PRINCIPAL COMPONENT ANALYSIS OF PRECIPITATION IN THESSALY REGION (CENTRAL GREECE)

D. STATHIS[1,*]
D. MYRONIDIS[2]

[1]*Department of Mountain Water Management and Control*
Faculty of Forestry and Natural Environment
Aristotle University of Thessaloniki, Greece
University Campus Box 268, 54124 Thessaloniki, Greece
[2]*Dept. of Forestry and Natural Environment Management*
Technological Educational Institute of Kavala
1st Km-Dramas Microhoriou, Drama, Greece, 66100

*to whom all correspondence should be addressed:
e-mail: dstatis@for.auth.gr

ABSTRACT

Characteristics of the precipitation pattern are determined for Thessaly region in central Greece by using the Principal Component Analysis (PCA). Mean monthly precipitation data from 75 meteorological stations are used covering a 35 year period (1960-1994). The first two principal components (PCs), which describe 87.7% of the total variance, are taken into consideration. It is demonstrated statistically that PC_1 is related to the frequency of the cyclones crossing the area, while PC_2 is related to the anticyclonic and non-gradient surface weather types and cold pools aloft. This means that PC_1 explains the annual amounts of precipitation and PC_2 the summer thundery rainfalls. The spatial distribution of precipitation is studied afterwards. Using the PCs scores, homogeneous precipitation sectors are determined over the territory of which the annual pattern of precipitation is discussed.

1. INTRODUCTION

Many disciplines (geography, hydrology, forest management, agriculture, ecology and others) use spatial information considering climatological data as a basis for the understanding of the processes they study. Precipitation is one of the main climatic elements whose distribution can be highly variable in space and time over a complex terrain. Accurate knowledge of the spatial and seasonal variations of long term mean precipitation is required for a variety of planning tasks in civil engineering, agriculture and forestry.

One of the fundamental problems that hydrology has to solve is the estimation of precipitation in an unmonitored site by using data from surrounding meteorological stations. This situation frequently occurs in mountainous terrains, where few stations are usually available and orographic effects can be large. An example of this type of area is the region of Thessaly, which is located in eastern-central Greece Peninsula (Figure 1). It consists of a plateau that is surrounded by semi-mountainous and mountainous areas. It is the basin of the Aheloos and the Pinios rivers and it covers an area of 13,929 km^2. This region has a great importance for the study of spatial precipitation as this is the mountainous basin of the biggest Greek rivers (Aheloos, Pinios), which supply water to a great part of the region. The study area has a mountainous character and its altitude varies from 0 m to 2,520 m.
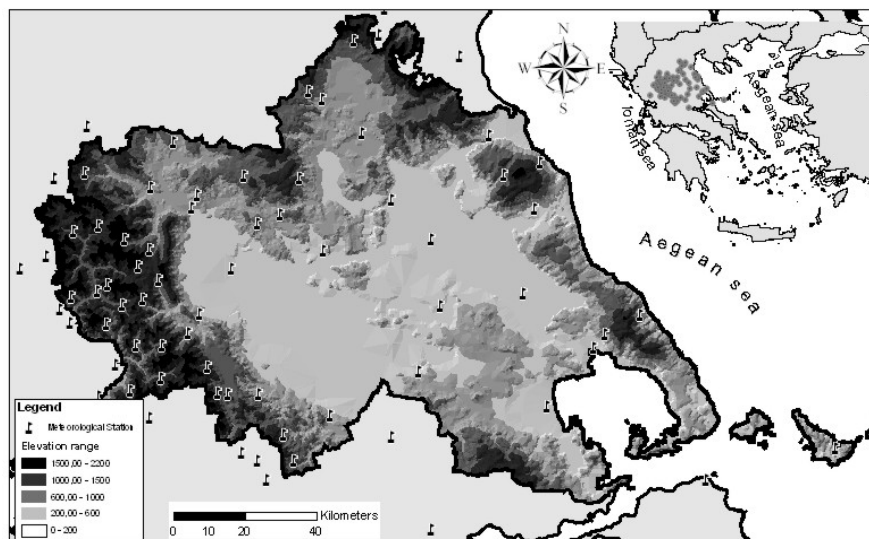
Figure 1. Location of the study area and the 75 meteorological stations

The precipitation regime of the region is Mediterranean (wet winters and dry summers) and it is very irregular at both intra-annual and inter-annual timescales. The main causes of this irregularity are the geographical location and the general atmospheric circulation. Moreover, local characteristics -such as the orography- play an equally important role (Katsoulis, 1986; Stathis *et al.,* 2005). The mean annual precipitation in the western mountainous area exceeds 2,000 mm, while the corresponding value in the central plain area comes up to only 420 mm (Flocas, 1994).

Many studies have contributed to the analysis of the precipitation variability over the Mediterranean and Greece. Precipitation studies have been performed for various time periods at regional and local scale. For example, (Maheras, 1983) has tried to classify the weather conditions over the eastern Mediterranean by defining 16 weather types. Recently, more objective methods are used and it is concluded that the precipitation regime in Greece can be explained in terms of variability of a small number of circulation type patterns (Maheras and Anagnostopoulou, 2003). Recent studies, which describe rainfall variability in the Greek region in relation to circulation types, have been carried out by (Maheras *et al.,* 2004; Tolika *et al.,* 2007).

The relationship between the atmospheric circulation and the environment of a region is considered to be crucial for the understanding of the regional climate and its changes. The influence of both the atmospheric circulation and the geographical factors on precipitation has been studied for a large number of Mediterranean and Greek stations (Maheras, 1985; Maheras and Kolyva-Mahera, 1990; Kassomenos *et al.*, 1998; Maheras *et al.,* 2000; Xoplaki *et al.,* 2000; Barzokas *et al.*, 2003; Feidas *et al.*, 2007).

In this study, an attempt is made to analyze the spatial structure of the present precipitation climate of an area, which was represented by only one or none station in previous studies. The aim of this work is firstly, to describe some characteristics of precipitation totals by using the method of Principal Component Analysis and secondly, to illustrate an objective method for the determination of homogeneous precipitation regions. This will contribute to a better understanding of the spatial and seasonal variability of the precipitation climate in a small region of Thessaly in central Greece. The results may be used in climatic models that are focused in this area of Greece as a source of surface precipitation parameters.

2. DATA AND METHODS
Precipitation data were obtained from 87 meteorological stations. Monthly totals of precipitation were used covering the period 1960-1994. This period was chosen because more than half of the stations there terminated their operation. Afterwards, these data were elaborated so as to obtain consistent homogeneous databases. All the selected stations had less than 5% of missing months for the specific period. A preliminary examination of the homogeneity of the precipitation data was first accomplished by selecting stations, where the measuring site has

not been altered during the period of the record. Double mass method was then used to adjust any heterogeneity of the precipitation records at each station (Stathis, 1998).

Double-mass analysis (Linsley *et al.*, 1988) examined the consistency of the record at a station by comparing its accumulated annual or seasonal precipitation with the concurrent accumulated values of mean precipitation for a group of surrounding stations. A change due to meteorological causes would not provoke a change in slope since all base stations would be similarly affected. The consistency of the record for each of the base stations had to be tested and records that showed inconsistency were dropped before other stations were tested or adjusted. Data from twelve (12) stations were identified as 'suspect' and as a result of this all stations were regarded as possessing an inhomogeneity problem. Furthermore, these stations were excluded from the analysis.

Additionally, two parametric statistical tests were applied in order to evaluate the homogeneity of the data for the remaining stations. The quality control was performed on monthly precipitation. Statistical analyses were performed so as to compare the mean values, the standard deviations and the distributions for every month between all stations by using the parametric, the Student's t and the Chi-Square tests. Details of these methods can be obtained from WMO (1986). The precipitation for the remaining 75 stations data were found to be homogeneous after running the previous tests of randomness. Figure 1 shows the location of the 75 stations. These stations are operated by the Public Power Corporation (31 meteorological stations), the Ministry of Agriculture (7 stations), the Ministry of National Works and Environment (19 stations), the Hellenic Meteorological Service (9 stations) and the National Agricultural Research Foundation (9 stations). The meteorological stations ranged in altitude from 3m to 1,180m, whle their density is one station for every 278 km^2.

Principal Component Analysis (PCA) is a multivariate statistical analysis, which attempts to simplify a complex set of interrelationships by creating one or more new variables, with respect to those that allow a more convenient examination of the overall spatial relationship. Furthermore, it attempts to explain the overall variance in a data set by isolating a number of components with respect to newly defined axes, each of which corresponds to a variable (Richman, 1986; Preisendorfer, 1988; Graham, 1988). The reader is referred elsewhere for a detailed exposition of these techniques since their description is lengthy (Kendall, 1980; Jolliffe, 1986; Richman, 1986). They have been applied in domains that use regionalization climatic variables with increasing frequency and they are important tools in meteorology/climatology (Wigley *et al.*, 1982; Sneyers and Goossens, 1988; Sneyers *et al.*, 1989; White *and Richman*, 1991; Prudhomme *and Reed*, 1999).

The monthly averages constitute a set of 12 variables, which were evaluated at 75 stations distributed throughout Thessaly region. These variables were correlated with one another. Their description in terms of a smaller number of uncorrelated variables simplified and reorganized the data in a way that made them more easily understood (Logue, 1984). Principal Component Analysis enabled this to be done. This was accomplished by identifying 12 linear combinations of the original variables, known as Principal Components, which are mutually uncorrelated and it determined the proportion of the total variance, which might be associated with each one of them. Those PCs, which explained only a small fraction of the variance, could then be neglected without much loss of information and the system might be described in terms of the remaining and more important Components. The Kaiser's criterion (Kaiser, 1959) was used to determine the statistically significant PC's. This criterion is probably the one most widely used. The final stage of the method consisted in the "rotation of the axes', a process necessary in order to achieve a better discrimination among the Components and, therefore, an easier physical interpretation of the results (Richman, 1986; Bartzokas *et al.,* 1994). These components (PC's) were further subjected to the orthogonal varimax rotation. Rotation of these hypothetical vectors often removes certain ambiguities that are sometimes evident in the direct solutions (White *and Richman*, 1991).

3. DISCUSSION
3.1 Physical explanation
In the first stage, the applied method (PCA) estimated the relation between variables, the relation between stations and the relation between variables and stations (Maheras, 1985) aiming at the PC's extraction that produced the most and more significant information. The

results of the analysis (eigenvalues, variances and PC loadings) are presented in Table 1. Kaiser's criterion indicated that the cut-off value for the eigenvalues was at the eigenvalue number 2. These first two components accounted for 87.7 % of the total variance of the rainfall whose 77.2% was extracted for the first component and 10.5% for the second one.

Table 1. Results of the Principal Component Analysis

A/A	Eigenvalue	Percent var	Cumul. percent	Loadings	
				PC-I	PC- II
1	9.26	77.2	77.2	0.92	0.30
2	1.26	10.5	87.7	0.93	0.29
3	0.53	4.4	92.1	0.90	0.30
4	0.24	2.0	94.1	0.85	0.41
5	0.22	1.9	95.9	0.61	0.68
6	0.17	1.4	97.3	0.61	0.70
7	0.11	0.9	98.3	0.06	0.94
8	0.10	0.8	99.1	0.47	0.78
9	0.04	0.3	99.4	0.75	0.33
10	0.03	0.2	99.6	0.92	0.18
11	0.03	0.2	99.9	0.88	0.39
12	0.02	0.1	100	0.93	0.31

The results of the analysis indicated that the colder months (September to April) were highly correlated with the first PC that explained 74.5% of the total variance, while the warmer months (May to August) were significantly correlated to the second PC. Thus, a close correlation exists between the Greek weather types and the rotated PC's. This can be supported by the simultaneous presentation of the frequencies of the annual course of the cyclonic and anticyclonic weather types over Greece (Maheras and Anagnostopoulou, 2003). The annual course of the cyclonic frequencies compared with the annual course of the first PC's loadings (Figure 2) showed a good agreement. Therefore, the first component can be related to the cyclonic precipitation type, which falls during the colder period over the country. This relation means that the PC_I explains the annual amounts of precipitation. The annual course of the frequencies of anticyclonic weather types compared with the annual course of the second PC's loadings (Figure 3) also showed a good agreement. Thus, the second component, which explains 10.5% of the total variance, is related to summer thunderstorm activity, which is intensified by the presence of a cold pool aloft.
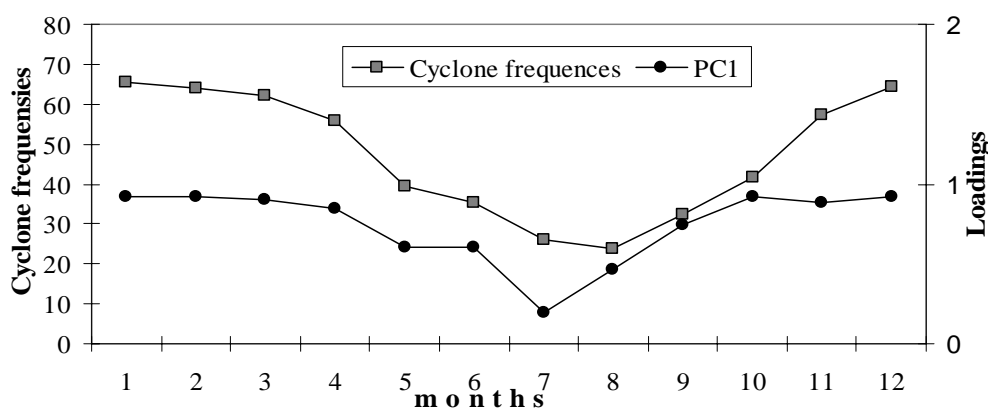

Figure 2. Monthly averages of the first standardised precipitation field varimax rotated (PC_1) and the annual course of cyclone frequencies
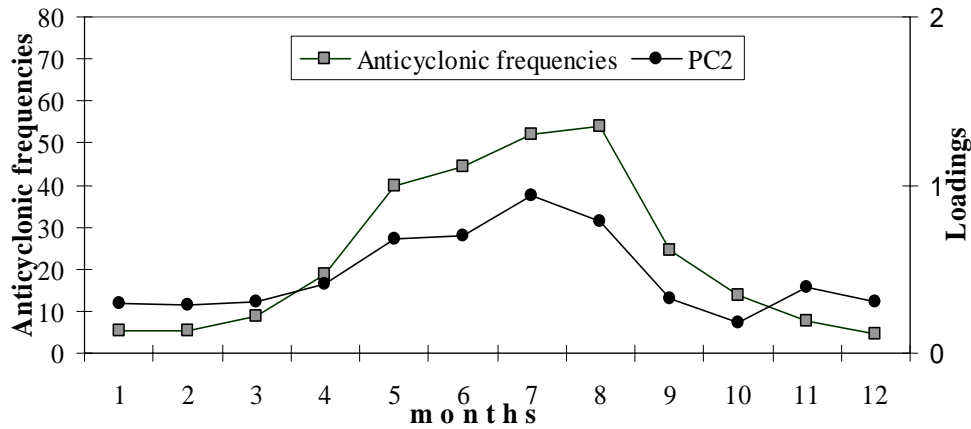
Figure 3. Monthly averages of the second standardised precipitation field varimax rotated (PC$_2$) and the annual course of anticyclonic frequencies

Synoptic conditions over the Balkans, especially over Greece during the warm months (May to July), reveal that a cold pool frequently appears (Balafoutis, 1991; Spanos *et al.*, 2003). This cold pool is well-recognised in the 1,000-500 hPa thickness level. It usually extends to the south as far as the 39th parallel and it is one of the main factors controlling the summer thunderstorm activity over an extensive part of the Balkans. This weather type is associated with clouds, thunderstorms and rain and is usually cut-off south of the main westerlies during the last stages of their life circle (Sioutas and Flocas, 2003, Spanos *et al.*, 2003). The situation usually occurs over the southeastern Aegean.

As far as it concerns the component scores, the first one presents positive scores mainly in the western and the eastern part of the studied area (Figure 4) and negative scores in the central plain area. Higher positive values mean higher amounts of annual precipitation and negative scores mean lower annual rainfall amounts. Thus, the first PC's scores divide the region in sectors with high annual rainfall amounts (positive areas C, D in Fig. 4) and low amounts (negative areas A, B). Comparing this map with the map of Figure 1, it becomes obvious that the stations with high scores are located in the higher mountainous areas. Consequently, a link between the atmospheric circulation and the topography is established.

The second component (Figure 5) presents positive scores in the western area and the northern part of Thessaly, where the higher mountains appear and negative scores mainly in the central and southern areas. The high positive scores denote that these areas have higher amounts of precipitation during the summer compared to the areas of negative scores, where a longer dry period is established. This distribution of the scores seems to be linked to the thundery showers during the warm period, as the summer rainfall mainly comes from the thunderstorm action, which is reinforced by the topography.
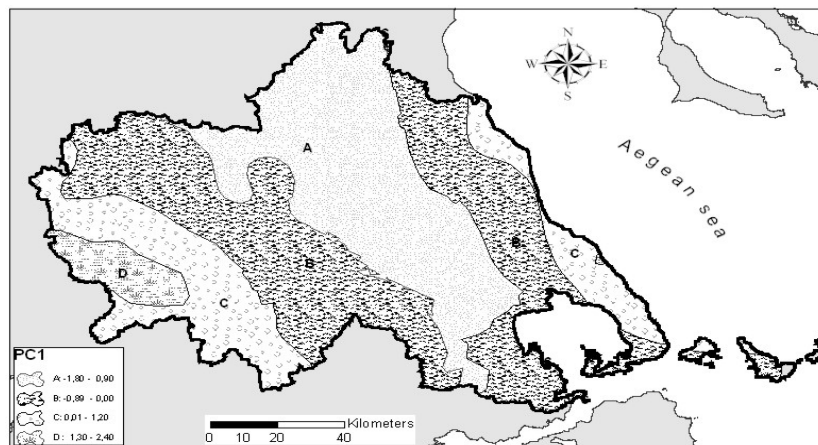

Figure 4. First Principal Component of the mean monthly precipitation in Thessaly
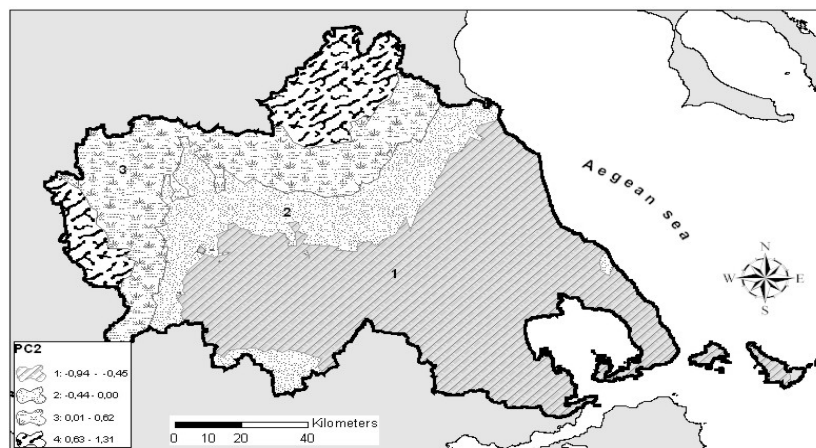
Figure 5. Second Principal Component of the mean monthly precipitation in Thessaly

3.2 Homogeneous precipitation sectors

An objective mapping and interpolation method were applied in order to determine the homogeneous precipitation groups of the stations in Thessaly by using Geographical Information System (GIS) techniques together with the first two PC's scores. The result is illustrated in Figure 6, where Thessaly region is divided into four sectors that constitute four homogeneous groups:

Group I: This group comprises stations with both positive scores. This means that these areas receive the highest annual amounts of precipitation (more than 1,700 mm). It appears in the western mountainous area of Thessaly, where the mountain chain of Pindos dominates and influences the precipitation climate.

Group II: This group is characterised by high annual precipitation amounts (more than 1,100 mm) and very little rain during summer. This group covers a variety of complex topographical areas in the eastern and the south-western region.

Group III: This group appears mainly in the northern areas having small amounts of precipitation (average annual totals 700 mm) and a very limited summer dry period due to thundery rainfall.

Group IV: This group is characterised by lower annual rainfall amounts (less than 700mm) in comparison with the previous groups and presents a pronounced dry summer period. It appears in the central plain and the semi-mountainous part of the region (far from the sea) with mountain heights less than 1,000 meters.

A representative course of the annual precipitation is shown separately for each group in Figure 7. The values are the monthly arithmetic means of the stations that are found in each group. This is a very good illustration of the dominating annual precipitation regime for each group of stations.
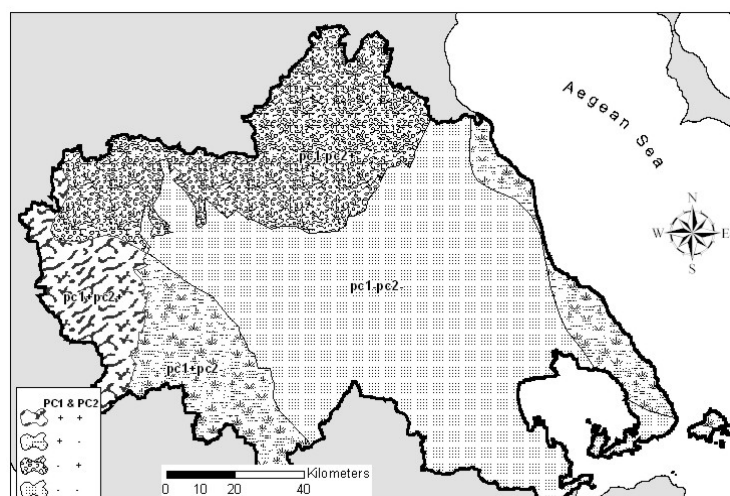

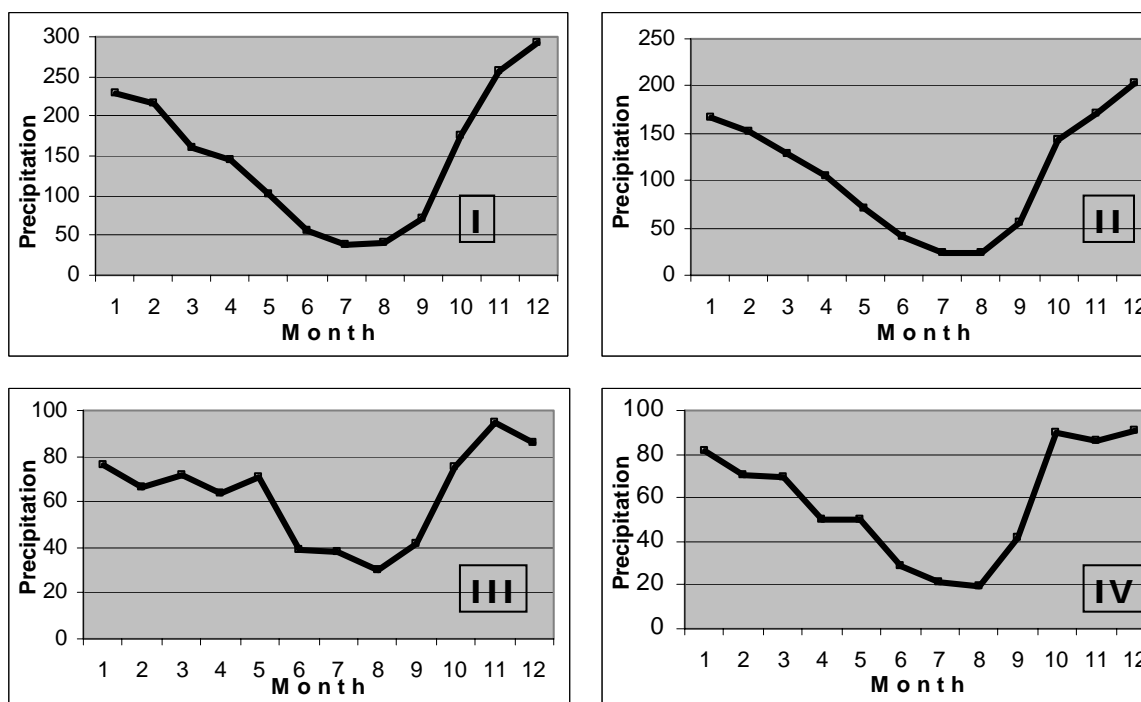*Figure 6.* The homogeneous precipitation groups in Thessaly

Figure 7. Annual course of the mean precipitation amounts for each group

The main pluviometric characteristics of the first group of stations are the following: a) Short dry period, b) the greatest amount of precipitation falls in December (Fig. 7). It includes 20 stations of high altitude (mean group's altitude 907 m). All stations are located in the western areas of Thessaly and in the windward sites of the Pindos mountain range. The mean annual amount of precipitation varies between 1,500 and 2,500 mm. These high amounts are mainly attributed to depression tracks, which usually follow a SW or NW path (Maheras, 1983; Flocas, 1988). Thus, the rain systems, which meet near the Ionian coast and the Pindos mountain range (Fig. 1) result in heavy orographic rainfalls over these areas (Baltas, 2007). Then, the drier air moves east, producing a rain shadow that affects this part of the region. The yearly distribution presents a simple profile (Fig. 7). The main maximum (December) occurs when there is the maximum of cyclonic activity in central and eastern Mediterranean. The main minimum occurs in July.

Fifteen stations with positive scores for the first component and negative scores for the second one belong in the second group. The mean annual amount of precipitation varies from 1,080 to 1,450 mm. Theses stations are situated in shorter altitudes (580-980 m) than those of group I. The annual distribution of precipitation is similar to the one in the first group (simple profile with the maximum in December and the minimum in July). This group covers a variety of complex topographical areas in east Aegean coast and south-western region.

The third group comprises 16 stations with negative scores for the PC1 and positive scores for the PC2. The mean annual amount of precipitation varies from 800 to 1,000 mm. In these stations, the yearly distribution of precipitation presents a double peak. Besides the main maximum in November, a secondary one appears in these stations during March or May due mainly to thunderstorms, which develop very often during this season in Mediterranean areas which located northern the 39^0 latitude (Maheras, 1985). Thessaly is the most southern end of the summer upper weather trough (500h Pa), which influences SE Europe. The typical characteristic of this trough all over SE Europe is the activity of the thermodynamic thunderstorms that predominate during the warm period and result in damages to the crops (Sioutas and Flocas, 2003; Spanos *et al.,* 2003).

Finally, the 24 stations with both PCs having negative scores are situated in the central plain area and at low altitudes. Those last stations are the driest with an annual amount of precipitation from 400 to 800 mm. The yearly distribution is simple. The main maximum takes place in October or in December.

4. CONCLUSIONS

The application of the Principal Component Analysis (T-mode) in the mean monthly values of observed precipitation in the Thessaly region during the period 1960-1994, revealed two components explaining 87.7 % of the total variance. The first component represents the annual cycle of precipitation and it may be linked to depression frequencies. The second one shows the maximum during the summer and seems to be linked to atmospheric instability during that period. Similar results were derived from the studies of (Maheras, 1985; Balafoutis, 1991; Fotiadi, 1999).

Principal Component Analysis can be of great help in defining groups of stations having relatively homogenous climatological characteristics. The applied method identified four groups of stations. The spatial distribution of precipitation is characterised by high values in the western areas and low values in the eastern region. The Pindos mountain range, which is located in the western region, constitutes a crucial factor in the spatial distribution of precipitation in central Greece. The intra-annual variation of the precipitation amount over the western Thessaly is characterised by a broad maximum in December. During this cold period, the maximum of cyclonic activity in central Mediterranean takes place. The depressions are usually formed along the northern coasts of Mediterranean Sea due to the orography and the strong surface baroclinicity (Trigo *et al.*, 1999; Flocas *et al.*, 2001; Maheras *et al.*, 2001) and they are enhanced during their movement.

The climatic study of small areas in regions characterized by complex topography services a particularly valuable agriculture and forestry purpose, as a mean of optimizing productivities (Galliani and Filippini, 1985). It is important to highlight that although several approaches have been designed to model precipitation at Mediterranean spatial scales, true spatial diversity is not well recorded (e.g. Thessaly region is represented by only one or none station). Therefore, these models show serious limitations when trying to identify spatial and temporal precipitation patterns in Mediterranean region, an area with high temporal and spatial climatic variability. Our results indicate that more detailed studies are required at local scales for appropriate spatial identification and regionalization of the precipitation in areas such as Thessaly.

Although this analysis was applied to precipitation in this region, it is easy to perform similar analysis on other variables of interest for either parameterization purposes or output from climatological models.

REFERENCES

Balafoutis C. (1991) Principal Component Analysis of Albanian rainfall, *Journal of Meteorology,* **16**, 90-96.

Baltas E. (2007) Spatial distribution of climate indices in northern Greece, *Meteorological Application,* **14**, 69-78.

Bartzokas A, Metaxas A. and Gana S. (1994) Spatial and temporal sea-surface temperature covariances in the Mediterranean, *International Journal of Climatology,* **14**, 201-213.

Bartzokas A., Lolis CJ. and Metaxas DA. (2003) A study on the inter-annual variation and the spatial distribution of precipitation amount and duration over Greece on a 10 day basis, *International Journal of Climatology,* **23**, 207–222.

Feidas X., Noulopoulou Ch., Makrogiannis T. and Bora-Senta E. (2007) Trend analysis of precipitation time series in Greece and their relationship with circulation using surface and satellite data: 1955–2001, *Theoretical and Applied Climatology,* **87**, 155-177.

Flocas A. (1988) Frontal depressions over the Mediterranean Sea and central southern Europe. *Mediterranee,* **4**, 43-52.

Flocas A. (1994) Meteorology and Climatology, Ziti, Thessaloniki, 465pp

Fotiadi AK., Metaxas DA. and Bartzokas A. (1999) A statistical study of precipitation in NW Greece, *International Journal of Climatology,* **19**, 1221–1232.

Galliani G. and Filippini F. (1985) Climatic clusters in a small area, *Journal of Climatology,* **5**, 487-501.

Graham S. (1988) Precipitation: Process and Analysis, Wiley, 455pp.

Jolliffe I.T. (1986) Principal Component Analysis, Springer-Verlag, NewYork, 271.

Kaiser H.F. (1959) Computer program for varimax rotation in factor analysis, *Psychological Measurement,* **19**, 413-420.

Kassomenos P., Flocas H., Lykoudis S. and Petrakis M. (1998) Analysis of mesoscale patterns in relation to synoptic conditions over an urban Mediterranean Basin, *Theoretical and Applied Climatology,* **59**, 215-229.

Katsoulis B. (1986) The effect of topography on the precipitation regime; the case of the mountain ranges crossing Greece from north to south, *Rivista di Meteorologia Aeronautica,* **XLVII(2)**, 125-133.

Kendall, M. (1980) Multivariate Analysis, Griffin, London.

Linsley, RK., Kohler, MA. and Paulhus, HL. (1988), Hydrology for Engineers, SI Metric Edition, McGraw-Hill Book Co: Singapore, 512pp

Logue J. (1984) Regional variations in the annual cycle of rainfall in Ireland as revealed by Principal Component Analysis, *Journal of Climatology,* **4**, 597-607.

Maheras P. (1983) *Climatologie de la mer Egee et de'ses marges contiinentals*, These d'Elat Atelier de Reproduction de These, Universite Lille III, 783pp.

Maheras P. (1985) A factorial analysis of Mediterranean precipitation, *Archives of Meteorological and Geophysical Bioclimatology* Series B, **36**, 1-14.

Maheras P. and Kolyva-Mahera F. (1990) Temporal and spatial characteristics of annual precipitation over Balkans in the twentieth century, *International Journal of Climatology,* **10**, 495–504.

Maheras P., Patrikas I., Karacostas Th. and Anagnostopoulou Chr. (2000) Automatic classification of circulation types in Greece: methodology, description, frequency, variability and trend analysis, *Theoretical and Applied Climatology,* **67**, 205-223.

Maheras P. and Anagnostopoulou, C. (2003) Circulation types and their influence on the interannual variability and precipitation changes in Greece. In: Mediterranean Climate Variability and Trends (ed. Bolle HJ), pp. 215–239. Springer, Berlin.

Maheras P., Tolika K., Anagnostopoulou C., Vafiadis M., Patrikas I. and Flocas H.A. (2004) On the relationships between circulation types and changes in rainfall variability in Greece. *International Journal of Climatology*, **24**, 1695-1712.

Preisendorfer R.W. (1988) Principal Component Analysis in Meteorology and Oceanography, Elsevier, New York, 425 pp.

Prudhomme C. and Reed D.W. (1999) Mapping extreme rainfall in a mountainous region using geostatistical techniques: A case study in Scotland, *International Journal of Climatology,* **19**, 1337-1356.

Richman M.B. (1986) Rotation of Principal Components, *Journal of Climatology,* **6**, 293–335.

Sioutas M. and Flocas A.H. (2003) Hailstorms in Northern Greece: synoptic patterns and thermodynamic environment, *Theoretical and Applied Climatology,* **75**, 189-202.

Sneyers R. and Goossens C. (1988) L'analyse par la methode des composantes principales, application a la Climatologie et a la Meteorologie.OMM., Institul Royal Meteeorologique de Belgique, 63pp.

Sneyers R, Vandiepenbeeck M. and Vanlierde R. (1989) Principal Component Analysis of Belgian Rainfall, *Theoretical and Applied Climatology,* **39**, 199-204.

Spanos S., Maheras P., Karacostas T. and Pennas P. (2003) Objective Climatology of 500-h Pa Cyclones in Central and East Mediterranean Region during Warm-Dry Period of the Year, *Theoretical and Applied Climatology,* **75**, 167-178.

Stathis D. (1998) Meteorological features of Pindos from hydrological point of view*, Ph.D. Thesis, Aristotle University of Thessaloniki, Thessaloniki, Greece.

Stathis D., Ivanova D., Balafoutis Ch. and Makrogiannis T. (2005) Orographic effect on heavy rainfall in Chalkidiki peninsula (Greece) induced by a Mediterranean cold front: a case study on 7 to 8 of October 2000, *Croatian Meteorological Journal,* **40**, 490-493.

Trigo I.F., Davies T.D. and Bigg G.R. (1999) Objective climatology of cyclones in the Mediterranean region, *Journal of Climate,* **12**, 1685–1696.

Tolika K., Maheras P., Vafiadis M., Flocas H.A., Arseni-Papadimitriou A. (2007) Simulation of seasonal precipitation and raindays over Greece: A statistical downscaling based on Artificial neural networks. *International Journal of Climatology,* **27**, 861-881.

Xoplaki E., Luterbacher J., Burkard R., Patrikas I. and Maheras P. (2000) Connection between the large scale 500 hPa geopotential height fields and precipitation over Greece during winter time, *Climate Research,* **14**, 129–146.

White D. and Richman M. (1991) Climate regionalization and rotation of Principal Components, *International Journal of Climatology,* **11,** 1-25.

Wigley T.M.L., Lough M.J. and Jones P.D. (1982) Spatial patterns of precipitation in England and Wales and a revised, homogenous England and Wales precipitation series, *Journal of Climatology,* **4**, 1-25.

WMO (1986) *Guidelines on the Quality Control of Surface Climatological Data,* Geneva, WCP-85.