Abstract
The objective on this study is to fully statistically determine and describe the rainfall regime characteristics of the very long precipitation records of 124-year period (1892–2015) in Thessaloniki and moreover to apply extreme value statistical methodologies to extreme daily rainfall information (period, 1931-2015), in order to define the extreme precipitation probability distributions. To meet these objectives, the observed annual, monthly and daily rainfall measurements obtained at the Meteorological Station of the Department of Meteorology and Climatology of the Aristotle University of Thessaloniki are used. The use of both Mann-Kendall test and moving average methods are used to determine the profile of the examined very long-period rainfall regime in Thessaloniki, on annual, seasonal, and monthly basis. Taking into consideration the climatic change effect upon precipitation regimes, the extreme daily rainfall measurements are statistically analyzed with extreme value statistical methodologies, defining thus the extreme precipitation probability distribution functions. The Taylor diagram analysis, through combinations of appropriate correlation coefficients, standard deviations, and root-mean-square differences, are constructed to statistically quantify the degrees of similarity between reference points (mean annual precipitation) and results on seasonal and agricultural (agro-) seasonal basis. All the above tests are applied to detect possible changes in precipitation characteristics, over the study region, throughout the extreme long period of 124 years. The resulted probability distributions might potentially contribute to the prediction of extreme rainfall events, and be used for future climatic change projections. Moreover, based upon the resulted-adopted probability distributions, emerging from this extreme long study period, return periods for different extreme values of precipitation observed are calculated.

Keywords
Extreme precipitation; statistical analysis; Taylor diagrams; Mann-Kendall test; extreme probability distributions; Thessaloniki (Greece)

Manuscript Details
Manuscript number
ATMOSRES_2017_477
Title
STUDY OF VERY LONG-PERIOD EXTREME PRECIPITATION RECORDS IN THESSALONIKI, GREECE
Article type
Research Paper
Abstract
The objective on this study is to fully statistically determine and describe the rainfall regime characteristics of the very long precipitation records of 124-year period (1892–2015) in Thessaloniki and moreover to apply extreme value statistical methodologies to extreme daily rainfall information (period, 1931-2015), in order to define the extreme precipitation probability distributions. To meet these objectives, the observed annual, monthly and daily rainfall measurements obtained at the Meteorological Station of the Department of Meteorology and Climatology of the Aristotle University of Thessaloniki are used. The use of both Mann-Kendall test and moving average methods are used to determine the profile of the examined very long-period rainfall regime in Thessaloniki, on annual, seasonal, and monthly basis. Taking into consideration the climatic change effect upon precipitation regimes, the extreme daily rainfall measurements are statistically analyzed with extreme value statistical methodologies, defining thus the extreme precipitation probability distribution functions. The Taylor diagram analysis, through combinations of appropriate correlation coefficients, standard deviations, and root-mean-square differences, are constructed to statistically quantify the degrees of similarity between reference points (mean annual precipitation) and results on seasonal and agricultural (agro-) seasonal basis. All the above tests are applied to detect possible changes in precipitation characteristics, over the study region, throughout the extreme long period of 124 years. The resulted probability distributions might potentially contribute to the prediction of extreme rainfall events, and be used for future climatic change projections. Moreover, based upon the resulted-adopted probability distributions, emerging from this extreme long study period, return periods for different extreme values of precipitation observed are calculated.

Keywords
Extreme precipitation; statistical analysis; Taylor diagrams; Mann-Kendall test; extreme probability distributions; Thessaloniki (Greece)
Manuscript category
Climatology and climate change
Corresponding Author
Nikoletta Pakalidou
Institution
The University of Manchester, School of Chemical Engineering and Analytical Science
Order of Authors
Nikoletta Pakalidou, Petroula Karacosta
Suggested reviewers
Constantinos Theodoropoulos, Ewa Bednorz, Vlado Spiridonov, Elsa Aristodemou, Eftychia (Efi) Rousi

Submission Files Included in this PDF
File Name [File Type]
Cover_letter.docx [Cover Letter]
Point-by-pointResponse.docx [Response to Reviewers]
Highlights.docx [Highlights]
Pakalidou_Karacosta.doc [Manuscript File]
To view all the submission files, including those not included in the PDF, click on the manuscript title on your EVISE Homepage, then click 'Download zip file'.
Re: Submission of research paper for the Special Issue MichaelidesMed

Dear Prof. Sanchez,
I hereby attach the research paper entitled "Study of very long-period extreme precipitation records in Thessaloniki, Greece." by Nikoletta Pakalidou and Petroula Karacosta, which is submitted for the above Special Issue of Atmospheric Research. The names of the proposed reviewers are as follows:

1) Professor Dr. Hab. Ewa Bednorz
   Adam Mickiewicz University in Poznań, Poland
   ewabedno@amu.edu.pl

2) Efi Rousi, Potsdam Institute for Climate Impact Research
   Germany
   rousi@pik-potsdam.de

3) Vlado Spiridonov, SS Cyril and Methodius University
   Skopje, FYROM
   vspiridonov@meteo.gov.mk

4) Professor Theodoropoulos Constantinos,
   University of Manchester
   k.theodoropoulos@manchester.ac.uk

5) Elsa Aristodemou,
   London South Bank University
   aristode@lsbu.ac.uk

Also, I attach a letter as a point-by-point response to the reviewers, in order to check the corrections have been done.

Looking forward to hearing from you,

Yours faithfully,

Dr. Nikoletta Pakalidou
University of Manchester
Response to Reviewers

Manuscript number ATMOSRES_2017_477

Title: Study of very long-period extreme precipitation records in Thessaloniki, Greece.

Authors: Pakalidou N. and Karacosta P.

Authors’ point-by-point Response to Reviewers’ comments

Both authors would like to thank the reviewers for their comments in order to improve the quality of the manuscript. All comments and suggestions have been adopted. The corrections have been done in the manuscript. About the point-by-point response to reviewers, green and blue colors are used for the reviewers’ comments and authors’ responses, respectively.

Reviewer 1:

Comment #1:
It is desirable to directly connect and interrelate the theoretical background information (equations, etc.) presented in paragraph 2.2, with the calculated values and results derived and presented through these equations (reference the eq. #).

Authors’ response:
We agree with the reviewer and the comment is adopted. There are 19 equations in paragraph 2.2, and we use them in order to statistically study the precipitation records. We connect and interrelate these equations with the results derived. We do that by using these equations as reference, i.e. Equation #. Hence, the text on page 11 appears as follows:

The annual M-K test was found with a positive trend (Z=0.376) by applying Equation 8 (Table 1).

Table 1. Results from statistical analysis on annual, seasonal, agro-seasonal, and monthly basis, for the period 1892-2015, in Thessaloniki. Equations 1 and 5-9 are used to compute correlation coefficient and Mann-Kendall parameters. Other values are derived by applying basic statistical analysis.

Moreover, the four last lines on the first paragraph on page 13 appear as follows:

Further, the M-K test (Z-value) applied in monthly precipitation records, showed negative trends for about half of the months (Equation 8). As it can be seen in Table 1, the highest positive slope had the magnitude of +1.369 in February and the lowest negative one was -1.611 in June.
(b) **Seasonal trends**

Concentrated to the maximum precipitation heights observed during the whole period (1892-2015) in normal seasons, appear to be similar for all seasons, ranging from 250 mm to 320 mm. In Figure 3, one can notice the maximum precipitation height observed for the normal seasonal and annual precipitation conditions in the study area. Autumn has the highest maximum amount of precipitation (312.4 mm in 1896) and summer the lowest maximum (245.8 mm in 2014), during the 124-year period (Figure 3). While, 912.2 mm were measured for the extreme precipitation year of 2014, and only 200.8 mm for the year 2000. Winter precipitation records showed a maximum in 1935 (285.6 mm) and a minimum in 1991 (6.8 mm) during the period 1892-2015. Nevertheless, the lowest minimum corresponded to summer in 1927 (3.4 mm). The seasonal precipitation trends, according to the M-K test (Table 1), appear to be positive, except of the summer, which indicates negative value (Equation 8). As one would expect, autumn presented the highest precipitation mean (~133 mm), while the lowest noted in summer (~82 mm), which is well understood, as it is a basic characteristic of the Mediterranean climatic type (Feidas et al., 2007).

Also, paragraph (c) on page 14 appears as follows:

(c) **Agro-seasonal trends**

Positive trends were noted in *agro*-summer and *agro*-winter, whereas the only negative trend was found in *agro*-autumn (Table 1). *Agro*-spring was found with Z-value in M-K test equals to zero (Equation 8). Further, the lowest precipitation value, among the *agro*-seasons, belonged to *agro*-spring, whereas the maximum value was observed in *agro*-winter. It is therefore obvious that the relocation of March and September into *agro*-winter and *agro*-summer respectively, causes a significant difference between normal seasons and *agro*-seasons. For example, during autumn, an increasing trend is evident. However, relocating September to *agro*-summer, leaves *agro*-autumn with the months of October and November, resulting to a decreasing trend.

The last three line on the first paragraph on page 15 are rewritten as follows:

To further investigate the results from the above-mentioned diagrams, the values of the correlation coefficient of each period with the mean annual rainfall were calculated (Equations 1, 2, and 4). The results are discussed in the following paragraph.

Caption of Figure 4 on page 16 is changed to:
**Fig. 4.** Taylor diagrams represent the degree of contribution of each observation to the reference point (mean annual precipitation record during the period 1892-2015). Equations 1, 2, and 4 are used to design these diagrams. (a-e) normal seasons, (f) agro-seasons; (a) 1981-2000 less-rain period, (b) 1961-1980 more-rain period, (c) 1892-2015 less-rain period, (d) 1892-2015 more-rain period, (e) 1892-2015 period with total data in normal seasons, (f) 1892-2015 period with total data in agro-seasons.

First paragraph of section 3.2 appears as follows:

In the field of management of extreme events, observing the past in order to predict the future, is of paramount importance. To this end, daily observations of the sub-period 1932-2015 were used in order to determine return periods of the extreme events, which correspond to specific distribution functions. The filter of the threshold of R30 mm was applied and 140 extreme precipitation cases were counted (Figure 5). These 140 cases of extreme daily precipitation events were used as input information in search for the best fitted probability distribution. Sixty statistical distributions were examined for this purpose. The GOF distribution, through the Kolmogorov-Smirnov (Equation 10), Anderson-Darling (Equation 11) and Chi-Squared tests (Equations 12 and 13), implied that the Generalized Pareto distribution was the best fitted one. This conclusion can be drawn by examining the statistics of the three final distributions (Equations 14-19) according to the applied tests (Table 2). Figure 6 depicts the best fitted probability density functions, based on the 140 cases, for the sub-period 1931-2015, in Thessaloniki. These are: the Generalized Pareto for the Kolmogorov-Smirnov test, the Johnson SB for the Anderson-Darling, and the Generalized Extreme Value for the Chi-Squared.

Finally, caption of Table 2 is changed as:

**Table 2.** Extreme daily precipitation records (140 cases) with the threshold of $R_{30}$, during the sub-period 1931-2015, in Thessaloniki. The goodness-of-fit distribution and the K-S, A-D, and $X^2$ tests concluded that the Generalized Pareto distribution is the best fitted one. Equations 10-13 are used to apply the tests and Equations 14-19 are used to calculate all distributions’ parameters.

**Comment #2:**

Figures 4 and 6 need to be improved on the inside written information and the numbers in the abscissa and ordinate.

**Authors’ response:**

Indeed, the colors and the size font of the numbers were not very clear in both Figures 4 and 6. We changed the colors and the font size of the texts in diagrams in order to easy be read by anyone. Hence, Figure 4 (Taylor diagrams) and Figure 6 (distributions) appear as follow:
Comment #3:

It would be very supportive and useful, and at the same time would add value to the manuscript, if the authors would try to relate, support and compare their findings, results and conclusions, with previous research and relative references.

Authors’ response:

We agree with the reviewer. It is important to relate the results derived in this work with previous research and relative references. Hence, the first paragraph in section 3.1.1. on page 10 is changed as follows:

The annual precipitation time series was analyzed in Figure 1, presenting the precipitation height for the very long period of 1892-2015. The extremist high and low precipitation years are identified and depicted through the examined period. It is worth mentioning the extremist rainfall measurement of 912.2 mm obtained in the wettest year of 2014 and the second extreme observed in 1919 with 722.1 mm. It is significant to notice that 133 raindays were counted in 2014 (Tolika et al., 2017) and it is possible this phenomenon belonged to a larger scale over the Balkan-Peninsula region (Tošić et al., 2017). On the other hand, it should be noticed that three years can be characterized as driest years (1932, 1977 and 2000).
Also, the last seven lines in the last paragraph in section 3.1.1. on page 12 are changed as follows:

The last could be well understood by observing in the annual precipitation time series the two maxima in 2014 (~900 mm) and in 2011 (~550 mm). It worth to remark that Feidas et al. (2007) studied the precipitation records in Greece during 1955-2001 using surface and satellite data, and they concluded to a downward trend in both winter and annual records. This is well understood as in Figure 2 one can observe the high positive percentage during the period 2011-2015. The last seems to change the above trends observed by Feidas et al. (2007).

Finally, the second paragraph on page 15 is changed as follows:

The 20-year-less-rain period and 20-year-more-rain period (Figure 4(a) and (b)) were examined. During both periods (less-rain during 1981-2000, more-rain during 1961-1980), autumn contributed more to the reference point (mean annual records), with the highest degree during the more-rain period. Between the less-rain and more-rain periods through the total period of 124 years, winter was found to contribute more to the reference point in the first case (Figure 4 (c)). However, in the Taylor diagram, autumn was presented with more similar characteristics in rainfall profile with the reference point (Figure 4 (d)). These results are in a good agreement with the study of Tolika et al. (2017). Indeed, the authors remarked that autumn contributed by a percentage of 27-28% to the annual precipitation during 1958-2013 in Thessaloniki region. The authors concluded that this percentage was little more (29%) during 2014.

It should be notice that extra references are added as follows:


Reviewer 2:

Comment #1:
I would rather suggest moving Figure 1 with its description (page 3, lines 22-27) to chapter 3, which contain results (to 3.1.2.b exactly), however, I leave it to the Author’s decision.

Authors’ response:
We agree with the reviewer. Initial Figure 1 is moved to page 14 and is renamed to Figure 3. Also, its description (lines 22-27 on page 3) is moved to page 13. The second paragraph on page 13 appears as follows:

(b) Seasonal trends

Concentrated to the maximum precipitation heights observed during the whole period (1892-2015) in normal seasons, appear to be similar for all seasons, ranging from 250 mm to 320 mm. In Figure 3, one can notice the maximum precipitation height observed for the normal seasonal and annual precipitation conditions in the study area. Autumn has the highest maximum amount of precipitation (312.4 mm in 1896) and summer the lowest maximum (245.8 mm in 2014), during the 124-year period (Figure 3). While, 912.2 mm were measured for the extreme precipitation year of 2014, and only 200.8 mm for the year 2000. Winter precipitation records showed a maximum in 1935 (285.6 mm) and a minimum in 1991 (6.8 mm) during the period 1892-2015. Nevertheless, the lowest minimum corresponded to summer in 1927 (3.4 mm). The seasonal precipitation trends, according to the M-K test (Table 1), appear to be positive, except of the summer, which indicates negative value (Equation 8). As one would expect, autumn presented the highest precipitation mean (~133 mm), while the lowest noted in summer (~82 mm), which is well understood, as it is a basic characteristic of the Mediterranean climatic type (Feidas et al., 2007).

Comment #2:
Figure 4 is very appropriate, informative and well stated. The quality of the figure is quite good and I believe acceptable. However, it is necessary to have some improvements: (a) on the inside of the figure written words to become more readable and (b) on the size of the written words and numbers describing both axes (horizontal and vertical).

Authors’ response:
We agree with the reviewer. These corrections on Figure 4 are made in order to be easy readable. To this end, the font size of both numbers and words of the figure and the colors of lines are changed and on page 16 Figure 4 appears as follows:
Comment #3:
Figure 6 needs some improvements on the written words and numbers describing the horizontal and vertical axes.

Authors’ response:
We agree with the reviewer. Some improvements should be done in Figure 6 in order to be more readable. To this end, we change the font size of the numbers/words. On page 19 Figure 6 appears as follow:
Comment #4:

It would be supportive and useful, and at the same time would add some value to the manuscript, if the authors would try to relate, support and compare their findings, results and conclusions, with previous research and relative references. In other words to relate their findings with the existed literature.
**Authors' response:**

We agree with the reviewer that the results of this study should be related with the results of previous studies. To this end, the first paragraph in section 3.1.1. on page 10 is changed as follows:

The annual precipitation time series was analyzed in Figure 1, presenting the precipitation height for the very long period of 1892-2015. The extremist high and low precipitation years are identified and depicted through the examined period. It is worth mentioning the extremist rainfall measurement of 912.2 mm obtained in the wettest year of 2014 and the second extreme observed in 1919 with 722.1 mm. It is significant to notice that 133 raindays were counted in 2014 (Tolika et al., 2017) and it is possible this phenomenon belonged to a larger scale over the Balkan-Peninsula region (Tošić et al., 2017). On the other hand, it should be noticed that three years can be characterized as driest years (1932, 1977 and 2000).

Also, the last seven lines in the last paragraph in section 3.1.1. on page 12 are changed and the paragraph is modified:

The last could be well understood by observing in the annual precipitation time series the two maxima in 2014 (~900 mm) and in 2011 (~550 mm). It worth to remark that Feidas et al. (2007) studied the precipitation records in Greece during 1955-2001 using surface and satellite data, and they concluded to a downward trend in both winter and annual records. This is well understood as in Figure 2 one can observe the high positive percentage during the period 2011-2015. The last seems to change the above trends observed by Feidas et al. (2007).

Finally, the second paragraph on page 15 is changed as follows:

The 20-year-less-rain period and 20-year-more-rain period (Figure 4(a) and (b)) were examined. During both periods (less-rain during 1981-2000, more-rain during 1961-1980), autumn contributed more to the reference point (mean annual records), with the highest degree during the more-rain period. Between the less-rain and more-rain periods through the total period of 124 years, winter was found to contribute more to the reference point in the first case (Figure 4 (c)). However, in the Taylor diagram, autumn was presented with more similar characteristics in rainfall profile with the reference point (Figure 4 (d)). These results are in a good agreement with the study of Tolika et al. (2017). Indeed, the authors remarked that autumn contributed by a percentage of 27-28% to the annual precipitation during 1958-2013 in Thessaloniki region. The authors concluded that this percentage was little more (29%) during 2014.

Hence, extra references are added to the manuscript:
Comment #5:

The mathematical expressions (equations) of the resulted probability distribution functions, that is: the Generalized Pareto, the Jonson SB and the Generalized Extreme Value, whose parameters have been calculated, could be incorporated in the manuscript (or at least the Generalize Pareto which is the best fitted one), integrating thus the whole procedure and adding some extra value to the manuscript.

Authors’ response:

We agree with the reviewer. Hence, the last paragraph of 2.2.6. section is added to a new section (2.2.7.) on page 8 that describes the continuous probability distributions. Generalized Pareto distribution, Jonson SB distribution and Generalized extreme value distribution are analytically defined. The extra 2.2.7. section appears as follows:

2.2.7. Continuous probability distributions

The three tests described above, use the extreme events as input and compare it with the values that are expected from the 60 probability distributions that are available. The best fitted probability distribution is the one with the minimum absolute deviation between the observed and the predicted values.

(a) Generalized Pareto distribution

The Generalized Pareto distribution is characterized by three parameters; the shape parameter \( k \), the location parameter \( \mu \), and the scale parameter, \( \sigma (\sigma > 0) \). The domain where this distribution can be defined is (Pickands III, 1975):

\[
\mu \leq x < +\infty, k \geq 0, \\
\mu \leq x \leq \mu - \frac{\sigma}{k}, k < 0.
\]  

[14]

The probability density function derived by the Generalized Pareto distribution is given as:
\[ f(x) = \begin{cases} \frac{1}{\sigma} (1 + k\frac{x-z}{\sigma})^{-1/\delta}, & k \neq 0 \\ \frac{1}{\sigma} \exp(-z), & k = 0 \end{cases}, \text{where } z = \frac{x-\mu}{\sigma}. \]  
\[ \text{[15]} \]

(b) Johnson SB distribution

The Johnson SB distribution is a four-parameter model; the shape parameters \( \gamma \) and \( \delta \) (\( \delta > 0 \)), the scale parameter \( \lambda \) (\( \lambda > 0 \)), and the location parameter \( \xi \). The domain of this distribution is defined as (Johnson, 1949):

\[ \xi \leq x \leq \xi + \lambda. \]  
\[ \text{[16]} \]

The probability density function derived by the Johnson SB distribution is given as:

\[ f(x) = \frac{\delta}{\lambda \sqrt{2\pi z(1-z)}} \exp\left(-\frac{1}{2} \left( \gamma + \delta \ln \left( \frac{z}{1-z} \right) \right)^2 \right), \text{where } z = \frac{x-\xi}{\lambda}. \]  
\[ \text{[17]} \]

(c) Generalized Extreme Value distribution

The Generalized Extreme Value distribution is a three-parameter model; the shape parameter \( k \), the location parameter \( \mu \), and the scale parameter, \( \sigma \) (\( \sigma > 0 \)). The domain where this distribution can be defined is (Jenkinson, 1955):

\[ 1 + k \frac{(x-\mu)}{\sigma} > 0, k \neq 0, \] 
\[ 1 + k \frac{(x-\mu)}{\sigma} < 0, k = 0. \]  
\[ \text{[18]} \]

The probability density function derived by the Generalized Extreme Value distribution is given as:
\begin{align*}
  f(x) = \begin{cases} 
    \frac{1}{\sigma} \exp\left(- (1 + kz)^{-1/k}\right) (1 + kz)^{-1-1/k}, & k \neq 0 \\
    \frac{1}{\sigma} \exp\left(- z - \exp(z)\right), & k = 0 \\
  \end{cases}, \text{where } z = \frac{x - \mu}{\sigma}.
\end{align*}

Moreover, three extra references are added:


\textbf{Comment #6:}

Although I am not a native English speaker, I feel that improving the language would lead to better reception of the text; maybe the Authors would consider the language-checking by a specialist.

\textbf{Authors' response:}

We agree with the reviewer and we have read very carefully the manuscript. Some details are changed in order to improve the language and the spelling on the text.
Authors’ response:

Indeed, all details presented in the above table should be modified. We changed all the suggested corrections-changes following the step-by-step suggestions of the reviewer. For example, the 7th line in Abstract is changed and so the Abstract appears as follows:

The objective on this study is to fully statistically determine and describe the rainfall regime characteristics of the very long precipitation records of 124-year period (1892–2015) in Thessaloniki and moreover to apply extreme value statistical methodologies to extreme daily rainfall information (period, 1931-2015), in order to define the extreme precipitation probability distributions. To meet these objectives, the observed annual, monthly and daily rainfall measurements obtained at the Meteorological Station of the Department of Meteorology and Climatology of the Aristotle University of Thessaloniki are used. The use of both Mann-Kendall test and moving average methods are used to determine the profile of the examined very long-period rainfall regime in Thessaloniki, on annual, seasonal, and monthly basis. Taking into consideration the climatic change effect upon precipitation regimes, the extreme daily rainfall measurements are statistically analyzed with extreme value statistical methodologies, defining thus the extreme precipitation probability distribution functions. The Taylor diagram analysis, through combinations of appropriate correlation coefficients, standard deviations, and root-mean-square differences, are constructed to statistically quantify the degrees of similarity between reference points (mean annual precipitation) and results on seasonal and agricultural (agro-) seasonal basis. All the above tests are applied to detect possible changes in precipitation characteristics, over the study region, throughout the extreme long period of 124 years. The resulted probability distributions might potentially
contribute to the prediction of extreme rainfall events, and be used for future climatic change projections. Moreover, based upon the resulted-adopted probability distributions, emerging from this extreme long study period, return periods for different extreme values of precipitation observed are calculated.

Also, the text on page 2 is modified according to the above suggestion by the reviewer and so the text on page 2 appears as follows:

A number of studies have described the rainfall pattern in the Mediterranean area; there is an unambiguous negative trend in precipitation’s variability during winter season since the 1960s (Tselioudis et al., 2008; Philandras et al., 2010). This could be explained by the North Atlantic Oscillation (NAO) and by climatic models scenarios from IPCC. Indeed, there is a positive trend in NAO index that links the negative trends of rainfall data (Krichak et al., 2014). Moreover, SRES A1B scenario of IPCC indicates rainfall relocation at north direction that will affect the southern areas in the Mediterranean region decreasing the rainfall records there. The results that refer to the future period 2080-2099 indicate a decreasing by 20% of rainfall in comparison with the observations over the period 1980-1999 (Dünkeloh and Jacob, 2003; Luterbacher et al., 2005; Nastos and Zerefos, 2009; Sindosi et al., 2015; Zittis et al., 2017). On the other hand, heavy-precipitation frequency is predicted to statistically increase over most areas in Mediterranean region (Partal and Kahya, 2006; Ruiz-Leo, 2013).

Greece is located at the southeast part of European Mediterranean area. The geomorphology and geography of the country influence its climatic type, as it can be described as Mediterranean (Csa) according to Köppen classification. Several researchers have studied the temporal or spatial characteristics of precipitation in Greece (Feidas et al., 2007; Nastos and Zerefos, 2008; Karpouzos et al., 2010; Kambezidis et al., 2010; Nastos et al., 2016); the results were more important for the management of extreme floods when they occurred. For example, it was on November 2nd 1977 when 38 people in Athens lost their lives because of a heavy flood impact (Nikolaidou and Hadjiichristou, 1995). In order to insert a screening of areas with continuing study of the flooding risk, it is important to fully statistical quantify the precipitation characteristics in a region, for instance in Thessaloniki for northcentral Greece. This method requires a long-time-series data of rainfall records and a temporal distribution of risk assessment in extreme heavy rainfall events would be available for screening (Dawson, 2008; Kandilioti and Makropoulos, 2012; Stoyanova and Georgiev, 2013).

Finally, we delete the double full stop on page 11 and we fix the captions of Table 2 and Figure 6.
The long precipitation time series of 124 years in Thessaloniki (Greece) is statistically analyzed (1892-2015).

Taylor diagrams and Mann-Kendall test are used to identify contributions and trends, respectively.

Advance statistical methodologies are applied in the extreme daily records over the sub-period (1931-2015).

The extreme probability distribution function for the extreme daily records is defined and the return periods for different extreme values of precipitation observed are calculated.
STUDY OF VERY LONG-PERIOD EXTREME PRECIPITATION RECORDS IN
THESSALONIKI, GREECE

Nikoletta Pakalidou* and Petroula Karacosta
1 School of Chemical Engineering and Analytical Science, University of Manchester, United Kingdom
2 School of Physics, Aristotle University of Thessaloniki, Greece
*corresponding author’s e-mail: nikoletta.pakalidou@manchester.ac.uk

Abstract The objective on this study is to fully statistically determine and describe the rainfall regime
characteristics of the very long precipitation records of 124-year period (1892–2015) in Thessaloniki and
moreover to apply extreme value statistical methodologies to extreme daily rainfall information (period,
1931-2015), in order to define the extreme precipitation probability distributions. To meet these objectives,
the observed annual, monthly and daily rainfall measurements obtained at the Meteorological Station of the
Department of Meteorology and Climatology of the Aristotle University of Thessaloniki are used. The use
of both Mann-Kendall test and moving average methods are used to determine the profile of the examined
very long-period rainfall regime in Thessaloniki, on annual, seasonal, and monthly basis. Taking into
consideration the climatic change effect upon precipitation regimes, the extreme daily rainfall
measurements are statistically analyzed with extreme value statistical methodologies, defining thus the
extreme precipitation probability distribution functions. The Taylor diagram analysis, through combinations
of appropriate correlation coefficients, standard deviations, and root-mean-square differences, are
constructed to statistically quantify the degrees of similarity between reference points (mean annual
precipitation) and results on seasonal and agricultural (agro-) seasonal basis. All the above tests are applied
to detect possible changes in precipitation characteristics, over the study region, throughout the extreme
long period of 124 years. The resulted probability distributions might potentially contribute to the
prediction of extreme rainfall events, and be used for future climatic change projections. Moreover, based
upon the resulted-adopted probability distributions, emerging from this extreme long study period, return
periods for different extreme values of precipitation observed are calculated.

Keywords: Extreme precipitation; statistical analysis; Taylor diagrams; Mann-Kendall test; extreme value
statistics; extreme probability distributions; Thessaloniki (Greece)

1. Introduction

Precipitation and temperature are the most important variables that can describe
the climate of a region. Changes in both variables may drive in difficult situations for the
people in a region. For example, they should migrate in the case of extreme changes in
precipitation as floods or droughts events push them to move to another part of their
country (or outside of their country), in order to increase again the agricultural
productivity. Except the deaths counted in cases of extreme flood events, these
phenomena have an impact in both economy and society of the impacted area (Wang et
al., 2010; Sayemuzzaman and Jha, 2014; Limsakul and Singhruck, 2016; Sun et al.,
2016). In Europe, the management of solutions when extreme disasters by natural
phenomena such as floods occurred is a vital important section. Focusing on the
Mediterranean region, the temporal analysis of precipitation data plays a crucial role in
protection of agricultural countries, located on it by natural hazards as a result of global
warming changes. The last is well understood if one considers the location of it between approximately 30° to 40° north and south latitude, namely at the middle latitudes. It seems to be affected by the global changes and it addresses a significant variability in rainfall regimes.

A number of studies have described the rainfall pattern in the Mediterranean area; there is an unambiguous negative trend in precipitation’s variability during winter season since the 1960s (Tselioudis et al., 2008; Philandras et al., 2010). This could be explained by the North Atlantic Oscillation (NAO) and by climatic models scenarios from IPCC. Indeed, there is a positive trend in NAO index that links the negative trends of rainfall data (Krichak et al., 2014). Moreover, SRES A1B scenario of IPCC indicates rainfall relocation at north direction that will affect the southern areas in the Mediterranean region decreasing the rainfall records there. The results that refer to the future period 2080-2099 indicate a decreasing by 20% of rainfall in comparison with the observations over the period 1980-1999 (Dünkeloh and Jacob, 2003; Luterbacher et al., 2005; Nastos and Zerefos, 2009; Sindosi et al., 2015; Zittis et al., 2017). On the other hand, heavy-precipitation frequency is predicted to statistically increase over most areas in Mediterranean region (Partal and Kahya, 2006; Ruiz-Leo, 2013).

Greece is located at the southeast part of European Mediterranean area. The geomorphology and geography of the country influence its climatic type, as it can be described as Mediterranean (Csa) according to Köppen classification. Several researchers have studied the temporal or spatial characteristics of precipitation in Greece (Feidas et al., 2007; Nastos and Zerefos, 2008; Karpouzos et al., 2010; Kambezidis et al., 2010; Nastos et al., 2016); the results were more important for the management of extreme floods when they occurred. For example, it was on November 2nd 1977 when 38 people in Athens lost their lives because of a heavy flood impact (Nikolaidou and Hadjichristou, 1995). In order to insert a screening of areas with continuing study of the flooding risk, it is important to fully statistically quantify the precipitation characteristics in a region, for instance in Thessaloniki for northcentral Greece. This method requires a long-time-series data of rainfall records and a temporal distribution of risk assessment in extreme heavy rainfall events would be available for screening (Dawson, 2008; Kandilioti and Makropoulos, 2012; Stoyanova and Georgiev, 2013).

In this study, the availability of an extremely long time-series with rainfall records in Thessaloniki affords the opportunity to examine positive or negative statistically significant or not trends. To do so, a variety of statistical tools is used performing advanced analyses in given data such as, calculations of correlation coefficients, the root-mean-square differences, the Mann-Kendall (M-K) test, and design of Taylor diagrams. This is the first objective of this study. The second objective is to analyze the extreme precipitation events in the study area. For this purpose, methodologies of Extreme Value Statistic (EVS) are applied to objectively determine the probability distributions of extreme daily records. This is important as the numerical analysis provides a prediction of extreme floods, as well as their recurrence in a future period.
2. Materials and Method

2.1. Study area and data availability

In this study the precipitation data are obtained from the Meteorological Station of the Department of Meteorology and Climatology of the Aristotle University of Thessaloniki, in Greece. The data refer to one basic period of 124 years and one sub-period of 85 years; which are the 1892-2015 and 1931-2015, respectively. Annual and monthly rainfall records correspond to both periods. In addition, hourly observations refer to the second period. Since the full study period includes 124 years, it is long enough to allow conclusions on a statistical basis (Burn and Elnur, 2002). Further, as the time series is very long, it was tested for homogeneity. For this purpose, the Standard Normal Homogeneity Test (SNHT) written by Alexandersson was used (Alexandersson and Moberg, 1997) and the data satisfied the applied test of homogeneity. In the process, the mean monthly, seasonal, agro-seasonal, and annual heights of precipitation (measured in mm) for both long study periods were computed.

The meteorological station of Thessaloniki is located at the north part of Greece, with latitude 40° 37’ N, longitude 22° 57’ E, and altitude 31 m from the sea level. Thermaikos Gulf is at the southern of Thessaloniki and Mount Hortiatis is on its northern. This topology gives a specific climate to Thessaloniki directly affected by the gulf. Indeed, a humid subtropical climate characterizes climate classification. Hence, winters are relatively dry and summers are hot with humid nights.

2.2. Method

In this study, several statistical parameters were used to quantify the trends during both, period and sub-period. In addition, the normal seasons were reorganized in agro-seasons (Karacostas and Pennas, 1994). This modification conducted by Karacostas and Pennas (1994), is considered important as Greece is an agricultural country and the knowledge of wet or dry conditions during the agro-seasons may improve the agricultural productivity. To do so, March and September are relocated; the specific geographic profile of Greece suggests that March bares precipitation similarities with the months in winter normal season, and September has similar characteristics with the months in summer normal season. Thus, the agro-groups are presented below: agro-winter: Dec, Jan, Feb, Mar, agro-spring: Apr, May, agro-summer: Jun, Jul, Aug, Sep, agro-autumn: Oct, Nov. The steps followed in this work are presented further in the paragraphs follow.

2.2.1. Correlation coefficient of precipitation data on time basis

In order to quantify any similarity, or even differences, between and among the precipitation data on monthly, seasonal, agro-seasonal, and annual basis, the correlation coefficient is used. Considering the two variables, $f_n$ and $r_n$ that indicate each one set, which are defined in $N$ discrete points (in time), the terms $< f >$, $< r >$ are the mean
values and $\sigma_f$, $\sigma_r$ are the standard deviations of the variables, respectively. The correlation coefficient $R$ between the variables is defined as:

$$R = \frac{1}{N} \sum_{n=1}^{N} (f_n - \langle f \rangle)(r_n - \langle r \rangle)$$

[1]

The term $R$ reaches 1 as a maximum value when for all $n$, $(f_n - \langle f \rangle) = a (r_n - \langle r \rangle)$, where $a$ is a positive constant. In this case, the two datasets have the same centred pattern of variation, but are not identical, unless $a = 1$. Hence, it is not possible to determine whether the two sets have the same amplitude of variation (Beck et al., 2015; Taylor, 2001).

2.2.2. Root-mean-square differences in precipitation data

In order to quantify differences between the datasets, the root-mean-square (RMS) difference $E$, is used. This is defined as:

$$E = \left[ \frac{1}{N} \sum_{n=1}^{N} (f_n - c_r)^2 \right]^{1/2}$$

[2]

However, in order to isolate the differences in the patterns from the differences in the means of the two datasets, it is defined that: $\langle E \rangle = \langle f \rangle - \langle r \rangle$ and the centred pattern RMS difference is given by (Beck et al., 2015; Taylor, 2001).

$$E_c = \left\{ \frac{1}{N} \sum_{n=1}^{N} [(f_n - \langle f \rangle) - (r_n - \langle r \rangle)]^2 \right\}^{1/2}$$

[3]

Hence, the full RMS difference is defined as: $E^2 = \langle E \rangle^2 + E_c^2$.

2.2.3. Theory of Taylor’s diagram

The statistics mentioned so far ($R$, $E_c$, $\sigma_f$, $\sigma_r$) are used to compare the datasets and can also be illustrated on Taylor’s diagram. The diagram is constructed on the grounds of the relationship among those four statistical quantities, based on the law of cosines:

$$E_c^2 = \sigma_f^2 + \sigma_r^2 - 2\sigma_f\sigma_r R$$

[4]
Using the above definitions, Taylor (2001) constructed a diagram that provides the degree of similarity between two variables; the first variable is known as reference point and the second variable is known as test point. The closest the test-point to the reference point, the highest the degree of similarity between them. In that case, the RMS can be calculated.

### 2.2.4. Theory of Mann-Kendall test

In order to identify the trend in time series of precipitation at the study station, the Mann-Kendall test (M-K) is performed for the 124-year examined period. More precisely, it is known that M-K test is a non-parametrical test and it indicates the presence of an increasing or decreasing trend during a long time series for annual, seasonal, and monthly, time scales at a region (Mann, 1945; Kendall, 1975; Hirsch et al., 1982; Onoz and Bayazit, 2003; Yue and Wang, 2004; Mavromatis and Stathis, 2011; Yürekli, 2015).

The M-K test compares two hypotheses, the null hypothesis ($H_0$) and the alternative ($H_a$) one: $H_0$ that there is no trend, as the data are random and independent, and $H_a$ that there is a trend. The basic equation of M-K test is given by (Kahya and Kalayci, 2004):

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1, j > i}^{n} \text{sign}(T_j - T_i),$$  \hspace{1cm} \text{[5]}

where the value of \( \text{sign}(T_j - T_i) \) depends on the difference between the two values. This means that it is equal to 1, equal to 0, or equal to -1, if this difference is positive, neutral, or negative, respectively. Also, \( n \) corresponds to the data points, \( T_j \) and \( T_i \) are the annual values in years \( j \) and \( i \), respectively. If \( n \geq 10 \), the mean, \( E(S) \), and variance, \( \sigma^2 \), can be calculated as follows:

$$E(S) = 0,$$  \hspace{1cm} \text{[6]}

$$\sigma^2 = \frac{n(n-1)(2n+5) - \sum_{i=1}^{p} t_i(t_i-1)(2t_i+5)}{18},$$  \hspace{1cm} \text{[7]}

where \( p \) is the number of ties groups and \( t_i \) is the number of data points in the \( i \)th tied group. Also, it is important to notice that for a positive \( S \), an increasing trend is observed, but for a negative \( S \), a decreasing trend is noted. To continue, a measure to compute the significance of the trend noticed is both the test statistic \( Z \) and Kendall’s \( \tau \) (Partal and Kahya, 2006):
The value of $Z$ indicates if the trend is significant or not; if $Z > Z_{a/2}$, where $a$ is the chosen significance level (for example, for significance level 5%, $Z_{0.025} = 1.96$), then $H_0$ hypothesis derives to a significant upward/downward trend.

### 2.2.5. Thresholds for extreme precipitation records

The extreme rainfall records are examined using the daily measurements for the 85-year sub-period. An absolute threshold index $R30$ mm is established and a separation of data set above and below the limit of $R30 = 30$ mm/day is considering. This threshold is obtained by the literature as a common level to study daily precipitation records (Dauphine, 1975; Costa and Soares, 2009). However similar thresholds within ±5 mm/day were examined ($R25$ mm, $R35$ mm) in the process of goodness-of-fit (GOF) tests described later. The results using R30 were similar in comparison with the results obtained using thresholds $R25$ and $R35$. It is important to highlight that daily rainfall records below the threshold $R30$ mm are ignored here after. This process leads to 140 cases of extreme daily precipitation events.

### 2.2.6. Goodness-of-fit tests

The 140 cases of extreme daily precipitation events are used as an input in a group-of-three goodness-of-fit (GOF) tests; Kolmogorov-Smirnov (K-S), Anderson-Darling (A-D), and Chi-Squared ($X^2$) tests (Yates, 1934). The basic aim of the GOF tests is to detect the best fitted probability distribution for the precipitation records in the study area in order to further project the rainfall profile in the future. To do so, the process of the tests compares the extreme records with a number of theoretical functions of the probability distributions, and pairs them with the best fitting distribution. The significance level used is 0.05. The three GOF tests used in this work are briefly presented followingly. For all tests, the null hypothesis ($H_0$) is first processed, according
to which the records follow a specified distribution. The alternative hypothesis \( (H_a) \) arises if the data does not follow the specified distribution.

(a) Kolmogorov-Smirnov test

The Kolmogorov-Smirnov test is defined by the statistical expression (Chakravart et al., 1967):

\[
D = \max_{1 \leq i \leq N} \left( F(Y_i) - \frac{i - 1}{N} \right) \left( F(Y_i) - \frac{i}{N} \right)
\]

where \( F \) is the cumulative distribution function, \( Y_i \) is the data analysed (here is the precipitation records) and \( N \) corresponds to the number of values used. The test calculates the maximum difference between the empirical distribution function and the cumulative distribution function.

(b) Anderson-Darling test

The Anderson-Darling test is a modified K-S test, as it concerns more the tails of distributions, compared to the K-S test. The expression of this test is given by (Stephens, 1974):

\[
A^2 = -N + \sum_{i=1}^{N} \left( \frac{2i - 1}{N} \right) [\ln F(Y_i) + \ln(1 - F(Y_{N+1-i}))]
\]

where \( F \) is the cumulative distribution function, \( Y_i \) is the data analysed in the order, and \( N \) corresponds to the number of values used.

(c) Chi-square test

In the Chi-square test a \( k \) number of bins is used to divide the data used. The test is given by (Snedecor and Cochran, 1989):

\[
X^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}
\]

where \( O_i \) and \( E_i \) correspond to the observed and expected frequency for the bin \( i \), respectively. The observed frequency is known, but the expected frequency is given as:
\[ E_i = N(F(Y_u) - F(Y_l)), \]  

where \( N \) is the number of values used, \( F(Y_u) \) and \( F(Y_l) \) correspond to the cumulative distribution functions when both an upper and a lower limit are defined, respectively.

### 2.2.7. Continuous probability distributions

The three tests described above, use the extreme events as input and compare it with the values that are expected from the 60 probability distributions that are available. The best fitted probability distribution is the one with the minimum absolute deviation between the observed and the predicted values.

**(a) Generalized Pareto distribution**

The Generalized Pareto distribution is characterized by three parameters; the shape parameter \( k \), the location parameter \( \mu \), and the scale parameter, \( \sigma \) (\( \sigma > 0 \)). The domain where this distribution can be defined is (Pickands III, 1975):

\[
\mu \leq x < +\infty, k \geq 0, \\
\mu \leq x \leq \mu - \frac{\sigma}{k}, k < 0.
\]

The probability density function derived by the Generalized Pareto distribution is given as:

\[
f(x) = \begin{cases} 
\frac{1}{\sigma} (1 + kz)^{-1-k}, k \neq 0 \\
\frac{1}{\sigma} \exp(-z), k = 0
\end{cases}, \text{where } z = \frac{x - \mu}{\sigma}.
\]

**(b) Johnson SB distribution**

The Johnson SB distribution is a four-parameter model; the shape parameters \( \gamma \) and \( \delta \) (\( \delta > 0 \)), the scale parameter \( \lambda \) (\( \lambda > 0 \)), and the location parameter \( \xi \). The domain of this distribution is defined as (Johnson, 1949):

\[
\xi \leq x \leq \xi + \lambda.
\]

The probability density function derived by the Johnson SB distribution is given as:
The Generalized Extreme Value distribution is a three-parameter model; the shape parameter $k$, the location parameter $\mu$, and the scale parameter, $\sigma$ ($\sigma > 0$). The domain where this distribution can be defined is (Jenkinson, 1955):

$$1 + k \frac{(x-\mu)}{\sigma} > 0, k \neq 0,$$

$$-\infty \leq x < +\infty, k = 0. \tag{18}$$

The probability density function derived by the Generalized Extreme Value distribution is given as:

$$f(x) = \begin{cases} 
\frac{1}{\sigma} \exp\left(- \left(1 + k z\right)^{-1/k} \left(1 + k z\right)^{-1-1/k}\right) & k \neq 0 \\
\frac{1}{\sigma} \exp\left(- \left(1 - \exp(z)\right)\right) & k = 0
\end{cases}, \text{where } z = \frac{x-\mu}{\sigma} \tag{19}$$

2.2.8. Study of return periods of extreme events

In the management of extreme events, one of the basic quantities is the return period. This period is defined as the interval within an extreme event-flood could recur. For this purpose, the observations of the extreme daily records (higher than or equal to 30 mm/day) during the 85-year period are taking into account and the return period $T$ for precipitation extreme events is given by: $P(x > x_T) = 1/T$, where $P(x > x_m)$ corresponds to the probability of the $m^{th}$ value, $P(x > x_m) = m/n$, $x_T$ describes the return level, $n$ is the total number of values, and $m$ is the rank of each value in a decreasing mode (Coles, 2001).

3. Results and Discussion

In this study, the monthly, seasonal, agro-seasonal and annual descriptions are taking place for the period of 124 years. Graphical presentation is used to analyze and depict the results concerning the differences and/or similarities among different time-
groups, as Mann-Kendall (M-K) test and Taylor diagrams illustrate. Furthermore, results are presented on daily basis through the time series of 85 years. The GOF tests are used to generate the best fitted distribution and the presentation of recurrence of extreme-rainfall events in Thessaloniki.

3.1. Period 1892-2015

3.1.1. Analysis of precipitation trends on annual basis

The annual precipitation time series was analyzed in Figure 1, presenting the precipitation height for the very long period of 1892-2015. The extremist high and low precipitation years are identified and depicted through the examined period. It is worth mentioning the extremist rainfall measurement of 912.2 mm obtained in the wettest year of 2014 and the second extreme observed in 1919 with 722.1 mm. It is significant to notice that 133 raindays were counted in 2014 (Tolika et al., 2017) and it is possible this phenomenon belonged to a larger scale over the Balkan-Peninsula region (Tošić et al., 2017). On the other hand, it should be noticed that three years can be characterized as driest years (1932, 1977 and 2000).

![Fig. 1. Annual precipitation time series over the period 1892-2015 in Thessaloniki. The two highest (red points) and three lowest (blue points) precipitation records are marked.](image)

Table 1 summarizes the results drawn from the statistical analyses within the period 1892-2015. The mean annual precipitation height for the examined period is 457.6 mm with the highest and lowest values to reach 912.2 mm and 230.8 mm, whereas the standard deviation and coefficient of variation were 105.4 mm and 23%, respectively.
The annual M-K test was found with a positive trend ($Z=0.376$) by applying Equation 8 (Table 1).

**Table 1.** Results from statistical analysis on annual, seasonal, agro-seasonal, and monthly basis, for the period 1892-2015, in Thessaloniki. Equations 1 and 5-9 are used to compute correlation coefficient and Mann-Kendall parameters. Other values are derived by applying basic statistical analysis.

<table>
<thead>
<tr>
<th>Time Periods</th>
<th>Mean (mm)</th>
<th>St Dev (mm)</th>
<th>CV (%)</th>
<th>Correl. with annual</th>
<th>Min value (mm)</th>
<th>Max value (mm)</th>
<th>M-K (tau)</th>
<th>M-K (S)</th>
<th>M-K (Z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual</td>
<td>457.6</td>
<td>105.4</td>
<td>23.0</td>
<td>---</td>
<td>230.8</td>
<td>912.2</td>
<td>0.023</td>
<td>175</td>
<td>0.376</td>
</tr>
<tr>
<td>Jan</td>
<td>37.1</td>
<td>26.7</td>
<td>72.1</td>
<td>0.37</td>
<td>0.0</td>
<td>111.0</td>
<td>-0.013</td>
<td>-98</td>
<td>-0.210</td>
</tr>
<tr>
<td>Feb</td>
<td>32.8</td>
<td>25.3</td>
<td>77.0</td>
<td>0.27</td>
<td>0.1</td>
<td>129.1</td>
<td>0.083</td>
<td>635</td>
<td>1.369</td>
</tr>
<tr>
<td>Mar</td>
<td>36.8</td>
<td>24.0</td>
<td>65.3</td>
<td>0.44</td>
<td>0.2</td>
<td>101.5</td>
<td>0.063</td>
<td>478</td>
<td>1.030</td>
</tr>
<tr>
<td>Apr</td>
<td>38.7</td>
<td>27.5</td>
<td>70.9</td>
<td>0.17</td>
<td>2.7</td>
<td>133.0</td>
<td>-0.005</td>
<td>-37</td>
<td>-0.078</td>
</tr>
<tr>
<td>May</td>
<td>45.9</td>
<td>31.2</td>
<td>67.9</td>
<td>0.18</td>
<td>0.2</td>
<td>164.3</td>
<td>0.017</td>
<td>127</td>
<td>0.272</td>
</tr>
<tr>
<td>Jun</td>
<td>34.6</td>
<td>26.1</td>
<td>75.6</td>
<td>0.25</td>
<td>0.0</td>
<td>133.3</td>
<td>-0.098</td>
<td>-747</td>
<td>-1.611</td>
</tr>
<tr>
<td>Jul</td>
<td>26.1</td>
<td>26.3</td>
<td>100.7</td>
<td>0.35</td>
<td>0.0</td>
<td>118.8</td>
<td>0.054</td>
<td>409</td>
<td>0.881</td>
</tr>
<tr>
<td>Aug</td>
<td>21.4</td>
<td>21.9</td>
<td>102.1</td>
<td>0.25</td>
<td>0.0</td>
<td>145.2</td>
<td>0.050</td>
<td>381</td>
<td>0.821</td>
</tr>
<tr>
<td>Sep</td>
<td>31.9</td>
<td>29.5</td>
<td>92.6</td>
<td>0.33</td>
<td>0.0</td>
<td>125.1</td>
<td>0.054</td>
<td>411</td>
<td>0.886</td>
</tr>
<tr>
<td>Oct</td>
<td>49.3</td>
<td>37.2</td>
<td>75.6</td>
<td>0.42</td>
<td>0.0</td>
<td>194.4</td>
<td>-0.048</td>
<td>-363</td>
<td>-0.782</td>
</tr>
<tr>
<td>Nov</td>
<td>51.5</td>
<td>34.6</td>
<td>67.2</td>
<td>0.27</td>
<td>4.8</td>
<td>185.6</td>
<td>-0.061</td>
<td>-463</td>
<td>-0.998</td>
</tr>
<tr>
<td>Dec</td>
<td>51.4</td>
<td>35.6</td>
<td>69.3</td>
<td>0.33</td>
<td>0.2</td>
<td>166.2</td>
<td>-0.001</td>
<td>-6</td>
<td>-0.011</td>
</tr>
<tr>
<td>Spring</td>
<td>121.5</td>
<td>45.6</td>
<td>37.5</td>
<td>0.46</td>
<td>24.1</td>
<td>274.8</td>
<td>0.049</td>
<td>376</td>
<td>0.810</td>
</tr>
<tr>
<td>Summer</td>
<td>82.1</td>
<td>47.0</td>
<td>57.2</td>
<td>0.46</td>
<td>3.4</td>
<td>245.8</td>
<td>-0.007</td>
<td>-56</td>
<td>-0.119</td>
</tr>
<tr>
<td>Autumn</td>
<td>132.7</td>
<td>56.2</td>
<td>42.4</td>
<td>0.61</td>
<td>1.7</td>
<td>312.4</td>
<td>0.012</td>
<td>90</td>
<td>0.192</td>
</tr>
<tr>
<td>Winter</td>
<td>121.3</td>
<td>53.2</td>
<td>43.9</td>
<td>0.26</td>
<td>6.8</td>
<td>285.6</td>
<td>0.041</td>
<td>316</td>
<td>0.680</td>
</tr>
<tr>
<td>Agro-Spring</td>
<td>84.7</td>
<td>39.2</td>
<td>46.3</td>
<td>0.26</td>
<td>20.3</td>
<td>216.9</td>
<td>0.000</td>
<td>0</td>
<td>0.000</td>
</tr>
<tr>
<td>Agro-Summer</td>
<td>114.0</td>
<td>59.0</td>
<td>51.7</td>
<td>0.53</td>
<td>11.6</td>
<td>346.4</td>
<td>0.027</td>
<td>208</td>
<td>0.447</td>
</tr>
<tr>
<td>Agro-Autumn</td>
<td>100.8</td>
<td>49.7</td>
<td>49.3</td>
<td>0.50</td>
<td>8.8</td>
<td>250.2</td>
<td>-0.055</td>
<td>-418</td>
<td>-0.901</td>
</tr>
<tr>
<td>Agro-Winter</td>
<td>158.1</td>
<td>63.2</td>
<td>39.9</td>
<td>0.62</td>
<td>30.3</td>
<td>361.9</td>
<td>0.076</td>
<td>580</td>
<td>1.251</td>
</tr>
</tbody>
</table>

The further analysis of the annual precipitation record did not show any significant trend at the 95% confidence level. Dry and wet periods were examined throughout the long period and they are presented in Figure 2. To do so, the percentage departures from the mean annual records were computed and grouped into thirteen 10-year groups. One can conclude that the negative percentage departures from the mean (dry periods) were most frequent than the positive percentage departures (wet periods). This was clearly observed at about half of the 124 years of the study period. Indeed, there was a main turning point starting in 1961, when the positive percentages increased,
providing an extreme high percentage during the period 2011-2015 (~8.0%). The last could be well understood by observing in the annual precipitation time series the two maxima in 2014 (~900 mm) and in 2011 (~550 mm). It worth to remark that Feidas et al. (2007) studied the precipitation records in Greece during 1955-2001 using surface and satellite data, and they concluded to a downward trend in both winter and annual records. This is well understood as in Figure 2 one can observe the high positive percentage during the period 2011-2015. The last seems to change the above trends observed by Feidas et al. (2007).

Fig. 2. 10-year groups of normalized sums of the percentage departures from the mean precipitation, during the period 1892-2015 in Thessaloniki. The positive and negative percentages are represented by red and blue horizontal columns, respectively.

3.1.2. Analysis of precipitation trends on monthly, seasonal and agro-seasonal basis

(a) Monthly trends

In Table 1, the monthly precipitation heights show that the highest values were marked in both November and December (~51 mm), characterizing them as the wettest
months. On the other hand, August could be characterized as the driest month during the examined period of the 124 years. From the correlation coefficient values it is implied that the months, March and October, contribute the most towards the mean annual precipitation (457.6 mm). On the contrary, April and May contribute less to the mean annual record. The majority of the months (~83%) appear to have no precipitation at all as their minimum value during the examined period, whereas October was found with the maximum monthly value (194.4 mm). Further, the M-K test (Z-value) applied in monthly precipitation records, showed negative trends for about half of the months (Equation 8). As it can be seen in Table 1, the highest positive slope had the magnitude of +1.369 in February and the lowest negative one was -1.611 in June.

(b) Seasonal trends

Concentrated to the maximum precipitation heights observed during the whole period (1892-2015) in normal seasons, appear to be similar for all seasons, ranging from 250 mm to 320 mm. In Figure 3, one can notice the maximum precipitation height observed for the normal seasonal and annual precipitation conditions in the study area. Autumn has the highest maximum amount of precipitation (312.4 mm in 1896) and summer the lowest maximum (245.8 mm in 2014), during the 124-year period (Figure 3). While, 912.2 mm were measured for the extreme precipitation year of 2014, and only 200.8 mm for the year 2000. Winter precipitation records showed a maximum in 1935 (285.6 mm) and a minimum in 1991 (6.8 mm) during the period 1892-2015. Nevertheless, the lowest minimum corresponded to summer in 1927 (3.4 mm). The seasonal precipitation trends, according to the M-K test (Table 1), appear to be positive, except of the summer, which indicates negative value (Equation 8). As one would expect, autumn presented the highest precipitation mean (~133 mm), while the lowest noted in summer (~82 mm), which is well understood, as it is a basic characteristic of the Mediterranean climatic type (Feidas et al., 2007).
(c) Agro-seasonal trends

Positive trends were noted in agro-summer and agro-winter, whereas the only negative trend was found in agro-autumn (Table 1). Agro-spring was found with Z-value in M-K test equals to zero (Equation 8). Further, the lowest precipitation value, among the agro-seasons, belonged to agro-spring, whereas the maximum value was observed in agro-winter. It is therefore obvious that the relocation of March and September into agro-winter and agro-summer respectively, causes a significant difference between normal seasons and agro-seasons. For example, during autumn, an increasing trend is evident. However, relocating September to agro-summer, leaves agro-autumn with the months of October and November, resulting to a decreasing trend.

(d) Quantifying the condition to the annual precipitation

Another aspect worth taking into consideration is the quantification of the degree of contribution of each normal season and agro-season to the mean annual precipitation. To this end, Taylor diagrams were created (Figure 4). Calculations of root-mean-square differences between the data sets, and the standard deviation of each data set, were
essential for the development of the Taylor diagrams, in order to demonstrate the degree of contribution of observations to a specific reference point. In particular, the mean seasonal and agro-seasonal rainfall records were used as the observation points, while the reference point corresponded to the mean annual precipitation, over the studied period 1892-2015. To further investigate the results from the above-mentioned diagrams, the values of the correlation coefficient of each period with the mean annual rainfall were calculated (Equations 1, 2, and 4). The results are discussed in the following paragraph.

The 20-year-less-rain period and 20-year-more-rain period (Figure 4(a) and (b)) were examined. During both periods (less-rain during 1981-2000, more-rain during 1961-1980), autumn contributed more to the reference point (mean annual records), with the highest degree during the more-rain period. Between the less-rain and more-rain periods through the total period of 124 years, winter was found to contribute more to the reference point in the first case (Figure 4 (c)). However, in the Taylor diagram, autumn was presented with more similar characteristics in rainfall profile with the reference point (Figure 4 (d)). These results are in a good agreement with the study of Tolika et al. (2017). Indeed, the authors remarked that autumn contributed by a percentage of 27-28% to the annual precipitation during 1958-2013 in Thessaloniki region. The authors concluded that this percentage was little more (29%) during 2014.
Fig. 4. Taylor diagrams represent the degree of contribution of each observation to the reference point (mean annual precipitation record during the period 1892-2015). Equations 1, 2, and 4 are used to design these diagrams. (a-e) normal seasons, (f) agro-seasons; (a) 1981-2000 less-rain period, (b) 1961-1980 more-rain period, (c) 1892-2015 less-rain period, (d) 1892-2015 more-rain period, (e) 1892-2015 period with total data in normal seasons, (f) 1892-2015 period with total data in agro-seasons.
Furthermore, regarding the degree of contribution of each normal season and agro-season to the annual rainfall data, autumn was found to contribute the most in the seasonal study, and agro-winter in the agro-seasonal study (Figure 4(e) and (f)). On the other hand, winter and agro-spring contributed the least to the mean annual precipitation. These results are explained by the above-mentioned relocation of March and September, due to which the rainfall patterns of the two season categories are affected. Moreover, the higher correlation coefficient in autumn and agro-winter, and also the lower one in winter and agro-spring, demonstrate that the results obtained by Taylor diagrams were in a good agreement with the basic statistical analysis results presented in Table 1.

3.2 Sub-period 1931-2015

In the field of management of extreme events, observing the past in order to predict the future, is of paramount importance. To this end, daily observations of the sub-period 1932-2015 were used in order to determine return periods of the extreme events, which correspond to specific distribution functions. The filter of the threshold of R30 mm was applied and 140 extreme precipitation cases were counted (Figure 5). These 140 cases of extreme daily precipitation events were used as input information in search for the best fitted probability distribution. Sixty statistical distributions were examined for this purpose. The GOF distribution, through the Kolmogorov-Smirnov (Equation 10), Anderson-Darling (Equation 11) and Chi-Squared tests (Equations 12 and 13), implied that the Generalized Pareto distribution was the best fitted one. This conclusion can be drawn by examining the statistics of the three final distributions (Equations 14-19) according to the applied tests (Table 2). Figure 6 depicts the best fitted probability density functions, based on the 140 cases, for the sub-period 1931-2015, in Thessaloniki. These are: the Generalized Pareto for the Kolmogorov-Smirnov test, the Johnson SB for the Anderson-Darling, and the Generalized Extreme Value for the Chi-Squared.

The last step to manage extreme rainfall events is to determine the return periods (frequencies) of the two highest daily rainfall cases: the 16th of July 2014 (107.3 mm) and the 24th of November 1985 (98.0 mm). The projections of these events arise by observing the past (historical) records. Figure 7 displays the frequency of extreme precipitation events as a function of the reoccurrence interval (return period). It becomes obvious, that 280 year return period corresponds to 107.3 mm, meaning that, in this study area and at that point of time, about 107 mm of rain are expected to be observed after 280 years. For the second extreme case of 98 mm, the expected reoccurrence time is 93 years. Moreover, a thorough study of Figure 7 confirms that the highest precipitation amount, the longer return period. For example, one extreme rainfall event of 30 mm/day (the adopted criterion of the resulted 140 cases) could reoccur in one (1) year, or one extreme rainfall event of 60 mm/day could reoccur in 10 years. It should be stated that these numbers are accurate under some confidence, as uncertainties are evident due to the unpredictability
of nature and, in addition, the rainfall-return-period methodology should be adjusted, due to climatic changes.

**Fig. 5.** Extreme daily precipitation records $\geq 30$ mm/day (threshold $R30$ mm); 140 cases during the sub-period 1931-2015 in Thessaloniki. The colors describe the individual intervals of mm/day; green: $30 \leq x < 40$, cyan: $40 \leq x < 50$, orange: $50 \leq x < 60$, pink: $60 \leq x < 90$, red: $x \geq 90$.

**Table 2.** Extreme daily precipitation records (140 cases) with the threshold of $R30$, during the sub-period 1931-2015, in Thessaloniki. The goodness-of-fit distribution and the K-S, A-D, and $X^2$ tests concluded that the Generalized Pareto distribution is the best fitted one. Equations 10-13 are used to apply the tests and Equations 14-19 are used to calculate all distributions’ parameters.

<table>
<thead>
<tr>
<th>Test</th>
<th>Distribution</th>
<th>Statistic</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K$-S</td>
<td>Gen. Pareto</td>
<td>0.0461</td>
<td>$k=0.047$, $\sigma=11.28$, $\mu=29.992$</td>
</tr>
<tr>
<td>$A$-D</td>
<td>Johnson SB</td>
<td>0.3380</td>
<td>$\gamma=2.822$, $\delta=1.00$, $\lambda=161.07$, $\xi=28.818$</td>
</tr>
<tr>
<td>$X^2$</td>
<td>GEV</td>
<td>5.2348</td>
<td>$k=0.269$, $\sigma=6.36$, $\mu=35.887$</td>
</tr>
</tbody>
</table>
Fig. 6. Results of the distribution functions for the 140 cases with ≥ 30 mm/day during the sub-period 1931-2015 in Thessaloniki; [Top] The Generalized Pareto/Kolmogorov-Smirnov test, [Medium] the Johnson SB/Anderson-Darling, and [Bottom] the GEV/\(X^2\).
Fig. 7. Return periods expressed by the diagram of frequency of the extreme events related to reoccurrence interval in Thessaloniki. The red open square describes the return period of the maximum of 107.3 mm observed during the sub-period 1931-2015 in Thessaloniki.

4. Conclusions

The extremely long period of 124 years, from 1892 to 2015, was used to estimate the trends of precipitation records on annual, seasonal, agro-seasonal, and monthly bases, in Thessaloniki. Also, daily precipitation data were analyzed to describe the daily rainfall distributions over the sub-period 1931-2015. Since an extended statistical analysis was performed over the whole period, only summer in normal seasons and autumn in agro-seasons indicated negative trends by Mann-Kendall results. The majority of the seasons and agro-seasons exhibited positive trends, which are not statistically significant at the 95% confidence level, in general. The degree of contribution of different periods to the mean annual precipitation rainfall for the whole period was examined and illustrated by using the Taylor diagrams. The results indicated a very good agreement with the statistical analyses (Mann-Kendall test) results. Both methodologies demonstrated that the maximum contribution to the mean annual precipitation came from autumn and agro-winter. It was obvious that the relocation of March and September, due to the definition of agro-seasons, has an impact to the differences observed between normal seasons and agro-seasons.
Further, the annual precipitation observations in 2014 set a new record since the beginning of measurements, that is 1892. That year might be characterized as the wettest year up to now, if one considers only the individual extreme event observed in the 16th of July. Nevertheless, this was an additional point for a more detailed study about the estimating extreme rainfall events by using specific probability density functions. Thus, the sub-period of 1931-2015 was examined for extreme events, considering the threshold of \( R30 \) mm on daily basis. Hence, 140 individual extreme rainfall events have been identified and considered to finally provide the Generalized Pareto probability distribution function as the best fitted distribution over the examined sub-period of 85 years. These very extreme rainfall events have catastrophic results, the most important of which is flooding. Thus, the purpose of defining a specific probability distribution function that can simulate the flooding events over the study region is to control flood disasters by minimizing the impact of these events on human’s life and to establish a protection system on a general basis. The Generalized Pareto probability distribution showed the return periods in 93 years and 280 years for the two observed highest precipitation records of 98.0 mm and 107.3 mm, respectively. Moreover, the return-period diagram was mapped out and an initial step for a future warning system of flooding events in Thessaloniki has been pointed out.

Obviously, the above results can be used as a driver for further studies on extreme precipitation events. The adopted methodology on deriving the best fitted probability distribution function could be a useful tool to fully describe and study the precipitation regime at a region, and at the same time to differentiate and compare the precipitation profile from one region to the other. Moreover, to provide the tools for the projection of the precipitation regime to future, through the climatic changes.

References


