**Natural Hazards
and Earth
System Sciences**

Assessing precipitation distribution impacts on droughts on the island of Crete

A.-E. K. Vrochidou [1] **and I. K. Tsanis**[1,*]

[1]Department of Environmental Engineering, Water Resources Management and Coastal Engineering Laboratory, Technical University of Crete, Greece
[*]on research leave: Department of Civil Engineering, McMaster University, Hamilton, Canada

Correspondence to: I. K. Tsanis (tsanis@hydromech.gr)

Abstract. Precipitation records from 56 stations on the island of Crete (Greece) revealed that areal mean annual precipitation is of a strong orographic type and its magnitude decreases in west-east direction by as much as 400 mm on average. Amongst many parameters that influence precipitation, the elevation and longitude were the most important and provided the highest spatial correlation. It was found that during the year with minimum precipitation, the precipitation shortage was greater at high elevations while the precipitation excess during the year with maximum precipitation was greater in the western part of the island. The assessment of the spatial and temporal distribution of droughts was carried out with the aid of the Spatially Normalized Standardized Precipitation Index (SN-SPI) for the period 1974–2005 in order to compare drought conditions between neighbouring areas of differing precipitation heights. The analysis showed that severe droughts occurred around the year 1992–1993, with a duration of up to 3 yr. Multiple linear regression (MLR) modeling of precipitation in conjunction with cluster analysis of drought duration exhibits the linkage between precipitation, droughts and geographical factors. This connection between spatial precipitation distribution and geographical parameters provides an important clue for the respective spatial drought pattern. The above findings on the spatio-temporal drought distribution will update the current drought management plans by developing more precise drought warning systems.

1 Introduction

The geographical distribution of precipitation depends on various parameters that include topography, orientation of topography and aspect, direction of wind and continentality (Naoum and Tsanis, 2003). Several studies deal with the variability of precipitation, analysing the relationship between mean annual precipitation and geographical factors (Basist et al., 1994; Guan et al., 2005; Harris et al., 1996; Naoum and Tsanis, 2004), and finding significant increasing relations between these characteristics. A common belief that precipitation amounts increase with elevation was then adopted, thereby proving that the mountainous environment is prone to extreme and frequent precipitation events (Allamano et al., 2009). Naoum and Tsanis (2003) studied the spatio-temporal rainfall characteristics on the island of Crete for a range of 12–50 yr and found that the rainfall-elevation correlation was significant. Drogue et al. (2002) used an operational software called PLUVIA, which distributes point measurements of monthly, annual and climatological rainfall to regularly spaced grid cells through a multiple regression analysis of rainfall versus morpho-topographic parameters derived from a digital elevation model. According to Wotling et al. (2000), a Gumbel rainfall distribution was performed by using a stepwise regression adjusted to rainfall records across Tahiti and an approximation of the pluviometric risk was provided.

The study of weather patterns associated with precipitation events can serve as a reliable early warning system and a non-structural approach for drought mitigation. It is known that storm events are often associated with the development of low-pressure systems (Koutroulis et al., 2010). Over the Mediterranean these systems originate from three main directions, which can be roughly distinguished as west (W), southwest (SW) and northwest (NW). According to Barry and Chorley (2003), depressions that enter the Mediterranean from the Atlantic Ocean (W source) and baroclinic waves from the Atlas mountain range (SW source) influence 9 % to 17 % of the low-pressure systems, respectively. The remaining 74 % form at the lee of the Alps and Pyrenees (NW source). Each class effects of this weather classification are mentioned in literature (Barry and Chorley, 2003) and are well documented.

Furthermore, the definition of relationships between factors can be assessed through clustering methodology, which is the process of grouping the data into classes or clusters so that objects within a cluster have high similarity in comparison to one another, but are very dissimilar to objects in other clusters. Clustering of time series has received considerable attention in recent years as it is a fundamental task in data mining. In literature, various clustering methods have been used in time series: hierarchical clustering (Oates et al., 1999), K-medoids clustering (Kalpakis et al., 2001), nearest neighbour clustering (Zhang et al., 2004), and self-organizing maps (Fu et al., 2001). Amongst all clustering algorithms, K-means clustering has become the most well-known and commonly used partitioning clustering method, because it works well for finding spherical-based clusters in small- to medium-sized databases in hydro-meteorological studies (Chavoshi and Soleiman, 2009; Soltani and Modarres, 2006; Ouyang et al., 2010). Hayward and Clarke (1996) made a statistical analysis of mean monthly rainfall in the Freetown Peninsula, performing division of the gauges with clustering, regarding their orientation. Additionally, multiple regression with the parameters of latitude and the distance from the sea was applied, providing various results for both groups of rain gauges.

Precipitation constitutes the most important key factor for the occurrence of droughts and will allow for a better understanding of the drought occurrence process. Drought is a recurrent climatic feature that is frequent in the Mediterranean region, one of the most vulnerable areas concerning the future precipitation extreme conditions, since it lies in the transitional zone between Northern Africa and Southern Europe. A large number of studies are included in the international literature on examining the effectiveness of various drought indices regarding detection and monitoring drought events and regional drought analysis (Ali and Lebel, 2009; McKee et al., 1993; Palmer, 1965; Rumman et al., 2009). Among the developed drought indices, Standardized Precipitation Index (SPI), Palmer Drought Severity Index (PDSI) and Reconnaissance Drought Index (RDI) are the most commonly used.

Several studies have been performed for the application of the SPI. Especially in the area of Crete, Tsakiris and Vangelis (2004) concluded that the eastern part of the island suffers more frequently from droughts, according to a method based on the estimation of the SPI and its use for characterizing drought. A digital terrain model, based on spatial distribution utilizing a grid analysis and a simple computer calculating process, was used and it was deduced that the proposed procedure could be easily applied to an area of mesoscale dimensions. It was concluded that a significantly persistent drought occurrence was noted during the period 1987 to 1994, while distinct drought events were observed in the years 1973–1974, 1976–1977, 1985–1986 and 1999–2000. Additionally, Tsakiris et al. (2007) estimated drought areal extent for eastern Crete using the SPI and RDI, and deduced

that the driest year during the examined period from 1962–1963 to 1991–1992 is 1989–1990.

The selection of the most appropriate drought index was carried out, according to previous studies. Tsakiris et al. (2007) concluded that SPI and RDI give comparable results. However, discrepancies occur due to the fact that RDI uses an additional meteorological determinant (PET) apart from precipitation. The PDSI, on the other hand, is very complex, spatially variant, difficult to interpret, and temporally fixed. The application of the SPI index covers a significant part of many studies that have been carried out over the last decades (Bonaccorso et al., 2003; Loukas and Vasiliades, 2004; Wu et al., 2007). Nevertheless, Vicente-Serrano and Begueria (2003) point out that drought indices are not as useful in identifying spatial patterns of drought risk since they are based on standardized or normalized shortages in relation to "average conditions", which relate to a given station and a given period.

Although SPI is widely used for assessing drought occurrence, there are some limitations in providing relative information when applied for different regions (at the river basin scale). In its original form it provides a local measure of drought, which as such is not necessarily suitable for comparisons across space and time. As a result, the frequency of drought spells is about the same for all stations no matter if they lie in extremely arid or extremely rainy regions, even though the rainy sites may receive several times more rain than the arid sites. In this context, Dubrovsky et al. (2009) applied the relative drought indices (rSPI and rPDSI), which are calibrated using a reference weather series as a first step, which is then applied to the tested series.

A similar approach for the creation of a modified index based on the SPI was realized by Koutroulis et al. (2011). The Spatially Normalized-SPI (SN-SPI) is a variant of the SPI and allows the comparison between watersheds with different mean annual precipitations. After the normalization, the index smooths the extreme conditions and makes it possible to compare stations by taking into account the spatial character of precipitation. In this paper, the SN-SPI was recommended as a drought index for the calculation of drought climatology for the island of Crete in Greece, because it is spatially normalized for improved assessment of drought severity (Koutroulis et al., 2011). The fundamental virtue of this index consists of its capacity for compacting and unifying spatial information, reducing it to a common language.

The objective of this study is to assess drought events over Crete at both the spatial and temporal scale and connect them with precipitation variability. More specifically, the comprehension of the evolution of this climatic phenomenon and identification of important drought episodes are provided, offering extra reliability to the results via the innovative use of multiple methods. Moreover, the connection between spatial precipitation and weather systems orientation will be of great interest as they comprise a significant clue for the characterization of droughts according to the respective spatial

pattern. Determining precipitation variability and its spatial extent and therefore drought risk periods and areas for the island of Crete, can be used as basic but effective drought mitigation and risk management planning. At the same time, this kind of information may be used to support a simplified drought alert system.

2 Methodology

2.1 The SPI

In the present paper, the SPI is used for assessing drought occurrence in Crete. The index offers the advantage of assessing drought conditions over a wide spectrum of time scales, while comparison between dry and wet periods on different locations is possible. Moreover, it is based on precipitation alone, so that a drought could be assessed even if other meteo-hydrological data are not available (Bonaccorso et al., 2003).

There is a general agreement about the fact that the SPI computed on shorter time scales (3 or 6 months) describes drought events that affect agricultural activities, whereas the longer ones (12, 24 or 48 months) it describes the effects of precipitation deficit on different water resources components such as soil moisture, streamflow, groundwater, and reservoir storage. In this paper, the longest time scale (48 months) is set in the calculation of the SPI.

The SPI index was developed by McKee et al. (1993). In its original version, precipitation for a long period at a station is fitted to a gamma probability distribution, which is then required to be transformed into a normal distribution such that the mean SPI value equals zero. The index values are then the standardized deviations of the transformed precipitation totals from the mean. The gamma distribution is defined by its frequency or probability density function:

$$g(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \qquad \text{for } x > 0 \tag{1}$$

where α is a shape parameter ($\alpha > 0$), β is a scale parameter ($\beta > 0$), x is the precipitation amount ($x > 0$) and $\Gamma(\alpha)$ is the gamma function.

Positive SPI values denote greater than median precipitation whereas negative values denote less than median precipitation. Periods with drought conditions are represented by relatively high negative deviations. Specifically, the "drought" part of the SPI range is arbitrarily divided in four categories; mildly dry ($0 > \text{SPI} > -0.99$), moderately dry ($-1.0 > \text{SPI} > -1.49$), very dry ($-1.5 > \text{SPI} > -1.99$) and extremely dry conditions ($\text{SPI} < -2.0$). A drought event is considered to start when SPI value reaches -1.0 and ends when SPI becomes positive again (McKee et al., 1993). Thresholds of the SPI for drought characterization are presented in Table 1.

Table 1. Thresholds of SPI for drought characterization.

SPI value	Category
2 or more	Extremely wet
1.5 to 1.99	Very wet
1 to 1.49	Moderately wet
0 to 0.99	Mildly wet
0 to −0.99	Mildly dry
−1 to −1.49	Moderately dry
−1.5 to −1.99	Very dry
−2 or less	Extremely dry

2.2 The SN-SPI

The objective of the new modified SPI is the potent comparison of drought events among different areas with different mean annual precipitations at different times. The procedure includes the normalization of SPI values through the incorporation of the precipitation values. The calculation of the SN-SPI is based on a two-step procedure. The first step is the normalization of the SPI index according to the relative average precipitation, based on a set of coefficients (a_i, b_i) that satisfy:

$$\frac{\bar{P}_i}{\bar{P}_{\text{all}}} = a_i \tag{2}$$

$$\frac{\bar{P}_{\text{all}}}{\bar{P}_i} = b_i \tag{3}$$

where $\bar{P}_i$ is the mean monthly precipitation for each watershed i, and $\bar{P}_{\text{all}}$ the mean monthly precipitation for all watersheds. Given a_i and b_i, SPI'_i for each watershed i is calculated through

$$\text{SPI}'_i = \text{SPI}_i \times a_i \quad \text{if } \text{SPI}_i > 0 \tag{4}$$

$$\text{SPI}'_i = \text{SPI}_i \times b_i \quad \text{if } \text{SPI}_i < 0 \tag{5}$$

With the above procedure SPI_i, time series of each watershed i are modified properly in order to include the information of relative average precipitation among all the watersheds of the study area, resulting to the corresponding SPI'_i time series. The second step includes the rescaling of SPI'_i in order to meet the scale of SPI_i, based on the coefficients c and d estimated through

$$\frac{\max(\text{SPI}_i)}{\max(\text{SPI}'_i)} = c \tag{6}$$

$$\frac{\min(\text{SPI}_i)}{\min(\text{SPI}'_i)} = d \tag{7}$$

where $\max(\text{SPI}_i)$ is the maximum SPI value of all watersheds, $\min(\text{SPI}_i)$ is the minimum SPI value of all watersheds,

max(SPI'_i) the maximum modified SPI value of all watersheds, and min(SPI'_i) the minimum modified SPI value of all watersheds. Given c and d, the SN-SPI calculation is defined by

$$\mathrm{SN} - \mathrm{SPI}_i = \mathrm{SPI}'_j \times c, \qquad \text{if } \mathrm{SPI}'_i > 0 \tag{8}$$

$$\mathrm{SN} - \mathrm{SPI}'_i = \mathrm{SPI}'_j \times d, \qquad \text{if } \mathrm{SPI}'_i < 0 \tag{9}$$

Koutroulis et al. (2011) examined the sensitivity of the SN-SPI to the variability of basin scale precipitation based on the example of three hypothetical watersheds with different mean annual precipitation.

2.3 Multiple linear regression (MLR)

The method of analysis used is the method of least squares (LS), which is simply a minimization of the sum of the squares of the deviations of the observed response from the fitted response (Naoum and Tsanis, 2003). This involves the initial assumption that a certain type of relationship, linear in unknown parameters, holds. With precipitation being the dependent (response) variable, the model function is of a specified form that involves both the predictor variables (elevation and longitude) and the parameters. Interaction effects between the variables also can be considered. The unknown parameters are then estimated under certain other assumptions with the help of available data so that a fitted equation is obtained. While the equation might be physically meaningless, it may, nevertheless, be extremely valuable for predicting the values of some variables from knowledge of other variables, perhaps under certain stated restrictions (Draper and Smith, 1998). The general form of the final model is

$$P = b_o + b_1 x_1 + b_2 x_2 \tag{10}$$

where P is precipitation (mm yr^{-1}), x_1 is elevation (m) and x_2 is longitude (km).

2.4 Cluster analysis

Cluster analysis is a multivariate method that aims to classify a sample of subjects (or objects) on the basis of a set of measured variables into a number of different groups such that similar subjects are placed in the same group. K-means clustering (McQueen, 1967) comprises a prototype-based, partitional clustering technique that attempts to find a user-specified number of clusters (K), which are represented by their centroids. The algorithm first selects initial cluster centers (essentially this is a set of observations that are far apart – each subject forms a cluster of one and its center is the value of the variables for that subject). Each object is assigned to its "nearest" cluster, defined in terms of the distance to the centroid and after the formation of the centroids of the clusters, the distance from each object to each centroid is re-calculated moving observations to the "nearest" cluster

once again. (Tan et al., 2006). This process is repeated until the criterion function converges and the centroids remain relatively stable, forming permanent clusters (Struyf et al., 1997). The algorithm aims at minimizing the sum of squared Euclidean distances, for n data points and k centroids, implicitly assuming that each cluster has a spherical normal distribution:

$$F = \sum_{j=1}^{k} \sum_{i=1}^{n} |x_i^j - c_j|^2 \tag{11}$$

where $|x_i^j - c_j|^2$ (squared Euclidean distance) is a chosen distance measure between a data point x_i^j and the cluster center c_j, and represents an indicator of the distance of the n data points from their respective cluster centers.

The algorithm was applied with the use of the MATLAB R2011b package, which also provides a graphical display of the silhouette plot; the silhouette width [$s(i)$] was recommended by Kaufman and Rousseeuw (1990) as a quality index allowing to select the number of clusters and validate the effectiveness of the analysis. The silhouette width of the i-th object is defined by

$$s(i) = \frac{b_i - a_i}{\max\{a_i, b_i\}} \tag{12}$$

where $a(i)$ is the average distance to other elements in the cluster and $b(i)$ the smallest average distance to other clusters. $a(i)$ is interpreted as how well matched i is to the cluster it is assigned (the smaller the value, the better the matching). An $s(i)$ close to one means that the data is appropriately clustered. If $s(i)$ is close to negative one, then the corresponding data is probably in the wrong cluster. An $s(i)$ near zero means that the elements are on the border of two natural clusters.

3 Study area and data description

The island of Crete is located in the southeastern part of the Mediterranean region and it is well known that it comprises an area that has been characterized as one of the most drought prone areas of Greece (Fig. 1). The island covers an area of 8336 km^2; the mean elevation is 460 m and the average slope is 22.8 %. Crete is divided into four prefectures, namely from west to east: Chania, Rethymnon, Heraklio and Lassithi. The mean annual precipitation is estimated to be 750 mm, which varies from east – 440 mm (Ierapetra – elevation: 10 m) to west – 2118 mm (Askifou – elevation: 740 m) and the potential renewable water resources reach 2650 Mm^3.

The actual water use is about 485 Mm^3 yr^{-1}. The main water use in Crete covers irrigation, with a high percentage of 83.3 % of the total consumption. The domestic use, including tourism, covers 15.6 % and the industrial use 1 % of the total consumption (Region of Crete, 2002). The eastern and southern parts are more arid than the west and northern parts, as there is higher precipitation in the northwestern
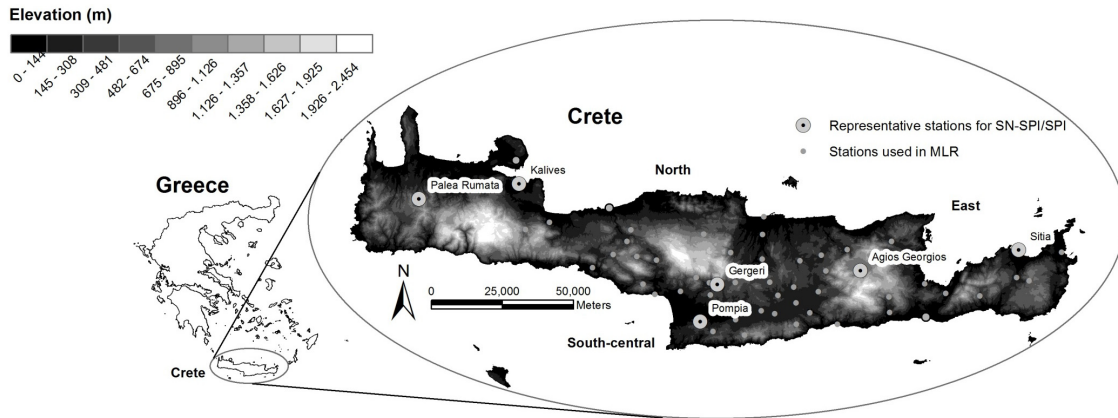
Fig. 1. Area of study, 56 stations used in MLR and 18 representative stations for SPI Crete, Greece.

coastal areas and lower in the southeastern part of the island, a fact that confirms regional variations in water availability (Chartzoulakis et al., 2005). There are significant effects when the uneven spatial and temporal precipitation distributions of Crete, although common in many Mediterranean areas, are related to intensive agricultural activities and the tourism industry (Tsanis and Naoum, 2003).

Monthly precipitation data was compiled by the WRDPC service (Water Resources Department of the Prefecture of Crete) for 56 precipitation stations (Fig. 1). The stations mainly cover the eastern part of the island, which has a higher level of agricultural activity than the western part. The gauges were located at elevations that ranged from sea level, in the prefecture of Iraklion (central Crete), to 905 m a.s.l, in the prefecture of Lassithi (eastern Crete) (Region of Crete, 2009). These data cover a thirty (30) year time period for each month of the hydrological year (September to August), from 1974 to 2005.

In order to simplify calculations, a relative coordinate system was defined by locating an origin (0,0) at the lower left corner of the island at latitude 3 800 000 m and longitude 461 000 m. GGRS87-Greek Geodetic Reference System 1987 was the coordinate system used. All Y coordinates employed in the regression were hence the result of subtracting 3 800 000 from the original latitudes of the different stations and dividing by 1000 to obtain latitudes in km. Similarly, all X coordinates were derived by subtracting 461 000 from the original longitudes of the different stations and dividing by 1000 to obtain longitude values in km. This manipulation of coordinates was important when performing the regression analysis. Large numbers for latitude and longitude could result in small values for model parameters (b_i) and any small error could result in significant changes in the model output. It is then more practical to put all variables into approximately the same order of magnitude to ensure that they receive appropriate weighting in the multiple regression analysis.

Geographic Information System (GIS) technology, represented by the commercial package ArcView GIS 10, was used to provide the tools for spatial data management and to generate maps, including the Digital Elevation Model (DEM). DEM comprises the source of spatially gridded data developed by the Greek Army Geographical Service, with a cell size of 30 m, resulting in a grid of 8723 columns and 3304 rows.

4 Results and discussion

Two parameters (elevation and longitude) were adopted to perform simple linear regression individually for 55 stations. The station of Askifou was excluded from this method in order to obtain more reasonable results, as Askifou (with elevation 740 m) exemplifies a special case of a region with orographic precipitation. The island was divided into three parts (northern, south-central and eastern), so as orientation of topography factors was taken into account. Upon examining the correlation between precipitation and elevation, taking into consideration the three parts, it was found that the regression plot for all stations justifies the positive correlation (Fig. 2a). The plot based on elevation provided a more physically meaningful interpretation of the effect of this variable on precipitation; the presence of the orographic effect at high elevations. On the other hand, further analysis covering the parameter of longitude showed that there is a negative trend between the aforementioned parameter and precipitation (Fig. 2b). The noticeable downward gradient of precipitation from one part to another denotes statistical evidence of the geographical factor. In other words, the decrease of precipitation is responsive to longitude increasing. It is then verified that topographic and geographic factors determine the spatial association in precipitation variations. More specifically, the precipitation magnitude decreases in west-east direction and increases in elevation by as much as 400 mm. Previous studies that tested statistical models relating annual precipitation to elevation or longitude (Naoum and Tsanis,
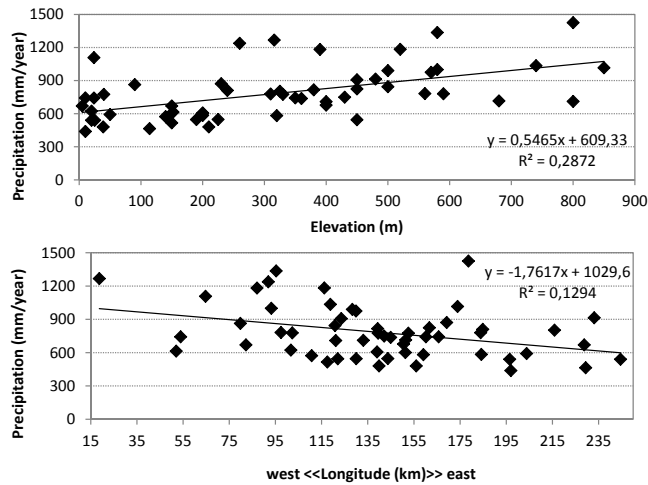
Fig. 2. (a) Elevation versus mean annual precipitation plot and **(b)** Longitude versus mean annual precipitation plot for Crete.

2004) showed that the best results were obtained were both elevation and longitude.

Therefore, simple and multiple linear regressions were used to study the level of association between the two variables for a thirty year period, from 1974 to 2005 (Table 2). The database of precipitation was used to generate and apply the regression models. The regression equations were produced for the 55 stations, representing the whole island. Simple linear regression was carried out considering just elevation as a explanatory variable and the additional independent variable of longitude was introduced in order to construct a more efficient model. The coefficient of determination (R^2) was used to determine the adequacy of the regression equation. The regression analysis showed that 1989–1990 represents the year with the minimum precipitation amount (499 mm) and 2002–2003 was the year with the maximum precipitation amount (1423 mm); moreover, an average year was taken into account (1978–1979, 902 mm). The R^2 varied between 9 % and 42 % for the one-variable model (maximum in 2004–2005) while a range of R^2 23 % to 53 % was obtained for the two-variable model (maximum in 1974–1975). The R^2 values significantly increased when longitude was added to the model, a fact that comprises a possible sign of model improvement.

Subsequently, the regression analysis was carried out for the three subareas (northern, south central and eastern parts) for the three representative years (Table 3). The relationship between precipitation and elevation was not as strong for the whole island (55 stations) as it was for individual parts, especially the eastern part for all years assessed, except 2002–2003. This was attributable to the relatively high elevations of stations in this part, which was likely to result in a higher association than the other parts (R^2 up to 79 % for 1989–1990). Developing separate regression equations and still

Table 2. Simple and two-variable linear regression for 55 stations, regarding each year of 1974–2005. The years 1989–1990 and 2002–2003 are in bold as years with the minimum and maximum precipitation amount, respectively.

Years	b_0	b_1	b_2	R^2	Extracted precipitation (mm)
1974–1975	555.39	0.58		38 %	834
	795.89	0.57	−1.70	53 %	832
1975–1976	728.52	0.56		23 %	996
	1045.06	0.54	−2.23	39 %	995
1976–1977	422.99	0.52		21 %	672
	652.21	0.51	−1.62	31 %	671
1977–1978	807.58	0.99		38 %	1283
	1196.40	0.97	−2.74	51 %	1281
1978–1979	572.08	0.68		40 %	902
	704.94	0.68	−0.94	43 %	901
1979–1980	633.64	0.67		27 %	958
	797.29	0.67	−1.15	31 %	957
1980–1981	703.19	0.64		29 %	1014
	1073.96	0.63	−2.61	49 %	1012
1981–1982	680.11	0.58		23 %	958
	1091.89	0.56	−2.90	48 %	956
1982–1983	452.91	0.55		28 %	719
	672.12	0.54	−1.55	37 %	718
1983–1984	627.88	0.59		35 %	911
	854.07	0.58	−1.59	46 %	910
1984–1985	695.33	0.62		34 %	993
	901.37	0.61	−1.45	42 %	992
1985–1986	432.43	0.39		23 %	619
	620.85	0.38	−1.33	34 %	618
1986–1987	778.09	0.48		12 %	1012
	918.45	0.48	−0.99	14 %	1011
1987–1988	525.55	0.73		31 %	875
	689.26	0.72	−1.15	35 %	875
1988–1989	515.68	0.46		21 %	736
	750.35	0.45	−1.65	33 %	735
1989–1990	**354.84**	**0.30**		**23 %**	**499**
	503.14	**0.29**	**−1.05**	**36 %**	**498**
1990–1991	521.82	0.36		21 %	697
	642.24	0.36	−0.85	25 %	696
1991–1992	590.59	0.49		16 %	828
	847.79	0.48	−1.81	25 %	827
1992–1993	421.71	0.36		16 %	597
	693.86	0.35	−1.92	34 %	596
1993–1994	504.21	0.58		31 %	785
	768.15	0.57	−1.86	44 %	784
1994–1995	566.77	0.70		31 %	903
	601.28	0.70	−0.24	32 %	903
1995–1996	688.61	0.61		29 %	980
	925.34	0.60	−1.67	39 %	979
1996–1997	635.13	0.64		17 %	943
	1077.78	0.62	−3.12	34 %	941
1997–1998	543.22	0.73		29 %	897
	821.98	0.72	−1.96	39 %	896
1998–1999	648.11	0.43		9 %	856
	1099.23	0.42	−3.18	29 %	854
1999–2000	424.11	0.37		16 %	605
	690.94	0.37	−1.88	33 %	604
2000–2001	578.34	0.62		24 %	876
	978.52	0.60	−2.82	46 %	875
2001–2002	632.34	0.63		17 %	934
	843.34	0.62	−1.49	21 %	933
2002–2003	**1050.17**	**0.77**		**18 %**	**1423**
	1552.64	**0.76**	**−3.54**	**35 %**	**1421**
2003–2004	656.62	0.57		23 %	933
	714.85	0.57	−0.41	23 %	933
2004–2005	499.13	0.65		42 %	811
	615.34	0.64	−0.82	45 %	810

using elevation and longitude as the predictor variables for the three divisions of Crete, the analysis provided even more descriptive models.

Results reported by Naoum and Tsanis (2003) demonstrated that, for a typical dry year, the island of Crete would receive up to 800 mm of precipitation; for an average year, precipitation has a range between 800 and 1100 mm and finally for a wet year, the island would receive precipitation greater than 1100 mm. Based on these values, the range of elevation–precipitation gradient (b_1) is 0.26–0.49, 0.72–0.96 and 0.5–1 mm m^{-1} for minimum, maximum and average precipitation years, respectively. The range of longitude-precipitation gradient (b_2) for a minimum precipitation year is -2.03 to 0.23 mm km^{-1}, -8.20 to -0.86 mm km^{-1} for a maximum precipitation year and -3.23–0.93 mm km^{-1} for an average year. In brief, the results show that there is a general increase in the coefficient of determination from the one-variable equation to the two-variable equation for all years under study.

Following the multiple linear regression method, Table 4 summarizes the precipitation characteristics of 55 stations. There is clear indication that in terms of average values, the south-central part experienced 5 % less mean annual precipitation than the eastern part and 23 % less than the northern part. During the year 1989–1990, the south-central part was characterized by 7 % and 29 % less precipitation than the eastern and northern part, respectively. It is also confirmed that during 2002–2003 the eastern part received 13 % and 24 % less precipitation than the south-central part and northern part. After merging both south-central and eastern parts, considering the low difference percentages, there is obvious statistical evidence of the decreasing longitude-precipitation gradient. However, differences between stations with different elevations relate to the orographic effect, so stations with high elevation represent higher precipitation values.

Multiple linear regression results are graphically demonstrated with the generation of maps in GIS. Each cell of the grid represents a precipitation value according to the produced regression models and the DEM. Figure 3a indicates the spatial distribution of precipitation for a long-term 30-yr average, obtained by the multiple linear regression method. Taking into account the minimum precipitation year 1989–1990, the precipitation shortage was derived from the difference between the precipitation of a 30-yr average at each cell and the precipitation of the year 1992–1993 at each cell as well (Fig. 3b). Considering the maximum precipitation year 2002–2003, the precipitation excess was the result of the difference between the precipitation of a 30-yr average at each cell and the precipitation of the year 2002–2003 (Fig. 3c). During 1989–1990, the mean annual precipitation was affected by elevation less than the mean annual precipitation of the 30-yr average, as shown in Table 3 ($b_1^{\min} < b_1^{\text{long}-\text{term}}$). This explains the fact that the precipitation shortage for the "dry" year was greater at high elevations. Contrarily, for the
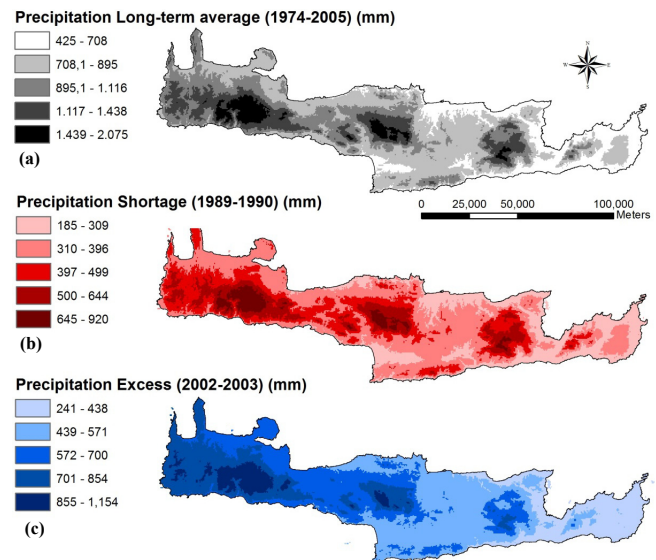

Fig. 3. Spatial distribution of **(a)** precipitation for a long-term average of the period 1974–2005, **(b)** precipitation shortage concerning the year 1989–1990 and **(c)** precipitation excess concerning the year 2002–2003.

"wet" year, elevation affected more the mean annual precipitation ($b_1^{\max} > b_1^{\text{long}-\text{term}}$). The downward gradient of precipitation from west to east was weaker for the "dry" year when compared to the 30-yr average ($|b_2^{\min}| < |b_2^{\text{long}-\text{term}}|$). On the other hand, during the "wet" year, precipitation decreased at a higher rate from west to east ($|b_2^{\max}| > |b_2^{\text{long}-\text{term}}| > |b_2^{\min}|$) and as a result the precipitation excess was higher in the western part. This denotes that the orographic effect was stronger during the wet year in the western part of Crete.

Regarding the spatial distribution of flood events over Crete (1990–2007), 66 % of them were reported for western Crete region, while the rest of the 34 % were reported for eastern Crete (Koutroulis et al., 2010). The east to west precipitation gradient can be attributed to the regional NW to SE dominant meteorological atmospheric patterns and the higher elevation and steepest slope morphology of western Crete in comparison to the eastern region of the island.

Long period characteristics, represented by 48-month time scale values of SN-SPI, were calculated for six (6) representative stations in order to provide an overview of prolonged drought occurrences in relation to the factor of elevation and longitude during the period 1974–2005. The results of drought analysis on the island of Crete during this period show a definite tendency towards prolongation and greater severity of drought episodes. Figure 4 illustrates the drought conditions of two representative stations in the north-western part of Crete. The period 1988–1997 was recorded to be a period of drought for both stations, with a greater intensity in Kalives. It is rather obvious that the sensitivity of precipitation variability is depicted in the index results. The

Table 3. Simple and linear regression for Crete and its parts for different years. The highest R^2 values are in bold.

	Spatial extent	Number of stations	b_0	b_1	b_2	R^2
Min precipitation year 1989–1990	Crete	55	354.84	0.30	–	23 %
		55	503.14	0.29	−1.05	36 %
	north	17	609.29	0.26	−1.36	26 %
	south-central	20	555.68	0.26	−2.03	35 %
	eastern	18	231.35	0.49	0.23	**79 %**
Max precipitation year 2002–2003	Crete	55	1050.17	0.77	–	18 %
		55	1552.64	0.76	−3.54	35 %
	north	17	1656.06	0.72	−3.98	17 %
	south-central	20	2034.33	0.85	−8.20	**51 %**
	eastern	18	1012.87	0.96	−0.86	39 %
Average precipitation year 1978–1979	Crete	55	572.08	0.68	–	40 %
		55	704.94	0.68	−0.93	43 %
	north	17	982.88	0.81	−3.23	64 %
	south-central	20	733.76	0.50	−1.78	49 %
	eastern	18	323.59	1.00	0.94	**70 %**
Long-term 1974–2005	Crete	55	609.33	0.55	–	29 %
		55	850.16	0.54	−1.70	41 %
	north	17	1040.35	0.68	−3.30	49 %
	south-central	20	1030.50	0.48	−3.90	42 %
	eastern	18	599.00	0.72	−0.35	**67 %**

SN-SPI behaves in a similar way to the SPI; however, a significant difference lies in the fact that the SPI is temporally comparable, but the SN-SPI is spatio-temporally comparable among different areas with different mean total annual precipitations. Obviously, the difference between the indices appears mainly at the peaks of the time series. The SN-SPI presents Palea Rumata as less dry during 1988–1997, and Kalives as less wet during 1980–1985 and 1999–2004. In the case of Palea Rumata (1267 mm), normalization took place during dry conditions (1988–1997) whereas in Kalives (742 mm) normalization was obvious during wet conditions (1980–1985, 1999–2004). SN-SPI values present Palea Rumata with mildly dry conditions and Petras with mildly wet conditions during the aforementioned periods. Stated in the simplest terms, the same SPI value occurs for different precipitation levels and, therefore, these stations cannot be compared via SPI. On the other hand, different SN-SPI values occur for different precipitation levels. The spatial pattern of precipitation of Kalives indicates the correlation between the low precipitation values of 1987–1989 and the SN-SPI negative peaks during 1991–1993 (as the 48-month index utilizes the precipitation total for 48 months).

The south-central stations experienced a drought period during 1988–1995, while Pompia was characterized by a mildly dry period during 1999–2002 (Fig. 5). Normalization took place for Gergeri during dry conditions and for Pompia during wet conditions. The severe drought spell of Pompia was a result of the low recorded precipitation during 1989–1990. Finally, as Fig. 6 shows, the eastern part of Crete experienced the most long-term drought period (Sitia, 464 mm) that signs its beginning in 1990 and ends in 2003 with no intervals. It is important to stress that Agios Georgios (1016 mm) had three drought periods and appeared the aforementioned drought period 8 yr earlier (1982). There was also a significant normalization for the wet periods for both stations. The precipitation variability revealed very low values for Sitia during the referring periods and a correspondence between the lowest peak of the index and the precipitation amount 48 months before. As a result, the additional drought period at the southern and eastern part was justified by the combination of the regional atmospheric patterns and the morphological variability among different parts of the island as stated by Koutroulis et al. (2010). Graphical examination of spatial evolution of drought confirms that the eastern part suffers more than the rest of the island from dry conditions. Generally speaking, for the 30-yr period, there are 26 stations with downward precipitation trend, 24 followed upward precipitation trend while 6 remained stable.

Similarities in time series of the SN-SPI across stations have been identified through SN-SPI drought months (SN-SPI < 0); this calculation comprised the main component for k-means cluster analysis with the use of MATLAB R2011b. The multiple correlation carried out, including the parameters of elevation and longitude for 55 stations, resulted in the generation of 2 optimum clusters (Fig. 7a). The definition of the number of the clusters a priori is a process that follows

Table 4. Characteristics of the stations used.

| Part of Crete	Stations	Mean annual precipitation (mm)	Elevation (m)	Longitude (km)	Min prec. year 1989–1990 Mean Precipitation	Max prec. year 2002–2003 Mean Precipitation	Average year 1978–1979 Mean Precipitation
North	Palea Rumata	1267	316	18.7	640	1738	1282
	Suda	614	152	52	356	930	608
	Kalives	742	24	53.9	519	1411	720
	Mouri	1108	24	64.7	695	1985	969
	Rethimno	670	5	82.2	493	1058	–
	Kavousi	1000	580	93.2	498	2189	1062
	Voleones	1238	260	91.9	716	1869	1149
	Spili	1182	390	87	807	1811	1087
	Vizari	778	310	102.3	519	841	773
	Gerakari	1336	580	95.3	602	2330	1235
	Anogia	1035	740	118.8	629	955	1220
	Krussonas	990	500	128.3	688	1448	965
	Iraklio	480	39	139.9	289	728	455
	Finikia	775	40	139.5	475	1555	655
	Metaksochori	749	430	142.1	486	1392	1130
	Voni	775	330	152.6	439	1256	920
	Profitis Ilias	816	380	139.3	518	1579	717
	Average	915	300	100	551	1475	902
South-central	Agios Kirillos	545	450	122	251	1224	602
	Kapetaniana	711	800	133.1	352	1442	785
	Lefkogia	863	90	79.8	458	1524	654
	Melabes	782	560	97.4	541	1359	783
	Agia Galini	623	20	101.7	365	1210	554
	Vorizia	1183	520	116.2	588	2162	1058
	Lagolia	573	140	110.7	345	1210	565
	Zaros	844	500	120.8	417	1603	795
	Agia Varvara	975	570	129.9	552	1500	937
	Gergeri	907	450	123.5	465	1700	772
	Partira	677	400	150.5	391	1156	633
	Asimi	607	200	139	350	972	718
	Vagionia	546	190	130	311	874	647
	Tefeli	737	360	144.9	487	1333	744
	Achentrias	715	680	151.4	459	1234	712
	Kalivia	601	200	151.3	357	1055	603
	Demati	480	210	156	196	987	537
	Moroni	709	400	121.1	363	1340	710
	Pompia	516	150	117.5	255	1135	414
	Pretoria	547	225	143.7	325	919	511
	Average	707	356	127	391	1297	687
Eastern	Kasteli	743	350	160.2	421	1200	873
	Armacha	824	450	161.7	514	1290	898
	Avdu	872	230	169.2	491	1739	1030
	Kassanoi	582	320	159.1	380	893	576
	Agios Georgios	1016	850	173.9	700	1137	–
	Kalo Chorio	540	20	196.5	389	911	482
	Malles	780	590	183.9	430	1234	920
	Kapsaloi	743	10	165.7	208	780	–
	Neapolis	810	240	184.8	477	933	890
	Exo Potamoi	1425	800	178.6	731	2406	1592
	Mithoi	584	200	184.2	311	945	594
	Ierapetra	439	10	197	235	705	445
	Pachia Ammos	592	50	203.8	301	1161	597
	Stavrochori	803	325	215.9	435	907	936
	Sitia	464	114	229.4	283	657	440
	Maronia	670	150	228.8	425	1075	713
	Katsidoni	914	480	233.1	510	1319	1173
	Paleokastro	540	25	244.5	290	940	556
	Average	741	290	193	418	1124	795
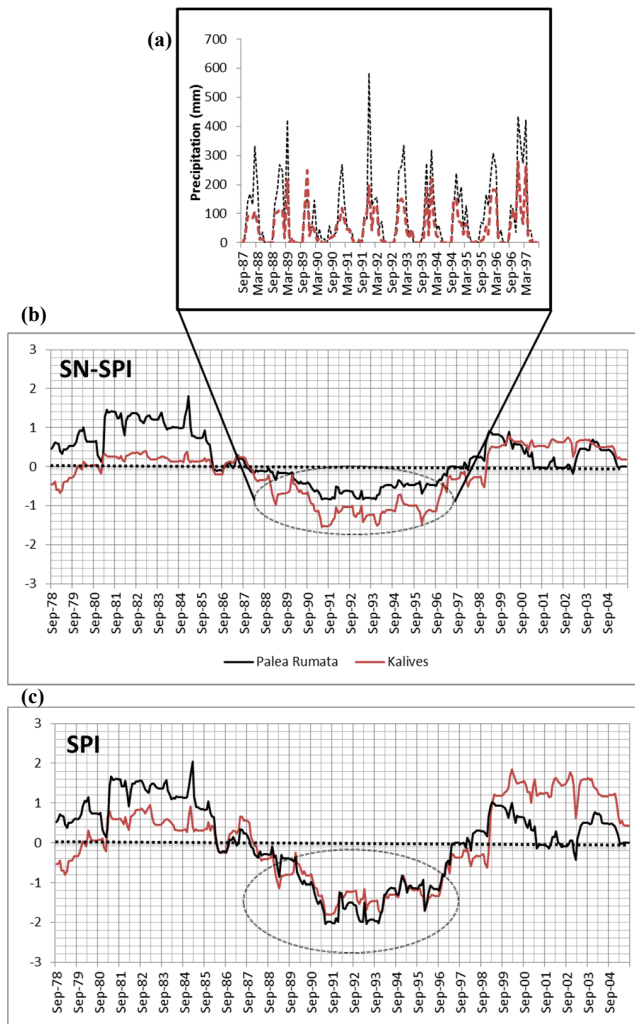
Fig. 4. (a) Monthly precipitation time series for drought events **(b)** 48-month time scale SN-SPI and **(c)** SPI for two representative stations of north-western Crete based on the period 1974–2005.

Fig. 5. (a) Monthly precipitation time series for drought events **(b)** 48-month time scale SN-SPI and **(c)** SPI for two representative stations of south-central Crete based on the period 1974–2005.

the minimization of the average silhouette width after several hypotheses (Fig. 7b). Then, the performance of multiple runs, each with a different set of randomly chosen initial centroids, led to the selection of the optimum set of clusters. The cluster (1), which includes low elevations (31 stations) corresponds to a higher value of drought months (162) in comparison with stations of high elevations (24 stations, 157 months, cluster 2), a fact that confirms the downward gradient of drought appearance as the elevation increases. The average silhouette width (0.72) indicates that a reasonable structure has been found. There is clear indication that cluster analysis results are connected to MLR results and justify the negative correlation between drought events and the most important factor in the bivariate model of the precipitation distribution, elevation.
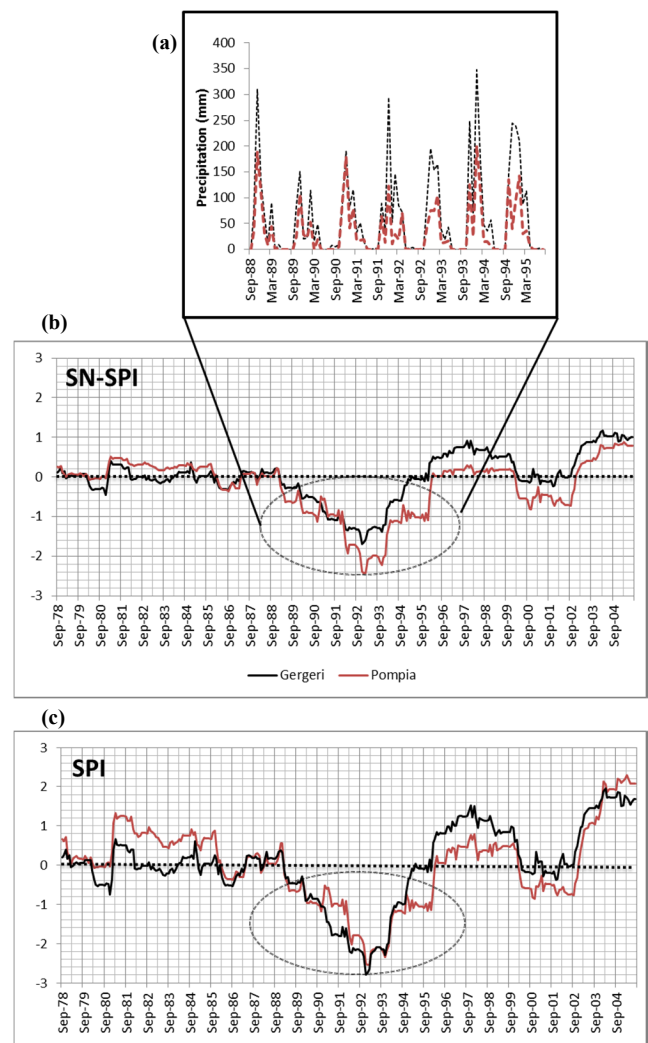
5 Conclusions

Results of spatio-temporal precipitation analysis for the period 1974–2005 for the island of Crete revealed valuable information within the context of drought occurrence. A positive rate between precipitation and elevation comprises a fact, whereas a negative correlation lies between precipitation and longitude. It is then deduced that amongst many topographic and geographic factors, elevation and longitude strongly determine the spatial association in precipitation variations.

The multiple linear regression method has been used to develop correlations to estimate the spatial distribution of orographic precipitation for a complex territory such as that of the island of Crete in Greece, using the parameters of elevation and longitude. The two-variable model is more reliable and realistic, especially when dealing with a relatively small
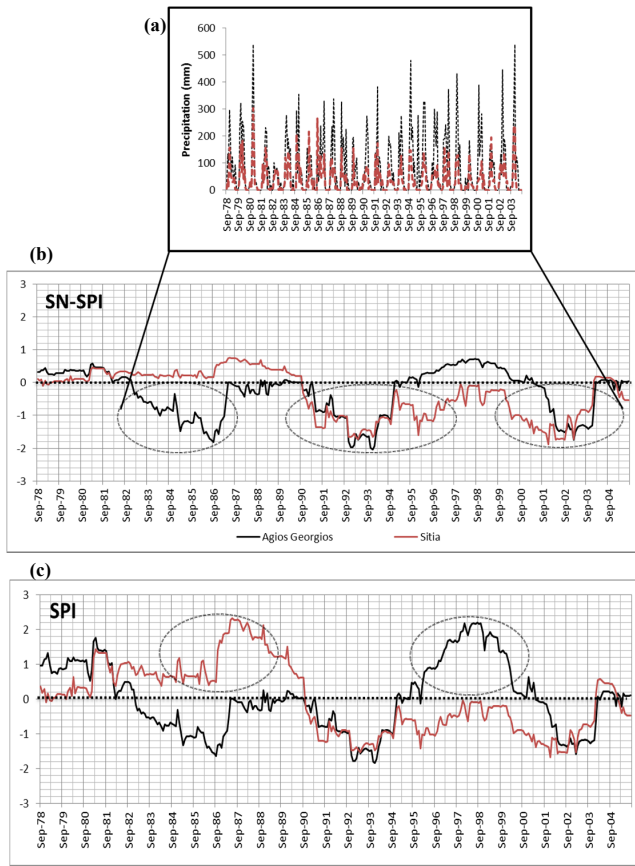
Fig. 6. (a) Monthly precipitation time series for drought events **(b)** 48-month time scale SN-SPI and **(c)** SPI for two representative stations of eastern Crete based on the period 1974–2005.

Fig. 7. Spatial separation of 55 stations using cluster analysis. **(a)** Stations plot illustrating drought months as the dependent variable and **(b)** silhouette plot of the clusters.

number of rain gauges. The coefficient of determination R^2 is lower when the one-variable model is used, whereas R^2 is higher for precipitation when both dependent variables, e.g. elevation and longitude, are used. Spatially, it was obvious that precipitation is of orographic type (precipitation is strongly correlated with elevation).

The SN-SPI was selected for the drought assessment as a variant of the common tool of drought assessment SPI, but represents a more suitable means of comparing drought conditions between neighbouring areas of differing precipitation heights, because the SN-SPI expands the meaning of the temporal character of drought to its spatial relativity. It appears that these new drought index values are reasonably comparable in their local significance, both in space and time, in contrast with SPI. The proposed methodology can evaluate the precipitation deficit and thus can become a practical tool for the assessment of regional drought events. Accordingly, the SN-SPI was successfully evaluated in 6 representative stations on the island of Crete for the period 1974–2005. The SN-SPI analysis revealed that Crete has faced a main drought period across the whole island and an additional
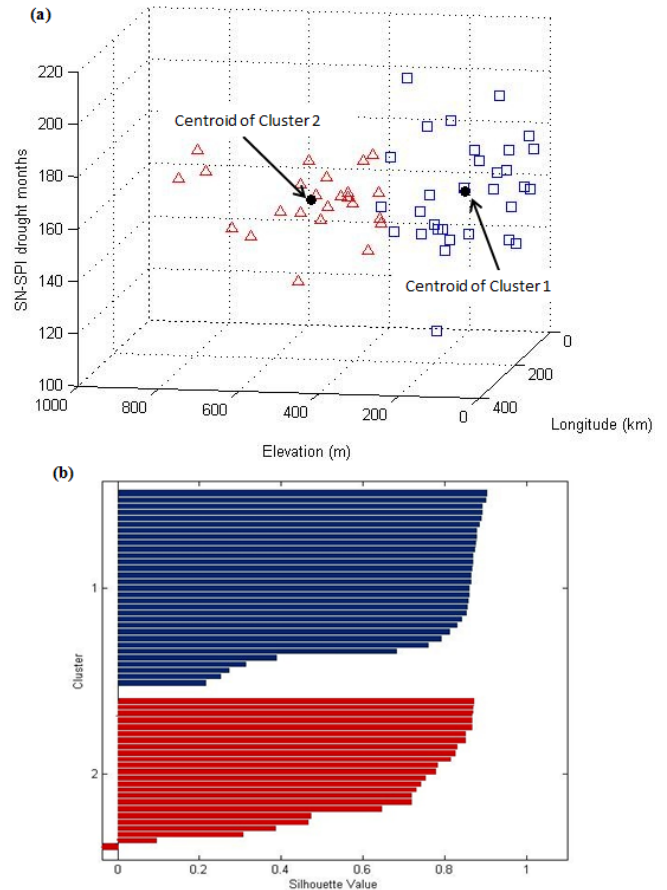
drought period at the southern and eastern parts. The additional drought period at the southern and eastern part is then justified by the combination of the regional atmospheric patterns and the morphological variability among different parts of the island. Cluster analysis confirmed the negative correlation between drought events and elevation, a fact that unifies the precipitation impact assessment on droughts. The selected clustering method proved to be effective, giving the opportunity to perform multiple runs until the optimum cluster set was achieved.

The statistical tools that were used proved to be very effective in the evaluation of the spatio-temporal variability of precipitation and accordingly of drought on the island of Crete. The combination of multiple methods, such as multiple linear regression, cluster and SN-SPI analysis offers a complete study that could be used in the context of climate change impact assessments.

Finally, the results signal an urgent need for the development of strategic water management and preparedness plans in all drought-prone areas in order to help mitigate most of the effects. Water resources could be seriously affected by

shorter rainy periods, with wide-ranging consequences for local human societies and ecosystems. The impact of these precipitation changes at the station level is required in order to develop strategies in long-term water supply and demand, and thus to attain sustainable water resource management. A specific policy framework on droughts is provided by the EU Water Framework Directive, which, in combination with the presented methodology, could be used for assessing the impact of climate change on water resources and, hence, for developing strategies dealing with the predictability of the phenomenon.

Edited by: R. Lasaponara
Reviewed by: two anonymous referees

References

Ali, A. and Lebel, T.: The Sahelian standardized rainfall index revisited, Int. J. Climatol., 29, 1705–1714, 2009.

Allamano, P., Claps, P., Laio, F., and Thea, C.: A data-based assessment of the dependence of short-duration precipitation on elevation, Phys. Chem. Earth, 34, 635–641, 2009.

Barry, R. G. and Chorley, R. J.: Atmosphere, weather, and climate, 8th Edn. Routledge, London, ISBN 0–415–27170–3, 2003.

Basist, A., Bell, G., and Meentmeyer, V.: Statistical relationships between topography and precipitation patterns, J. Climate, 7, 1305–1315, 1994.

Bonaccorso, B., Bordi, I., Cancelliere, A., Rossi, G., and Sutera, A.: Spatial Variability of Drought: An Analysis of the SPI in Sicily, Water Resour. Manag., 17, 273–296, 2003.

Chartzoulakis, K. and Psarras, G.: Global change effects on crop photosynthesis and production in Mediterranean: the case of Crete, Greece, Agr. Ecosyst. Environ., 106, 147–157, 2005.

Chavoshi, S. and Soleiman, W. N. A.: Delineating Pooling Group for Flood Frequency Analysis Using Soft Computing, Euro. J. Sci. Res., 35, 181–187, 2009.

Drogue, G., Humbert, J., Deraisme, J., Mahr, N., and Freslon, N.: A statistical-topographic model using an omnidirectional parameterization of the relief for mapping orographic rainfall, Int. J. Climatol., 22, 599–613, 2002.

Draper, N. R. and Smith, H.: Applied regression analysis, New York, John Wiley and Sons, Inc, 1998.

Dubrovský, M., Svoboda, M. D., Trnka, M., Hayes, M. J., Wilhite, D. A., Žalud, Z., and Hlavinka, P.: Application of relative drought indices in assessing climate change impacts on drought conditions in Czechia, Theor. Appl. Climatol., 96, 155–171, 2009.

Fu, T. C., Chung, F. L., Ng, V., and Luk, R.: Pattern discovery from stock time series using self-organizing maps, in: Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, 27–37, 2001.

Guan, H., Wilson, J., and Makhin, O.: Geostatistical mapping of mountain precipitation incorporating autosearched effects of terrain and climatic characteristics, J. Hydrometeorol., 6, 1018–1030, 2005.

Harris, D., Menabde, M., Seed, A., and Austin, G.: Multifractal characterization of rain fields with a strong orographic influence, J. Geophys. Res., 101, 26405–26414, 1996.

Hayward, D. and Clarke, R. T.: Relationship between rainfall, altitude and distance from the sea in the Freetown Peninsula, Sierra Leone, Hydrol. Sci., 41, 377–384, 1996.

Kalpakis, K., Gada, D., and Puttagunta, V.: Distance measures for effective clustering of ARIMA time series, in: Proceedings of the 2001 IEEE International Conference on Data Mining, San Jose, CA, 273–280, 2001.

Kaufman, L. and Rousseuw, P. J.: Finding groups in data: An introduction to cluster analysis, New York, John Wiley and Sons, Inc, 1990.

Koutroulis, A., Tsanis, I., and Daliakopoulos, I.: Seasonality of floods and their hydrometeorologic characteristics in the island of Crete, J. Hydrol., 394, 90–100, doi:10.1016/j.jhydrol.2010.04.025, 2010.

Koutroulis, A. G., Vrochidou, A.-E. K., and Tsanis, I. K.: Spatiotemporal Characteristics of Meteorological Drought for the Island of Crete, J. Hydrometeor., 12, 206-226, doi:10.1175/2010JHM1252.1, 2011.

Loukas, A. and Vasiliades, L.: Probabilistic analysis of drought spatiotemporal characteristics inThessaly region, Greece, Nat. Hazards Earth Syst. Sci., 4, 719–731, doi:10.5194/nhess-4-719-2004, 2004.

McKee, T., Doesken, N., and Kleist, J.: The relationship of drought frequency and duration to time scales. In Proceedings of the 8th Conference on Applied Climatology, Boston, Am. Met. Soc., 179–184, 1993.

McQueen, J. B.: Some Methods for classification and Analysis of Multivariate Observations. In Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1, 281–297, 1967.

Naoum, S. and Tsanis, I. K.: Temporal and spatial variation of annual rainfall on the island of Crete, Greece, Hydrol. Process., 17, 1899–1922, 2003.

Naoum, S. and Tsanis, I. K.: Orographic precipitation modeling with multiple linear regression, J. Hydrolo. Eng., 9, 79–102, 2004.

Oates, T., Firoiu, L., and Cohen, P. R.: Clustering time series with Hidden Markov Models and Dynamic Time Warping, in: Proceedings of IJCAI-99 Workshop on Neural, Symbolic and Reinforcement Learning Methods for Sequence Learning, Stockholm, Sweden, 17–21, 1999.

Ouyang, R., Ren, L., Cheng, W., and Zhou, C.: Similarity search and pattern discovery in hydrological time series data mining, Hydrol. Process., 24, 1198–1210, 2010.

Palmer, W.: Meteorological drought. Research Paper No 45, US Weather Bureau, Washington, DC, 1965.

Region of Crete: Sustainable management of water resources in Crete. Region of Crete Information Bull., 24 pp., 2002 (in Greek).

Rumman, M. A, Hiyasat, M., Sweis, G. J., and Sweis, R. J.: Assessment of droughts in Jordan: The Yarmouk and Zarqa basins, Manag. Environ. Quality, 20, 696–711, 2009.

Soltani, S. and Modarres, R.: Classification of Spatio -Temporal Pattern of Rainfall in Iran Using A Hierarchical and Divisive Cluster Analysis, J. Spatial Hydrol., 6, 1–12, 2006.

Struyf, A., Hubert, M., and Rousseeuw, P. J.: Integrating robust

clustering techniques in S-PLUS, Comput. Stat. Data An., 26, 17–37, 1997.

Tan, P. N., Steinbach, M., and Kumar, V.: Introduction to Data Mining. Chapter 8, Cluster Analysis: Basic Concepts and Algorithms, Michigan State University and University of Minnesota, 2006.

Tsakiris, G. and Vangelis, H.: Towards a Drought Watch System based on spatial SPI, Water Resour. Manag., 18, 1-12, 2004.

Tsakiris, G., Pangalou, D., Tigkas, D., and Vangelis, H.: Assessing the areal extent of drought, Water Resources Management, New Approaches and Technologies, European Water Resources Association, Chania, Crete-Greece, 2007.

Tsanis, I. K. and Naoum, S.: The effect of spatially distributed meteorological parameters on irrigation water demand assessment, Adv. Water Res., 26, 311–324, 2003.

Vicente-Serrano, S. M. and Beguería, S.: Estimating extreme dry-spell risk in the middle Ebro valley (Northeastern Spain): a comparative analysis of partial duration series with a General Pareto distribution and annual maxima series with a Gumbel distribution, Int. J. Climatol., 23, 1103–1118, 2003.

Wotling, G., Bouvier, Ch., Danloux, J., and Fritsch, J.-M.: Regionalization of extreme precipitation distribution using the principal components of the topographical environment, J. Hydrol., 233, 86–101, 2000.

Wu, H., Svoboda, M., Hayes, M., Wilhite, D., and Wen, F.: Appropriate application of the Standardized Precipitation Index in arid locations and dry seasons, Int. J. Climatol., 27, 65–79, 2007.

Zhang, H., Ho, T. B., and Lin, M. S.: A non-parametric wavelet feature extractor for time-series classification, in: Proceedings of the Eighth Pacific-Asia Conference on Knowledge Discovery and Data Mining, Sydney, Australia, 595–603, 2004.