the same, 93.6 mm in every set for all the time periods. This happens because the bigger annual extreme for this grid cell happens around 2045, so remains steady for the 5 time periods. On the other hand, for the top cell (grid 20), the extreme occurred after 2085 and that is why it shifts at the time period of 2031-2090, from 124.9 mm to 152.4 mm.

In general, we conclude that in order to understand the extremes, it is difficult to get the whole picture with 30 years of data, and then we analyze for 60 years. But, 60-year shifting time period also is not good enough to represent the future, because at the first sets we cannot notice the change that happens in the top grid cell, as it happens at the end.

Table 8: First, second and third moment (kappa, lambda, psi) of GEV Max distribution using L-moments method without specified κ and maximum likelihood method (MLE) for shifting 60 years annual daily rainfall extremes for modeled data of grid cells 3 and 20; p-values of Kolmogorov-Smirnov (KS) and Chi-square (x2) tests; predicted rainfall extremes for return periods (T) of 20, 61 and 100 year; and empirical extreme value for return period T=61 years.

Grids	Method	kappa	lambda	psi	KS-p (%)	x^2-p (%)	Predicted values (mm/d) T=20 yr	T=61 yr	T=100 yr	Emp. value T=61
					2021-2080					
3	L-mom	0.20	7.33	3.18	80.31	1.7*	53.2	70.3	79.1	93.6
3	MLE	0.12	7.92	2.99			52.1	66.0	72.8	93.6
20	L-mom	-0.03	19.87	2.73	96.46	30.8	110.5	130.7	139.2	124.9
20	MLE	-0.05	19.41	2.81			108.1	126.4	134.1	124.9
					2026-2085					
3	L-mom	0.17	7.07	3.30	90.99	8.2**	50.6	65.3	72.6	93.6
3	MLE	0.09	7.62	3.10			49.7	61.8	67.5	93.6
20	L-mom	0.004	19.43	2.72	91.96	18.9	110.8	133.2	143.0	124.9
20	MLE	-0.04	19.21	2.79			107.0	125.6	133.5	124.9
					2031-2090					
3	L-mom	0.16	7.75	2.99	95.06	23.1	52.9	68.6	76.5	93.6
3	MLE	0.11	8.11	2.88			52.1	65.8	72.5	93.6
20	L-mom	0.04	20.14	2.67	90.63	15.5	117.4	143.8	155.7	152.4
20	MLE	-0.02	20.13	2.71			113.0	134.5	143.8	152.4
					2036-2095					
3	L-mom	0.20	7.31	3.07	97.49	1.1*	52.1	68.8	77.4	93.6
3	MLE	0.14	7.73	2.94			51.1	65.3	72.2	93.6
20	L-mom	0.05	19.78	2.70	89.72	30.8	117.1	144.1	156.5	152.4
20	MLE	-0.01	19.86	2.74			112.2	133.8	143.1	152.4
					2041-2100					
3	L-mom	0.22	6.69	3.40	87.63	11.3	50.8	67.5	76.3	93.6
3	MLE	0.19	6.90	3.31			50.3	65.5	73.2	93.6
20	L-mom	0.10	17.35	3.05	95.20	11.3	112.6	140.3	153.5	152.4
20	MLE	0.03	17.20	3.12			107.4	129.4	139.3	152.4

*Not accepted at alpha 0.05;**Not accepted at alpha 0.10

4.2.3. Modeled extreme rainfall analysis for future 80 year time period (2021-2100)

For the 80-year time period we analyze the extremes by using all three methods of fitting the GEV-Max distribution, L-moments with and without κ specified (0.15) and maximum likelihood (MLE) methods, in order to compare the performance of these methods. The results are presented in Table 9. For larger time periods, L-moments method with κ=0.15 may not be the ideal one to use (Kharin et al. 2005).

Comparing the values of the chi-square test for L-moments method with κ=0.15 with L-moments method without specified κ, we notice that 7 out of 13 grid cells have a lower value with specified κ than without, and 6 out of 13 are higher. Mean absolute error is almost the same for all the grid cells for all three methods, with negligible differences.

Figure 12 shows predicted extremes for return period T=100 years for the modeled data of 80 years of 13 grid cells using L-moments method versus L-moments with specified κ (0.15) method. All values are similar when using these two methods, but the higher values become more biased, with 2 out of 13 being higher and 2 lower for the L-moments method with specified κ compared with L-moments without specified κ.

Figure 13 shows the predicted extremes for return period T=100 years for the modeled data of 80 years of 13 grid cells using L-moments method without specified κ versus maximum likelihood method (MLE). Here we note that only for the higher values of extremes, maximum likelihood method underestimates the extremes, with 5 out of 13 grid cells' extreme values being larger for L-moments.

By comparing the modeled data series of the future (2021-2100) with the modeled data series of the past (1981-2010, Table 5) we conclude that the 80 years data have a better fit than the 30 years. P-values of the KS goodness of fit test are larger for the 80 years data than the 30 years data except grid points 9, 12 and 20. Furthermore, for the L-moments method all κ values for the 80 years future data series are positive for all 13 cells. On the other hand, for the 30 years past data series most of the grid cells have a negative κ value. This also indicates that the 80 years future data get a better fit than the 30 years past data.

For the future, extremes appear to be higher than the past. The predicted rainfall extreme for return period of 20 years is expected to be fitted reasonably well by both periods. Indeed, values for almost all 13 grid cells are similar with the exception of grid cells 12 and 20, which are smaller for the future but still close to the modeled past values. The 100-year extreme is higher for the 11 grid cells for the 80-year period, compared to the past, with only grid cells 12 and 20 having smaller values.

Table 9: First, second and third moment (kappa, lambda, psi) of GEV Max distribution using L-moments method without and with specified κ (0.15) and maximum likelihood method (MLE), for modeled annual daily rainfall extremes of 13 grid cells for 2021-2100; p-values of Kolmogorov-Smirnov (KS) and Chi-square (x^2) tests; bias and mean absolute error (MAE) of predicted versus empirical extremes (T=27, 40.50, 81 years); and predicted rainfall extremes for two return periods (T).

Grids	Method	k	lambda	psi	KS-p (%)	x^2-p (%)	Bias (mm)	MAE (mm/d)	Predicted extremes (mm/d)	
									T=20yr	T=100yr
1	L-mom-k	0.15	6.96	3.50	36.86	0.06*	-8.1	4.2	50.5	70.5
1	L-mom	0.11	7.27	3.37	50.27	0.3*	-11.4	4.7	50.2	68.4
1	MLE	0.14	7.08	3.46			-8.9	4.4	50.4	69.9
2	L-mom-k	0.15	7.27	3.39	53.63	10.1	-17.9	7.6	51.9	72.8
2	L-mom	0.09	7.82	3.18	78.75	10.5	-23.7	8.4	51.3	69.0
2	MLE	0.08	7.86	3.17			-24.4	8.6	51.3	68.6
3	L-mom-k	0.15	7.54	3.06	69.25	2.5*	-22.6	8.9	51.3	73.0
3	L-mom	0.17	7.40	3.11	75.08	0.3*	-21.2	8.7	51.4	74.0
3	MLE	0.12	7.73	3.00			-26.6	9.4	50.8	70.6
4	L-mom-k	0.15	7.81	2.72	58.50	6.6**	-40.8	13.6	50.5	72.9
4	L-mom	0.13	8.03	2.66	68.95	3.6*	-43.2	14.4	50.3	71.4
4	MLE	0.11	8.08	2.65			-44.9	15.0	50.0	70.4
8	L-mom-k	0.15	9.17	3.53	96.00	40.3	-16.6	11.8	66.7	93.1
8	L-mom	0.14	9.30	3.49	96.61	39.6	-18.0	11.8	66.6	92.2
8	MLE	0.17	8.95	3.61			-13.8	11.7	67.0	94.9
9	L-mom-k	0.15	9.96	3.38	51.65	20.3	-61.4	21.8	71.0	99.7
9	L-mom	0.17	9.68	3.47	44.80	17.4	-58.4	21.1	71.2	101.7
9	MLE	0.11	10.35	3.27			-67.8	23.1	70.2	95.7
10	L-mom-k	0.15	8.74	3.51	89.40	21.9	-45.3	18.8	63.4	88.6
10	L-mom	0.12	9.04	3.41	95.97	25.7	-48.3	19.5	63.2	86.6
10	MLE	0.10	9.17	3.37			-51.1	20.1	62.8	84.9
11	L-mom-k	0.15	8.85	2.96	92.52	72.1	-24.5	13.4	59.3	84.8
11	L-mom	0.05	9.89	2.69	99.98	84.0	-35.5	15.8	58.3	77.8
11	MLE	0.07	9.72	2.73			-32.4	15.2	58.7	79.5
12	L-mom-k	0.15	8.71	2.91	88.47	23.5	-33.0	13.3	57.9	83.1
12	L-mom	0.08	9.47	2.71	99.65	57.6	-41.1	15.1	57.2	77.8
12	MLE	0.09	9.36	2.74			-39.8	14.8	57.4	78.6
19	L-mom-k	0.15	13.76	3.15	74.89	17.5	-6.7	8.1	94.8	134.5
19	L-mom	0.05	15.30	2.87	98.13	8.1**	-23.0	7.7	93.2	123.9
19	MLE	0.02	15.44	2.87			-34.6	11.5	91.2	117.9
20	L-mom-k	0.15	16.02	3.29	58.63	25.3	13.3	4.6	112.6	158.8
20	L-mom	0.02	18.45	2.91	98.15	25.7	-12.7	5.8	110.0	142.3
20	MLE	-0.01	18.31	2.95			-23.5	7.8	107.9	137.0
21	L-mom-k	0.15	17.30	2.90	80.04	42.9	-111.8	37.3	114.9	164.8
21	L-mom	0.24	15.49	3.20	67.76	64.3	-93.3	32.6	116.1	177.9
21	MLE	0.16	16.73	3.00			-115.4	38.5	113.8	163.8
22	L-mom-k	0.15	17.73	2.52	97.28	51.1	-1.1	4.9	110.9	162.1
22	L-mom	0.16	17.42	2.55	98.31	48.2	2.1	5.7	111.2	164.3
22	MLE	0.06	17.84	2.56			-44.7	14.9	103.3	139.2

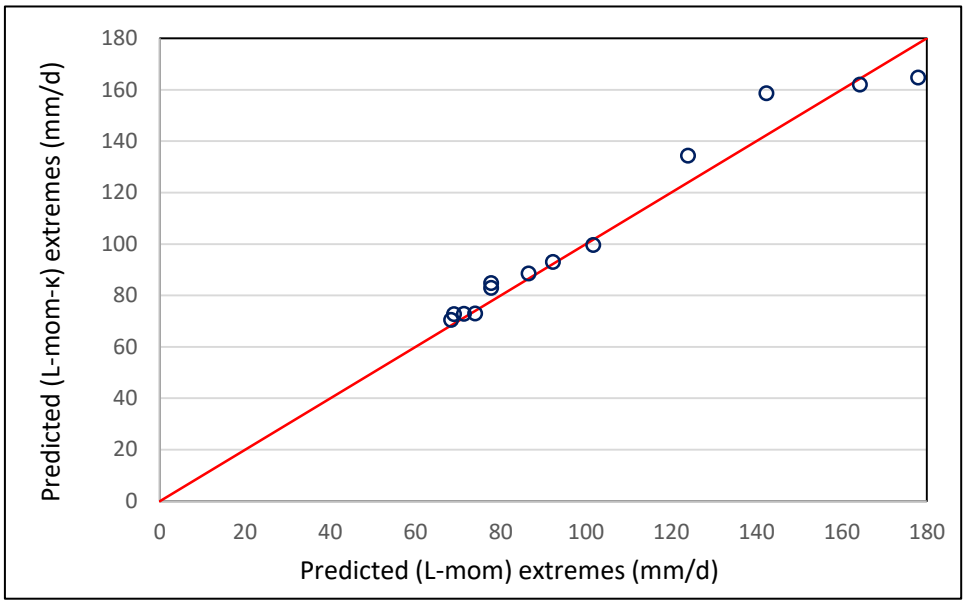

Figure 12: Predicted extremes for return period T=100 years for the modeled data of 80 years of 13 grid cells using L-moments method vs L-moments with specified κ=0.15

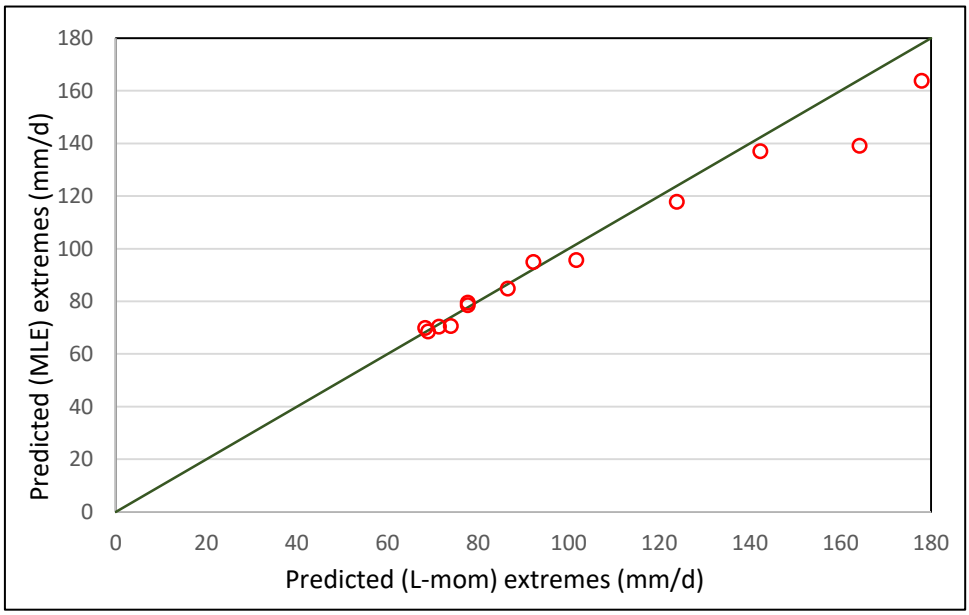

Figure 13: Predicted extremes for return period T=100 years for the modeled data of 80 years of 13 grid cells using L-moments method vs maximum likelihood method (MLE)

4.3. Comparison of the extreme values with rainfall data used by the WDD for flood maps

As explained in Section 2.2, for the flood modeling of the Flood Directive the Water Development Department used the rainfall extremes for return periods of 20 and 100 years for several rain stations in Cyprus, from the report of the Meteorology Department (Pashiardis, 2009). We used the data for four stations in Limassol, 311, 313, 394 and 400 and compared them with our data of nearby grid cells of this project with similar elevation (Table 10). The duration of the rainfall extremes from the report is 24 hours and our data are daily. Thus, we expect the observed and modeled grid cell extremes to be smaller, because of the larger area (12-km) and daily data instead of 24-hour data. From table 10 we notice that even if we don't compare exactly the same quantities with the same durations, indeed the data of the most of our grid cells are smaller than the WDD data. Only extreme values of grid cell 9 and 10 for both return periods are larger compared to station 311 for the observed data and extreme value of grid cell 9 for T=100 years is larger than the one of station 311 for the modeled future daily data. In the case of our modeled future data being higher than their report, that means that the modeled flood will be larger.

Table 10: Elevation of rain stations and 24-hour extreme rainfall for the past observed data (1970-2008) used by WDD and elevation, annual daily extreme rainfall for return period of 20 and 100 years for the past observed and modeled data (1981-2010) and for the future modeled data (2021-2100) of the selected grid cells. For location of stations see table 3 and for location of grid cells see table 4 and figure 5.

Station	**394**	**313**	**311**		**400**	
Elevation (m above sea level)	8	220	340		730	
1970-2008 (observed)						
T=20, 24-hour extreme (mm)	79.2	81.6	72		132.0	
T=100, 24-hour extreme (mm)	117.6	103.2	88.8		218.4	
Grid cell	**4**	**3**	**9**	**10**	**21**	**22**
Elevation (m above sea level)	174	205	572	528	813	693
1981-2010 (observed)						
T=20, daily extreme (mm)	73.7	70.0	96.4	101.2	121.3	124.5
T=100, daily extreme (mm)	101.8	95.2	135.0	144.9	169.8	175.4
1981-2010 (modeled)						
T=20, daily extreme (mm)	43.3	46.5	59.9	57.3	109.9	105.0
T=100, daily extreme (mm)	60.0	65.2	79.0	77.2	155.3	154.5
2021-2100 (modeled)						
T=20, daily extreme (mm)	50.5	51.3	71.0	63.4	114.9	111.0
T=100, daily extreme (mm)	73.0	73.0	99.7	88.6	164.8	162.1

5. Conclusion

We analyzed modeled and observed annual daily rainfall extremes for 30 years' time period, from 1980 to 2010 and compared several methods to derive the parameters (moments) of the GEV-Max distribution in order to find which method gives the best fitting distribution for our data. The L-moments method with specified κ (0.15) is a good method that describes our data and gives a good fit of the distribution.

From the analysis of this data, 1980 to 2010, we compared observed with modeled annual daily rainfall extremes. We concluded that although we found a method that fits out data in the best possible way, L-moments method with specified κ (0.15) still cannot create a perfect fit. The fitted modeled and observed data are not representative of each other because they have many differences between their distribution parameters. Also, for the predicted 100-year extremes in most cases observed data are higher than the modeled.

By studying the climate change with modeled rainfall data, we notice that the first 30 years of the future (2021-2050) are more extreme than the past, as for the return periods of 31 and 100 years, empirical and predicted extremes for 2021-2050 are higher than for 1981-2010.

Also, we conclude that using 60 years' moving time periods is not enough to understand the extremes and represent the future because we cannot notice the change in extremes in all the moving windows, which may lead to wrong conclusions.

Last, by comparing the modeled data series of the future (2021-2100) with the modeled data series of the past we conclude that the 80 years data have a better fit than the 30 years (1981-2010), as one would expect, because the data set is larger. Also, we find that for the 100 year return period, the predicted extreme value of the 80-year data is higher than the predicted 100 year extreme of the past 30-year data.

Because for the past we didn't find a good fit, more research is necessary. Higher resolution modeling is needed to capture the extremes better because it seems that the 12-km modeled data are not good enough to represent extremes. However, we still notice change in extremes for the future compared with the past, as higher extremes appear.

References

Beranová R. & J. Kyselý & M. Hanel (2018). Characteristics of sub-daily precipitation extremes in observed data and regional climate model simulations. Theor. Appl. Climatol. 132:515 – 527.

Bruyère, C.L., J.M. Done G.J. Holland, S. Fredrick. 2014. Bias Corrections of Global Models for Regional Climate Simulations of High-Impact Weather. Clim Dyn, 43, 1847 – 1856.

Bruyère, C.L., A.J. Monaghan, D.F. Steinhoff, D. Yates. 2015. Bias-Corrected CMIP5 CESM Data in WRF/MPAS Intermediate File Format. TN-515+STR, NCAR, pp. 27.

Camera C., A. Bruggeman, P. Hadjinicolaou, S. Pashiardis, and M. A. Lange (2014). Evaluation of interpolation techniques for the creation of gridded daily precipitation (1×1 km^2); Cyprus, 1980–2010, J. Geophys. Res. Atmos., 119, 693 – 712.

Coles S. G., M. J. Dixon (1999). Likelihood-Based Inference for Extreme Value Models. Extremes 2:1, 5 – 23.

Cooke R. A., S. Mostaghimi, F. E. Woeste (1993). VTFIT: A Microcomputer-Based Routine For Fitting Probability Distribution Functions To Data. Vol. 9(4): 401-408.

Gao M., D. Mo, X. Wu (2016). Nonstationary modeling of extreme precipitation in China. Atmospheric Research 182, 1 – 9.

Jonkman S. N. (2005). Global Perspectives on Loss of Human Life Caused by Floods. Natural Hazards (2005) 34: 151 – 175.

Kharin V. V., F. W. Zwiers (2000). Changes in the Extremes in an Ensemble of Transient Climate Simulations with a Coupled Atmosphere– Ocean GCM. Vol. 13, 3760 – 3788.

Kharin V.V., F. W. Zwiers (2005). Estimating Extremes in Transient Climate Change Simulations. Vol. 18, 1156 – 1173.

Kharin V. V., F. W. Zwiers (2007). Changes in Temperature and Precipitation Extremes in the IPCC Ensemble of Global Coupled Model Simulations. Vol. 20, 1419 – 1444.

Kharin V. V., F. W. Zwiers, X. Zhang, M. Wehner (2013). Changes in temperature and precipitation extremes in the CMIP5 ensemble.Climatic Change 119: 345 – 357.

Koutsoyiannis D. (2004). Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation. 591 – 610.

Koutsoyiannis D. (2004). Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records. 575 – 590.

Kozanis S., A. Christofides, N. Mamassis, A. Efstratiadis and D. Koutsoyiannis (2005). Hydrognomon: A hydrological data management and processing software tool. Vol. 7, 04644 (https://www.itia.ntua.gr).

Martins E. S., Jery R. Stedinger (2000). Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. Vol. 36, No. 3, 737 – 744.

Makkonen L. (2005). Plotting Positions in Extreme Value Analysis. Volume 45, 334 – 340.

M.C. Llasat, M. Llasat-Botija, M. A. Prat, F. Porcu, C. Price, A.Mugnai, K.] Lagouvardos, V. Kotroni, D. Katsanos, S. Michaelides, Y. Yair, K. Savvidou, and K. Nicolaides (2010). High-impact floods and flash floods in Mediterranean countries: the FLASH preliminary database. Adv. Geosci., 23, 47 – 55, 2010.

M.L. Anderson, M. ASCE, Z.-Q. Chen, M. ASCE, M. L. Kavvas, M. ASCE, and Arlen Feldman, M. ASCE (2002). Coupling HEC-HMS with Atmospheric Models for Prediction of Watershed Runoff. J. Hydrol. Eng., 2002, 7(4): 312 – 318.

M.R. Knebl, Z.-L. Yanga, K. Hutchison, D.R. Maidment (2005). Regional scale flood modeling using NEXRAD rainfall, GIS, and HEC-HMS/RAS: a case study for the San Antonio River Basin, Summer 2002 storm event. Journal of Environmental Management 75 (2005) 325 – 336.

Pashiardis, Department of Meteorology, Ministry of agriculture, Nicosia (2009). Series of meteorological notes No. 15, Setting up rain curves in Cyprus, Appendix 3.

Papalexiou S. M., D. Koutsoyiannis (2013). Battle of extreme value distributions: A global survey on extreme daily rainfall. VOL. 49, 187 – 201.

Salas J. D., M.ASCE, J. Obeysekera, M.ASCE (2014). Revisiting the Concepts of Return Period and Risk for Nonstationary Hydrologic Extreme Events. J. Hydrol. Eng., 2014, 19(3): 554 – 568.

Tsakiris G., I. Nalbantis, A. Pistrika (2009). Critical Technical Issues on the EU Flood Directive. European Water 25/26: 39 – 51.

Water Development Department (WDD) (2016). Program of Measures of Flood Risk Management Plan for Cyprus river basin (Period 2016 – 2021).

Water Development Department (WDD) Prisma Consulting Engineers S.A. – T.C. Geomatic Ltd – Ofek Aerial Photography (1987) Ltd (Nicosia, 2014). Report of hydrological model of the old and new beds of Garyllis river. Ministry of agriculture, Nicosia, Cyprus (in Greek).

Zittis G., A. Bruggeman, C. Camera, P. Hadjinicolaou, J. Lelieveld (2017). The added value of convection permitting simulations of extreme precipitation events over the eastern Mediterranean. Atmospheric Research 191 (2017) 20 – 33.

Appendix A: Equations for L-moments and Maximum Likelihood Estimation Methods

L moments estimator

The first three L moments of a random variable X are defined in terms of its distribution function F(x) as follows (Kharin et al. 2000):

$$\lambda_1 = EX = \int_0^1 x(F)dF\,,$$

$$\lambda_2 = \frac{1}{2}E(X_{2:2} - X_{1:2}) = \int_0^1 x(F)(2F-1)dF\,,$$

$$\lambda_3 = \frac{1}{3}E(X_{3:3} - 2X_{2:3} + X_{1:2}) = \int_0^1 x(F)(6F^2 - 6F + 1)dF$$

where

$\{X_{1:r}, X_{2:r}, \ldots, X_{r:r}\}$ are the order statistics obtained by sorting the sample of size r in ascending order,

L moments λ_1, λ_2, and λ_3 are in some ways analogous to the conventional central moments and are regarded as measures of location, shape, and scale parameters of the distribution.

Maximum likelihood estimator

The log-likelihood function for θ based on data x is given by (Kharin et al. 2005):

$$l_{x1,\ldots,x2}(\theta) = \sum_{i=1}^{n} \ln fx(x_i;\theta)$$

where

$fx(x;\theta)$ is the probability density function of a random variable X with parameters $\theta = \{\theta_1, \ldots, \theta_p\}$.

$x = \{x_i, i=1,\ldots,n\}$ is the n independent realizations of the random variable X.

The log-likelihood function for the GEV distribution with constant location, scale, and shape parameters ψ, λ and κ is given by (Kharin et al. 2005):

$$l_{x1,\ldots,x2}(\psi,\lambda,\kappa) = \sum_{i=1}^{n}\left\{-\ln\lambda - (1-\kappa)yi - e^{-yi}\right\}$$

$$y_i = \begin{cases} -\frac{1}{\kappa}\ln\left(1-\kappa\frac{x_i-\psi}{\lambda}\right), & \kappa \neq 0, \\ \frac{x_i-\psi}{\lambda}, & \kappa = 0 \end{cases}$$

Appendix B: Tables for fitted Extreme Value Distributions for 1981-2010 data

Table A1: First (kappa), second (lambda) and third (psi) moments of GEV distribution using L-moments method, without κ specified for the 30 years annual daily rainfall extremes for modeled and observed data of 13 grid cells; p-values of Kolmogorov-Smirnov (KS) and Chi-square (x^2) tests; mean absolute error (MAE) and bias of predicted versus empirical extremes (T=10.33, 15.50, 31 years); and predicted rainfall extremes for return periods (T) of 20, 50 and 100 year.

Grids	Data	kappa	lambda	psi	KS-p (%)	x^2-p (%)	Bias	MAE	Predicted values		
									T=20	T=50	T=100
1	Mod	-0.56	8.70	3.38	99.9	41.4	0.3	1.2	42.0	43.1	43.7
1	Obs	0.145	7.29	4.78	98.7	19.7	-6.4	2.1	61.9	73.1	82.5
2	Mod	-0.18	9.03	2.79	96.6	6.8	-7.2	2.4	46.0	50.6	53.5
2	Obs	0.12	7.94	4.68	100.0	19.7	-15.0	5.0	65.4	76.5	85.7
3	Mod	-0.15	8.58	2.72	91.5	3.7	-7.7	2.6	43.8	48.6	51.7
3	Obs	-0.03	10.57	3.59	100.0	1.0	-13.1	4.4	67.9	76.8	83.2
4	Mod	-0.22	8.05	2.82	91.5	24.8	-4.1	1.4	40.2	43.8	46.0
4	Obs	-0.03	11.77	3.22	99.6	19.7	-5.2	4.1	71.4	81.3	88.6
8	Mod	-0.06	9.88	3.49	97.6	41.4	-8.3	2.9	61.5	69.0	74.4
8	Obs	-0.106	18.37	2.70	98.1	31.7	-0.7	3.1	96.3	108.2	116.4
9	Mod	-0.16	8.81	4.12	95.8	41.4	-3.6	1.2	57.1	61.8	65.0
9	Obs	-0.08	16.76	2.86	99.7	56.4	-8.4	2.8	92.4	104.3	112.7
10	Mod	-0.38	10.32	3.25	93.8	2.5	-5.4	1.8	52.2	54.9	56.4
10	Obs	0.11	15.83	2.83	99.6	19.7	-19.0	9.0	100.6	122.3	140.1
11	Mod	-0.026	9.85	2.88	98.6	24.8	-0.9	2.1	56.5	64.9	71.0
11	Obs	-0.03	16.83	2.58	86.1	12.7	4.6	5.2	91.4	105.6	115.9
12	Mod	0.1	9.88	2.68	100.0	24.8	-14.4	4.8	60.8	73.8	84.4
12	Obs	0.0016	19.75	2.32	98.4	56.4	-14.1	5.6	104.6	123.1	137.0
19	Mod	0.14	11.67	3.94	71.6	0.7	-48.0	16.0	88.9	106.4	121.0
19	Obs	-0.15	17.54	3.07	89.3	2.5	-9.7	6.3	95.9	105.7	112.1
20	Mod	0.27	16.94	3.18	80.3	6.8	-47.8	15.9	130.7	170.3	207.2
20	Obs	-0.15	27.22	2.83	99.4	100.0	-8.9	3.7	142.2	157.4	167.4
21	Mod	-0.039	19.12	2.74	87.7	10.3	3.0	2.1	106.1	121.6	133.0
21	Obs	-0.18	22.73	2.69	99.4	19.7	0.0	3.6	113.6	125.0	132.4
22	Mod	0.33	13.37	2.97	87.3	12.7	-82.6	27.5	106.7	145.0	182.4
22	Obs	-0.20	24.17	2.56	96.5	4.6	8.2	4.1	115.9	127.1	134.3

Table A2: First (kappa), second (lambda) and third (psi) parameters of GEV distribution using Maximum likelihood method for the 30 years annual daily rainfall extremes for modeled and observed data of 13 grid cells; mean absolute error (MAE) and bias of predicted versus empirical extremes (T=10.33, 15.50, 31 years); and predicted rainfall extremes for return periods (T) of 20, 50 and 100 year.

Grids	Data	kappa	lambda	psi	Bias	MAE	Predicted values			Notes
							T=20	T=50	T=100	
1	Mod	-0.94	4.17	2.61	-79.6	26.5	15.0	15.2	15.2	*
1	Obs	0.013	7.37	4.82	-17.3	5.8	57.8	65.0	70.4	
2	Mod									**
2	Obs	0.025	8.27	4.55	-20.9	7.0	63.2	71.6	78.0	
3	Mod									**
3	Obs	-0.028	10.14	3.76	-15.7	5.2	67.0	75.6	81.9	
4	Mod									**
4	Obs	-0.049	11.20	3.42	-11.2	4.3	69.2	78.1	84.4	
8	Mod	-0.06	9.49	3.65	-10.8	3.6	60.6	67.9	73.1	
8	Obs									**
9	Mod	-0.19	8.57	4.26	-6.8	2.3	55.9	60.0	62.6	
9	Obs	-0.11	16.29	2.98	-15.9	5.3	89.6	99.9	107.0	
10	Mod									**
10	Obs	-0.016	16.10	2.87	-39.6	13.2	93.0	107.1	117.6	
11	Mod	-0.05	9.39	3.05	-5.9	2.0	54.7	62.1	67.5	
11	Obs	-0.07	16.06	2.75	-6.3	5.0	87.4	99.4	107.8	
12	Mod	0.0088	9.97	2.73	-23.9	8.0	57.2	66.8	74.0	
12	Obs	-0.034	18.97	2.45	-26.3	8.7	100.2	115.9	127.3	
19	Mod	0.075	12.49	3.71	-50.2	16.7	88.0	103.1	115.1	
19	Obs	-0.31	15	3.55	-47.8	15.9	82.4	87.2	90.0	***
20	Mod	0.14	16.73	3.31	-86.6	28.9	116.5	141.3	162.0	
20	Obs									**
21	Mod	-0.066	18.09	2.93	-8.6	2.9	101.9	115.3	124.8	
21	Obs	-0.42	14.31	3.83	-98.5	32.8	79.1	82.3	84.0	****
22	Mod	0.29	13.91	2.89	-85.4	28.5	105.6	140.5	173.6	
22	Obs									**

*29 values are discarded
** Floating point overflow
***2 values are discarded
****8 values are discarded

Table A3: First and second moment (lambda, psi) of GEV distribution using L-moments method with specified $\kappa = 0{,}10$ for the 30 years annual daily rainfall extremes for modeled data of 13 grid cells; p-values of Kolmogorov-Smirnov (KS) and Chi-square (x^2) tests; bias and mean absolute error (MAE) of predicted versus empirical extremes (T=10.33, 15.50, 31 years); and predicted rainfall extremes for return periods (T) of 20, 50 and 100 year.

Grids	kappa	lambda	psi	KS-p (%)	x^2-p (%)	Bias	MAE	Predicted values		
								T=20	T=50	T=100
1	**0.1**	5.79	4.68	13.0	5.9	13.5	4.5	47.2	54.8	61.0
2	**0.1**	7.09	3.41	45.3	13.53	-0.6	0.6	48.7	58.02	65.6
3	**0.1**	6.85	3.28	43.5	11.5	-1.9	2.7	46.2	55.2	62.5
4	**0.1**	6.16	3.51	26.7	5.0	2.6	2.0	42.9	51.0	57.6
8	**0.1**	8.49	3.99	86.4	60.7	-4.1	1.5	63.2	74.3	83.4
9	**0.1**	7.01	5.04	78.2	60.7	2.5	1.5	59.5	68.7	76.2
10	**0.1**	7.32	4.31	14.5	2.6	6.5	2.7	56.9	66.5	74.3
11	**0.1**	8.69	3.20	94.1	18.9	2.5	1.4	57.9	69.3	78.6
12	**0.1**	9.89	2.68	99.9	51.3	-14.4	4.8	60.8	73.8	84.3
19	**0.1**	12.19	3.79	78.3	2.6	-49.1	16.4	88.4	104.4	117.4
20	**0.1**	21.00	2.63	92.9	22.3	-54.1	18.0	127.9	155.5	177.9
21	**0.1**	16.69	3.07	79.4	31.1	10.2	4.6	109.0	130.9	148.8
22	**0.1**	18.18	2.27	82.6	43.5	-88.2	29.4	104.1	128.0	147.4