Theor Appl Climatol (2010) 100:105–121
DOI 10.1007/s00704-009-0162-5

ORIGINAL PAPER

Rainfall uncertainty in the Mediterranean: Intraseasonal rainfall distribution

Hadas Reiser · H. Kutiel

Received: 8 September 2008 / Accepted: 23 April 2009 / Published online: 25 July 2009

Abstract The rainfall distribution within the rainy season has crucial implications on a variety of disciplines. According to one approach of analyzing the intraseasonal rainfall distribution, it is essential to examine the date of different accumulated percentage (DAP hereafter), as presented in Paz and Kutiel (Isr J Earth Sci 52:47–63, 2003). The present study identifies various intraseasonal temporal distributions of rainfall, in 41 stations within the Mediterranean basin. Furthermore, classifications of these distributions according to their time, yield, and length are presented. The accumulated percentage was calculated for each Julian day for every available year in all stations. A correlation matrix between every possible pair of years, in each station, was calculated, and a cluster analysis (average linkage method) was performed. Finally, the averages of the entire dataset and the average of every cluster were compared in order to classify the clusters by using three parameters: timing represented by $DAP_{(25\%,\ 50\%,\ 75\%)}$ annual rainfall total and the rainy season length (RSL). Between 2 and 5 different types of clusters, with various probabilities, were defined for every station. Out of 132 overall clusters, which were found in 41 stations, the most frequent type (cluster 1) was the median in all three parameters. There were 16 clusters identified as short in their RSL, and 18 were identified as having a long classification. There were 19 dry clusters, and only eight were identified as wet. As for the parameter of timing, 39 clusters were classified as early and 38 as late. One conclusion of this study was that the probability of a dry year is higher than a wet one, and likewise, the probability of a long year is higher than of a short one.

Abbreviations

AP	accumulated percentage
CA	cluster analysis
DAP	date of accumulated precentage
DRT	daily rainfall threshold
RSBD	rainy season beginning date
RSED	rainy season ending date
RSL	rainy season length
RUEM	rainfall uncertainty evaluation model
SAD	starting analysis date
TOTAL	annual rainfall total

1 Introduction

The rainfall distribution within the year is one of the basic characteristics of the rainfall regime in a certain location. It has crucial implications on a variety of disciplines, regardless of the rainfall total (TOTAL hereafter). The temporal rainfall accumulation varies from year to year. Therefore, the same exact TOTAL in two different years, but with different temporal distributions, will have a different effect on the hydrological budget of the area. For example, if in a station, characterized by a Mediterranean climate, the same TOTAL in 1 year fell in December–January–February and in another in February–March–April with exactly the same daily distribution, the impacts on a

H. Reiser (✉) · H. Kutiel
Laboratory of Climatology,
Department of Geography and Environmental Studies,
University of Haifa,
Haifa 31905, Israel
e-mail: hreiser@geo.haifa.ac.il

variety of systems, such as water resources, ecosystems, regional agriculture management, and economic and social development, may be completely different due to the different PET rates, antecedent soil moisture, etc. in the different months.

Studies of rainfall variability have tended to focus on inter-annual time scales, while the intraseasonal scale has been relatively unexplored. Most of the researchers try to find consistency patterns in order to improve the predictabilities or to identify changes in climate (e.g., Maheras 1988; García-Oliva et al. 1991; Türkeş 1996; Romero et al. 1998; Rodriguez-Puebla et al. 1998; Lázaro et al. 2001; Rodrigo 2002; Tomozeiu et al. 2002; Norrant and Douguédroit 2005).

Everywhere in the world a few months tend to be wetter than others, even in regions where there are no defined wet or dry seasons, and rainfall is rather evenly distributed throughout the year, as in northwestern Europe, for example (Logue 1984), all the more so in a Mediterranean climate, where the wet and dry seasons are very distinguishable. Ramos (2001) characterized the Mediterranean climate as a complex pattern of seasonal variability, with large and unpredictable rainfall fluctuations from 1 year to the other. The variation in the mean annual rainfall did not follow a consistent trend in northeastern Spain, where dry, normal, and wet years have not been regularly distributed over time. However, during the most recent decades, less annual variability (from 1 year to the other) has been observed. The number of years with normal rainfall TOTAL has increased, but rainfall has been more irregularly distributed throughout the year.

According to one approach of analyzing the intraseasonal rainfall distribution, it is essential to examine the date of different accumulated percentage (DAP hereafter), as presented in Paz and Kutiel (2003). These are dates in which any percentile of the TOTAL was accumulated. For instance, $DAP_{(25\%)}$ is the date when the first quartile of the annual rainfall was accumulated. Similarly, the $DAP_{(50\%)}$ is the mid-season date or the date in which half of the annual rainfall was accumulated and so on.

On a given date in a certain year, for example, 75% of the TOTAL could accumulate, while in another year, only less than 25% of the TOTAL was accumulated at that date. In this example, the first year represents an early rainy season, whereas the second represents a late one. Furthermore, the same accumulated percentage (AP) may be reached at a variety of dates, spreading over several months (Paz and Kutiel 2003). Thus, the rainy season can be characterized according to these dates; for example, seasons are characterized according to abundant early season, abundant late season, and so on.

The purposes of the present study are:

1. To identify various intraseasonal temporal distributions of rainfall in different stations within the Mediterranean basin.
2. To characterize each temporal distribution in every station, identified in 1, their yield, and length.

2 Study area and data

The study area is located in the Mediterranean basin, between longitudes 10° W and 40° E and latitudes 30° and 46°N, excluding the North African coasts (Fig. 1). Forty-one rain stations were selected to represent the various rainfall regimes within the study area, not all having a typical Mediterranean climate.

Daily rainfall data within the period of 1931 and 2006 were obtained from the European Climate Assessment & Dataset. A dataset of at least 36 years was used for each station (Table 1). The selection of stations was based on the availability of the entire datasets without missing data (with very few exceptions, see Table 1), and therefore, no further process of completing data was necessary. Data was tested for homogeneity and was proven to be homogenous.

3 Methodology

The data were analyzed by a specially designed statistical model, entitled rainfall uncertainty evaluation model (RUEM). The RUEM was developed by the authors at the Laboratory of Climatology in the Department of Geography and Environmental Studies, University of Haifa. Description and more details can be found in the following website: http://geo.haifa.ac.il/~geoweb/RUEM%204.pdf.

Figure 2 presents a schematic flow chart illustrating the various steps in the process of rainfall data analysis:

1. Starting analysis date (SAD) determines the date on which the analyses start in each station as defined and detailed by Reiser and Kutiel (2008) and listed in Table 1.
2. The daily rainfall threshold (DRT) in each station was determined to exclude the lowest rain days that contribute altogether 4% of the annual rainfall. Table 1 presents the DRT (mm) for each station as suggested by Reiser and Kutiel (2009a).
3. The determination of the rainy season beginning and ending dates filters out sporadic rain events at the beginning and/or the end of the rainy season, regardless of their amount. These parameters were calculated

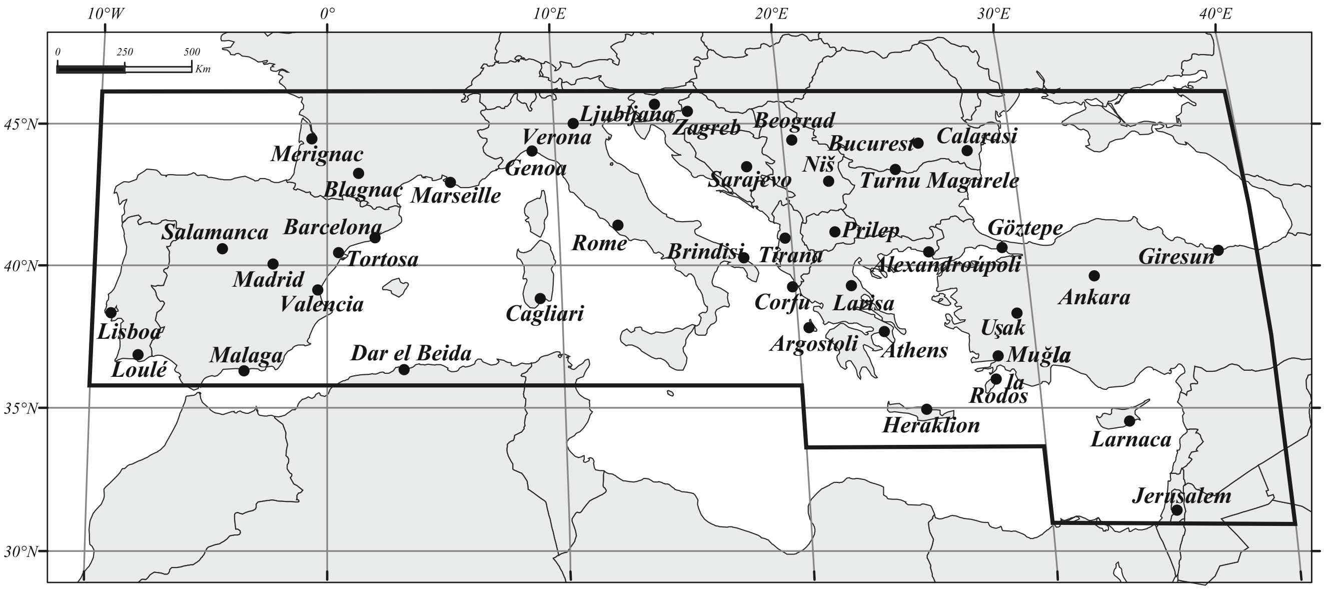

Fig. 1 The location of rain stations and the study area

separately for each year according to the distribution of the AP and in turn determined the rainy season length (RSL; Reiser and Kutiel 2009a).

These steps have already been done and reported in details in the two abovementioned studies.

The methodology used in the present study is demonstrated on data from the station of Madrid as an example. The selection of Madrid does not imply anything regarding its representation of the Mediterranean climate.

3.1 Cluster analysis

The APs were calculated for each Julian day, thus creating an annual course for every available year in all stations. A matrix was established, in which each column represents an annual course in a different year and each row a different Julian day. A correlation matrix between every possible pair of years in each station was calculated, and a CA (average linkage method) was performed. This methodology has been used previously mainly for clustering stations with similar characteristics into coherent regions (e.g. Tennant and Hewitson 2002). In the present study, however, it is used for merging different *annual course* in each station into groups of years according to similarities in their temporal course of AP.

The distance from one step to the following in the CA were calculated and averaged. The distance indicates the coherence within the grouped years in each cluster and the inconsistency between one cluster and the others. Years with very similar annual courses, and hence highly correlated, were clustered first at a very short distance (the vertical axis in Fig. 3a). As the differences between years or clusters increased, they were clustered at longer distances. The distance sums the squared differences of each of the variables in question, also known as the squared Euclidian distance. All distances are standardized, i.e., transformed to variables with mean=0 and variance=1. Once the distance between one step and the next one was higher than twice the average distance, the process was stopped and the obtained clusters were retained.

The dendrogram in Fig. 3a illustrates an example of the clustering process in Madrid. One can see that the first clustering steps occurred at short distances. The differences between one step and the next tended to increase as the process advanced. As long as the difference between one step and the following is short, it means that adding years to a cluster, or joining clusters, do not reduce considerably the coherence. When this difference is large, it means that the added years/clusters reduced considerably this coherence. The aim is to stop the clustering process at the shortest distance in order to ensure the highest coherence on the one hand and detect the main clusters in order to characterize the main temporal rainfall distributions patterns on the other hand.

The straight horizontal line in Fig. 3b represents twice the average difference (10.04). Step 49 occurred at a difference of 25.22 from step 48, and therefore, the process was stopped at step 48. At this step cluster 2 was created by the union of a cluster containing 7 years (1954, 1985, 1981, etc.) and a cluster build up of 5 years (1958, 1995, 1962, etc., see Fig. 3a). This step occurred at a total distance of 49.2, and the analysis retained five clusters.

At this stopping step, 52 years out of 56 (92.9%) were grouped into five different clusters: cluster 1 grouped 25 years (44.6%), cluster 2 grouped 12 years (21.4%),

Table 1 Main characteristics of rain stations

Country	Station name	Longitude (deg)	Latitude (deg)	Available data	SAD	DRT (mm)	RSL (days)
Albania	Tirana	19 47E	41 20N	1946–1986	July 1st	2.5	280
Algeria	Dar el Beida	03 15E	36 43N	1940–2006	July 1st	2.0	222
Bosnia	Sarajevo	20 60E	42 70N	1931–2005	February 1st	1.5	334
Croatia	Zagreb	15 59E	45 49N	1931–2006	March 1st	2.0	315
Cyprus	Larnaca	33 38E	34 55N	1931–2006	July 1st	2.0	158
France	Blagnac	01 23E	43 37N	1947–2006	February 1st	1.5	330
	Marseille	05 23E	43 18N	1931–2004	July 1st	2.0	272
	Merignac	00 41W	44 50N	1947–2006	July 1st	1.5	313
Greece	Alexandroúpoli	25 55E	40 51N	1958–2000	July 1st	2.0	261
	Athens	23 40E	38 03N	1955–2006	July 1st	1.5	200
	Argostoli	20 29E	38 11N	1958–2000	July 1st	2.5	209
	Corfu	19 55E	39 37N	1955–2006	July 1st	3.0	229
	Heraklion	25 11E	35 20N	1955–2006	July 1st	2.0	189
	Larisa	22 27E	39 39N	1955–2006	July 1st	1.5	283
	Rodos	29 10E	36 23N	1958–2000	July 1st	2.5	163
Israel	Jerusalem	35 13E	31 46N	1950–2004	July 1st	2.5	127
Italy	Brindisi	17 56E	40 38N	1956–2004	July 1st	1.5	255
	Cagliari	09 03E	39 14N	1951–2006	July 1st	1.0	234
	Genoa	09 00E	44 24N	1949–1996	July 1st	3.5	284
	Rome	12 35E	41 47N	1951–1998[d]	July 1st	2.0	272
	Verona	10 52E	45 23N	1951–2006	March 1st	2.0	304
FYROM	Prilep	21 57E	41 33N	1949–1995	February 1st	1.0	324
Portugal	Lisboa	09 09W	38 43N	1947–2006	July 1st	2.0	232
	Loulé[b]	08 00W	37 15N	1931–2005	July 1st	2.5	203
Romania	Bucurest	26 06E	44 25N	1931–2006	March 1st	1.5	314
	Calarasi	27 20E	44 12N	1938–2006	March 1st	1.5	313
	Turnu Magurele	24 53E	43 45N	1938–2000	March 1st	1.0	319
Serbia	Beograd	20 28E	44 48N	1936–2006	March 1st	1.5	324
	Niš	21 54E	43 20N	1949–2006	March 1st	1.0	329
Slovenia	Ljubljana	14 31E	46 04N	1950–2006	March 1st	2.5	310
Spain	Barcelona[c]	02 09E	41 23N	1965–2001	February 1st	2.0	294
	Madrid	03 41W	40 25N	1950–2006	July 1st	1.5	271
	Malaga	04 29W	36 40N	1943–2006	July 1st	2.5	199
	Tortosa	00 29E	40 49N	1946–2006	February 1st	1.5	279
	Salamanca	05 28W	40 56N	1945–2006	July 1st	1.0	288
	Valencia	00 21W	39 28N	1938–2006	July 1st	1.5	265
Turkey[a]	Ankara	32 53E	39 57N	1948–2003[e]	July 1st	1.0	282
	Giresun	38 23E	40 55N	1948–2003	July 1st	2.0	312
	Göztepe	29 05E	40 58N	1948–2003	July 1st	1.5	264
	Muğla	28 22E	37 13N	1948–2003	July 1st	3.5	194
	Uşak	29 24E	38 41N	1948–2003	July 1st	1.5	246

The SAD, the DRT and the RSL are based on Reiser and Kutiel (2008, 2009a) respectively

[a] Data obtained from Turkish State Meteorological Service (TSMS)

[b] Data obtained from Laboratorio Nacional de Engenaria Civil (LNEC), Portugal

[c] Data obtained from University Politecnica de Catalunya

[d] The season of 1978/9 was omitted due to lack of data

[e] The seasons of 1964/5 and 1965/6 were omitted due to lack of data

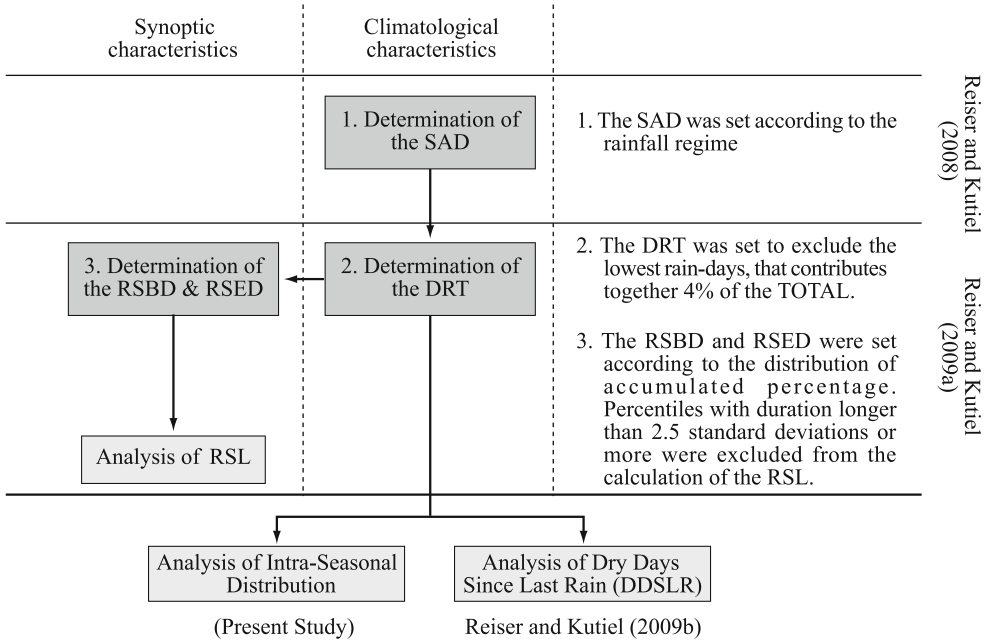

Fig. 2 A schematic flow chart representing the various steps needed for the analysis of the rainfall regime

cluster 3 grouped 6 years (10.7%) and clusters 4 and 5 grouped 5 and 4 years (8.9% and 7.1%, respectively). The remaining 4 years (1970, 1972, 1974, and 1980; 7.1%) were not clustered at this step.

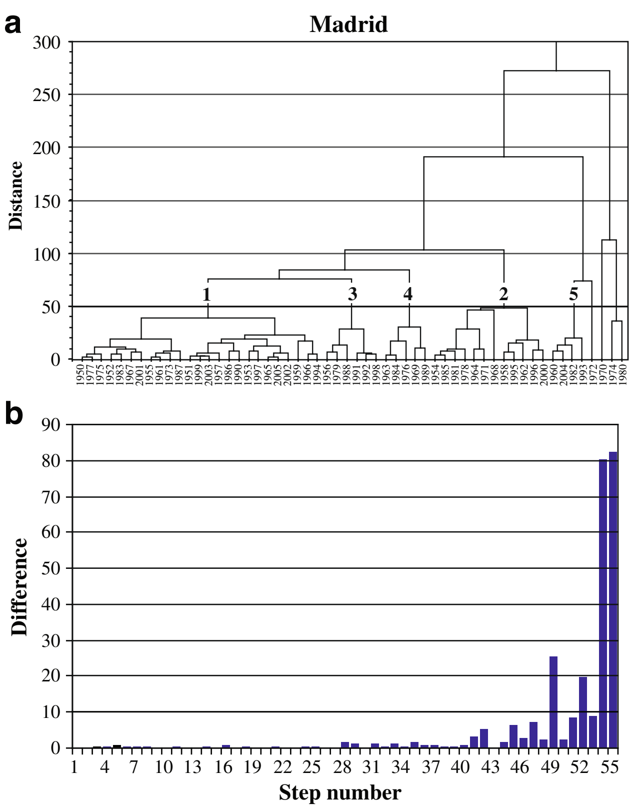

Fig. 3 The clustering dendrogram (case number 1=1948, 2=1949 and so on) (**a**) and the distribution of the difference between one step to another (**b**) in Madrid

To complete the analysis process, two conditions were set:

1. A cluster was defined only when it comprised of at least 2 years and not less than 5% of all years.
2. For each station, a minimum of two clusters and no more than five clusters were retained. This condition was made in order to assure that there would be enough information to characterize the variety of the temporal distributions of the AP in each station on the one hand. On the other hand, it was meant not to present too many clusters with very small differences between them and that were hardly identifiable one from the other.

The clusters are numbered according to their probabilities. The probability of a cluster is defined as the number of years included in that cluster divided by the total number of analyzed years. Cluster 1 was defined to be the cluster, which grouped the largest number of years, while cluster 5 the smallest number.

The clusters were obtained only on the basis of the annual distribution of the AP. Their characterization was defined on the basis of three parameters: their accumulation rate represented by DAPs (25%, 50%, and 75%), their TOTAL, and RSL.

3.2 Characterization of the cluster

3.2.1 Characterization according to the DAPs

The average timing of $DAP_{(25\%)}$, $DAP_{(50\%)}$, and $DAP_{(75\%)}$ of the entire dataset and their standard deviations were calculated. Correspondingly, the average annual AP courses of the years in *each cluster* were calculated.

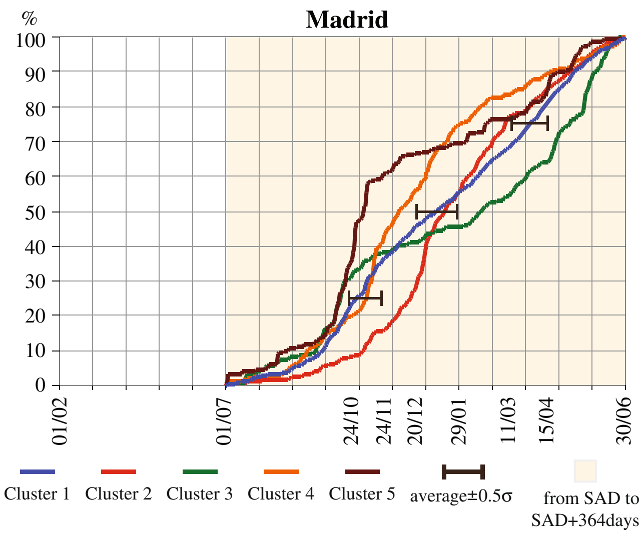

Fig. 4 Five courses of the AP according to the cluster analysis results in Madrid

A cluster was characterized as late when at least two out of the three of the above DAPs occurred at a certain date, which was at least 0.5σ *later* than the average date of the relevant DAP.

Similarly, a cluster was characterized as early when at least two out of the three of the above DAPs occurred at a certain date, which was at least 0.5σ *earlier* than the average date of the relevant DAP.

A cluster was characterized as median in two cases:

1. When at least two out of the three above DAPs occurred within the average date±0.5σ of the entire dataset
2. When each DAP was classified differently in each of the three percentiles (e.g. $DAP_{(25\%)}$=early, $DAP_{(50\%)}$= median, $DAP_{(75\%)}$=late).

The selection of a deviation ≥|0.5σ| was done in order to differentiate between clusters that behave similarly to the normal distribution, from clusters differing significantly from this distribution. Increasing this threshold to, e.g. ≥|1.0σ| would cause that the majority of clusters would be regarded as median, and only the very extreme cases would remain out of this classification. On the other hand, reducing this threshold, say to ≥|0.25σ| would cause that most clusters would be classified as extreme and would differ from the normal. Therefore, ±0.5σ seems to be a reasonable selection.

Figure 4 presents the five clustered annual AP distributions in the case study of Madrid. The time scale represented by the horizontal axis is longer than a year. However, the shadow represents the period from the SAD (July 1st in Madrid) until SAD+364 (June 30th of the following year), which is the analyzed year in Madrid (Reiser and Kutiel 2008). The average $DAP_{(25\%)}$±0.5σ of the entire dataset, representing the median period for $AP_{(25\%)}$, is between October 24th and November 24th. Similarly, the average period for ±0.5σ $AP_{(50\%)}$ is between December 20th and January 29th, and for $AP_{(75\%)}$, the period is between March 11th and April 15th. It can be seen that all three DAPs (25%, 50%, and 75%) of cluster 1 fall within these average periods ±0.5σ of each DAP, and therefore, cluster 1 was characterized as median. The same information is illustrated in Fig. 5a.

The annual course of cluster 2 demonstrates a slower accumulation rate at the beginning of the season (up to 40%). $DAP_{(25\%)}$ was reached only on December 16th, much later than the average ±0.5σ and therefore was characterized as late. However, in the second half of the season [between $DAP_{(50\%)}$ and $DAP_{(75\%)}$], the accumulation rate increased. $DAP_{(50\%)}$ was obtained within the median period, on January 17th, and $DAP_{(75\%)}$ on March

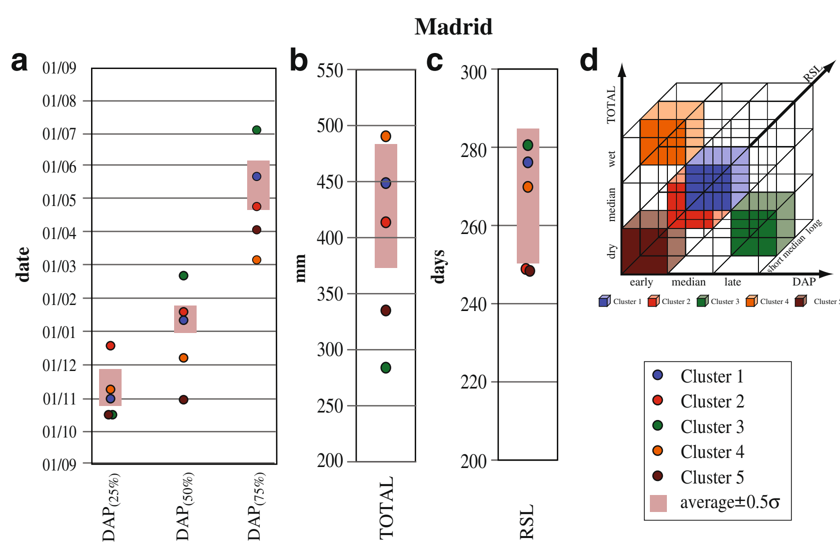

Fig. 5 The average $DAP_{(25\%, 50\%, 75\%)}$ ±0.5σ of the entire dataset and the $DAP_{(25\%, 50\%, 75\%)}$ of every cluster (**a**), the average TOTAL ±0.5σ for the entire dataset and the clusters' average TOTAL (**b**), the average RSL ±0.5σ for the entire dataset and the clusters' averages RSL (**c**), and a schematic illustration for the definitions of the five clusters, according to the average TOTAL, RSL, and DAP (**d**) in Madrid

Table 2 The number of clusters, percentage of clustered, and non clustered years and the probability of every cluster in all stations

	Number of clusters	Clustered years (%)	Non clustered years, percent (*n*)	Probability (%)				
				Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Bucurest	2	96.0	4.0 (3)	89.3	6.7			
Rodos		95.2	4.8 (2)	88.1	7.1			
Dar el Beida		95.5	4.5 (3)	87.9	7.6			
Calarasi		92.6	7.4 (5)	86.8	5.9			
Loulé		94.6	5.4 (4)	86.5	8.1			
Zagreb		93.3	6.7 (5)	85.3	8.0			
Malaga		90.5	9.5 (6)	82.5	11.1			
Argostoli		90.5	9.5 (4)	81.0	9.5			
Corfu		92.2	7.8 (4)	76.5	15.7			
Larisa		90.2	9.8 (5)	76.5	11.8			
Rome		91.3	8.7 (4)	65.2	26.1			
Verona	3	96.4	3.6 (2)	81.8	7.3	7.3		
Valencia		98.5	1.5 (1)	80.9	10.3	7.4		
Sarajevo		91.9	8.1 (6)	78.4	8.1	5.4		
Tortosa		98.3	1.7 (1)	78.3	13.3	6.7		
Uşak		96.5	3.5 (2)	76.4	10.9	9.1		
Blagnac		94.9	5.1 (3)	74.6	10.2	10.2		
Cagliari		98.2	1.8 (1)	74.5	16.4	7.3		
Prilep		91.3	8.7 (4)	71.7	13.0	6.5		
Turnu Magurele		98.4	1.6 (1)	69.4	22.6	6.5		
Merignac		88.1	11.9 (7)	66.1	13.6	8.5		
Beograd		91.4	8.6 (6)	58.6	22.9	10.0		
Genoa		97.8	2.2 (1)	57.4	29.8	10.6		
Marseille		91.8	8.2 (6)	56.2	20.5	15.1		
Tirana		92.7	7.3 (3)	56.1	22.0	14.6		
Jerusalem		87.0	13.0 (7)	46.3	33.3	7.4		
Larnaca	4	94.7	5.3 (4)	70.7	12.0	6.7	5.3	
Niš		93.0	7.0 (4)	66.7	15.8	5.3	5.3	
Giresun		92.7	7.3 (4)	63.6	12.7	9.1	7.3	
Alexandroúpoli		97.6	2.4 (1)	61.9	16.7	11.9	7.1	
Salamanca		91.8	8.2 (5)	60.7	13.1	9.8	8.2	
Lisboa		94.9	5.1 (3)	54.2	23.7	10.2	6.8	
Athens		92.2	7.8 (4)	52.9	25.5	7.8	5.9	
Heraklion		94.1	5.9 (3)	41.2	23.5	21.6	7.8	
Muğla		90.9	9.1 (5)	40.0	38.2	7.3	5.5	
Brindisi		89.6	10.4 (5)	33.3	25.0	20.8	10.4	
Ankara	5	94.3	5.7 (3)	64.2	9.4	9.4	5.7	5.7
Ljubljana		87.5	12.5 (7)	48.2	16.1	8.9	7.1	7.1
Madrid		92.9	7.1 (4)	44.6	21.4	10.7	8.9	7.1
Barcelona		97.2	2.8 (1)	44.4	13.9	13.9	13.9	11.1
Göztepe		98.2	1.8 (1)	43.6	30.9	12.7	5.5	5.5

Stations are sorted first according to the number of clusters in an ascending order, then according to the probability of cluster 1, in a descending order. Valencia and Jerusalem are emphasized since they are presented in the text as examples

12th. Thus, cluster 2 was characterized as median despite its slow accumulation rate at the beginning of the year.

In cluster 3, the accumulation was fast in the first quarter of the season, when the $DAP_{(25\%)}$ was obtained as early as October 10th. Further on, the accumulation rate decreased, and both $DAP_{(50\%)}$ and $DAP_{(75\%)}$ were late. Thus, this cluster was characterized as late.

Finally, at least two out of the three DAPs, in both clusters 4 and 5, were earlier than the median period of the entire dataset, and thus, the characterizations for both clusters were early.

3.2.2 Characterization according to the TOTAL

The averages TOTAL of the entire dataset and the standard deviations were calculated. The TOTAL for every cluster was averaged and compared to the average TOTAL of the whole series in order to define the clusters as follows:

A cluster was characterized as wet when the average TOTAL was at least 0.5σ *above* the average TOTAL of the entire dataset.

Similarly, a cluster was characterized as dry, when the average TOTAL was at least 0.5σ *below* the average TOTAL of the entire dataset.

A cluster was characterized as median, when the average TOTAL was within the average ±0.5σ TOTAL of the entire dataset.

3.2.3 Characterization according to the RSL

Similarly, the averaged RSL of the entire dataset and the standard deviations were calculated, and the average RSL of every cluster was compared to it and the characterizations are as follows:

A cluster was characterized as long when the average RSL was at least 0.5σ *longer* than the average RSL of the entire dataset.

Similarly, a cluster was characterized as short when the average RSL was at least 0.5σ *shorter* than the average length of the entire dataset.

A cluster was characterized as Median when the average RSL was the average ± 0.5σ length of the entire dataset.

In the case study of Madrid, the averages TOTAL ±0.5σ and the averages RSL ±0.5σ for the entire dataset are illustrated in Fig. 5b and c, respectively. The averages TOTAL and RSL of all clusters are also presented. As can be seen, the TOTAL range is between 371.9 and 482.1 mm, and

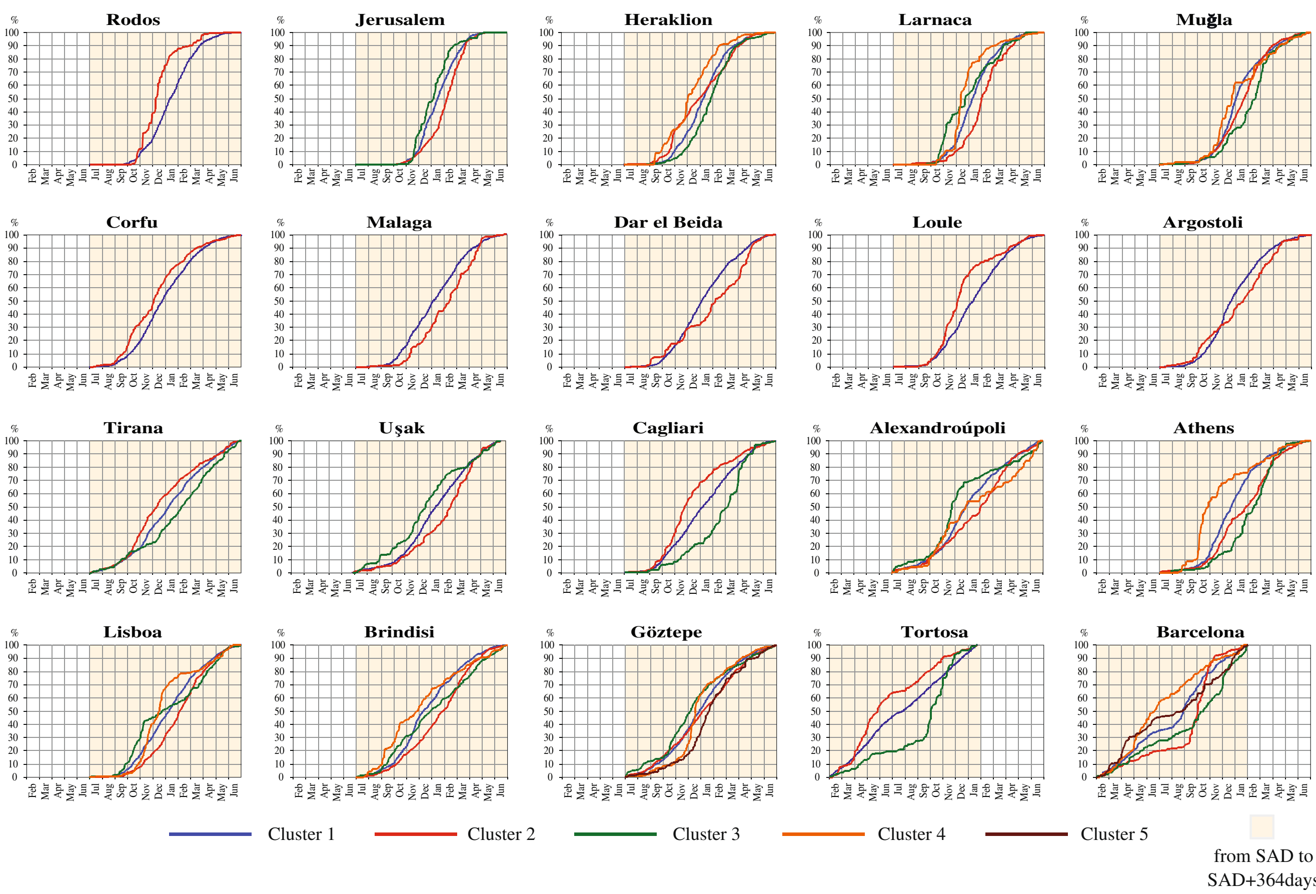

Fig. 6 The clustered annual courses in all stations

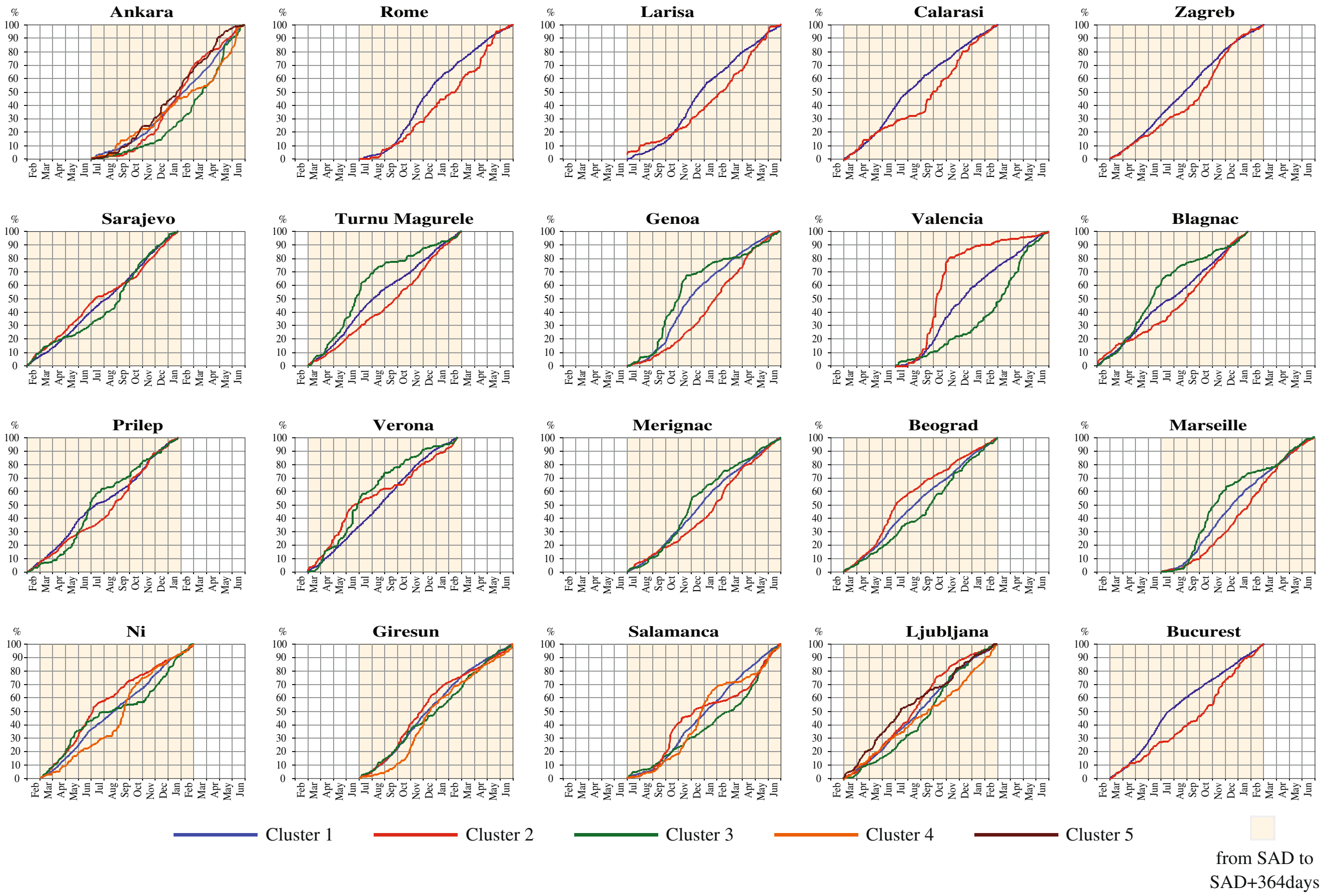

Fig. 6 (continued)

the RSL range is between 251 and 283 days. The departure of cluster 1 in both the TOTAL and the RSL are within these ranges (449.5 mm and 278 days, respectively). Therefore, cluster 1 is classified as median for these parameters.

Cluster 2 is classified as median in the TOTAL with 413.7 mm and short in the RSL with 246 days. Cluster 3 has a dry classification with only 336.7 mm on average, and the classification for the RSL is median (with 281 days). In cluster 4, the TOTAL is 489.1 mm, and it is classified as wet. However, the RSL is 269 and thus classified as median. Finally, cluster 5 is classified as dry for the TOTAL (345.7 mm) and short with only 245 days.

In the final analysis process, all three characteristics (the DAPs, the TOTAL, and the RSL) were combined in each cluster. Figure 5d presents a three-dimensional schematic illustration of this combination for the five clusters found in the case study. It can be seen that Madrid demonstrates five different types of annual rainfall courses:

1. Cluster 1 (blue) represents a Median course in all three parameters with the highest probability of more than twice in 5 years.
2. Cluster 2 (red) represents a median course in its DAPs and TOTAL and a short feature in its RSL, with a probability of once every 5 years.
3. Cluster 3 (green) represents a dry and late course, with a median length and with a probability of once every 9 years.
4. Cluster 4 (orange) represents a wet and early course, with median length and a probability of once in 11 years.
5. Cluster 5 (brown) represents a dry, early and short course with a probability of once in 14 years.

Once in 14 years, on the average, the annual course does not fit in any of the above five clusters.

This methodological process was applied on all stations.

The spatial distribution of the stations according to their intraseasonal accumulated course was found using the same methodology of CA as detailed above. For this purpose, the annual course of *cluster 1*, which is the most frequent cluster in all stations, was analyzed and mapped.

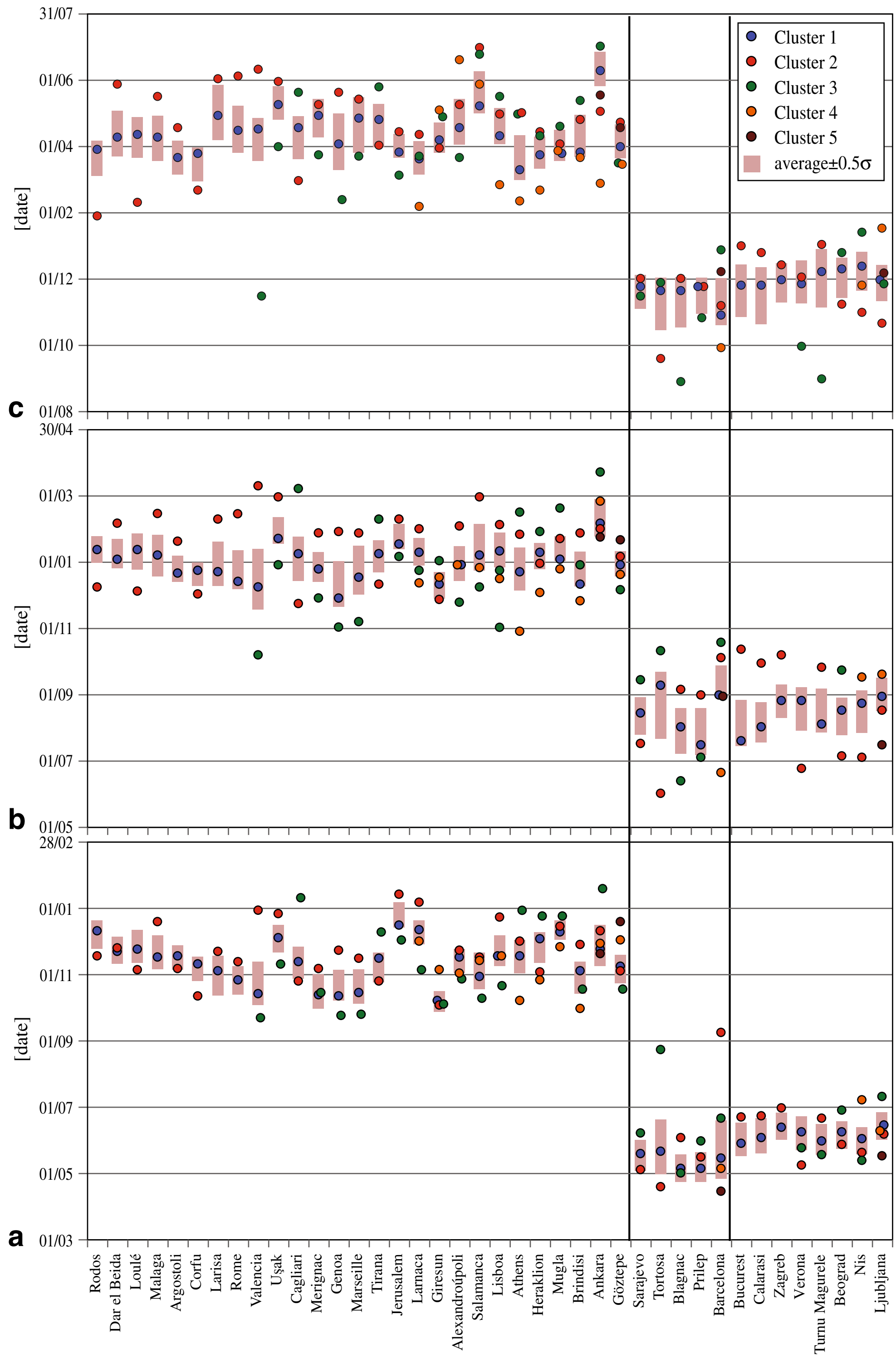

Fig. 7 The average $DAP_{(25\%)} \pm 0.5\sigma$ of the entire dataset and the $DAP_{(25\%)}$ of every cluster (**a**). The average $DAP_{(50\%)} \pm 0.5\sigma$ of the entire dataset and the $DAP_{(50\%)}$ of every cluster (**b**) and the average $DAP_{(75\%)} \pm 0.5\sigma$ of the entire dataset and the $DAP_{(75\%)}$ of every cluster in all stations (**c**). Stations are presented according to their SAD as recommended in Reiser and Kutiel (2008)

4 Results and discussion

Table 2, summarizes the number of clusters as a result of the clustering process and the percentage of the clustered and non-clustered years in all stations. Stations are sorted first according to the number of clusters in an ascending order, then according to the probability of cluster 1 in a descending order. In 11 stations, the available years were grouped into two clusters, in 15 stations into three clusters, in ten stations into four clusters, and in the remaining five stations, five clusters were found. In all stations on average, 93.6% of the years were grouped into clusters (two to five). In Valencia, 98.5% of the years (67 out of 68 years) were clustered, whereas the lowest percentage of years was

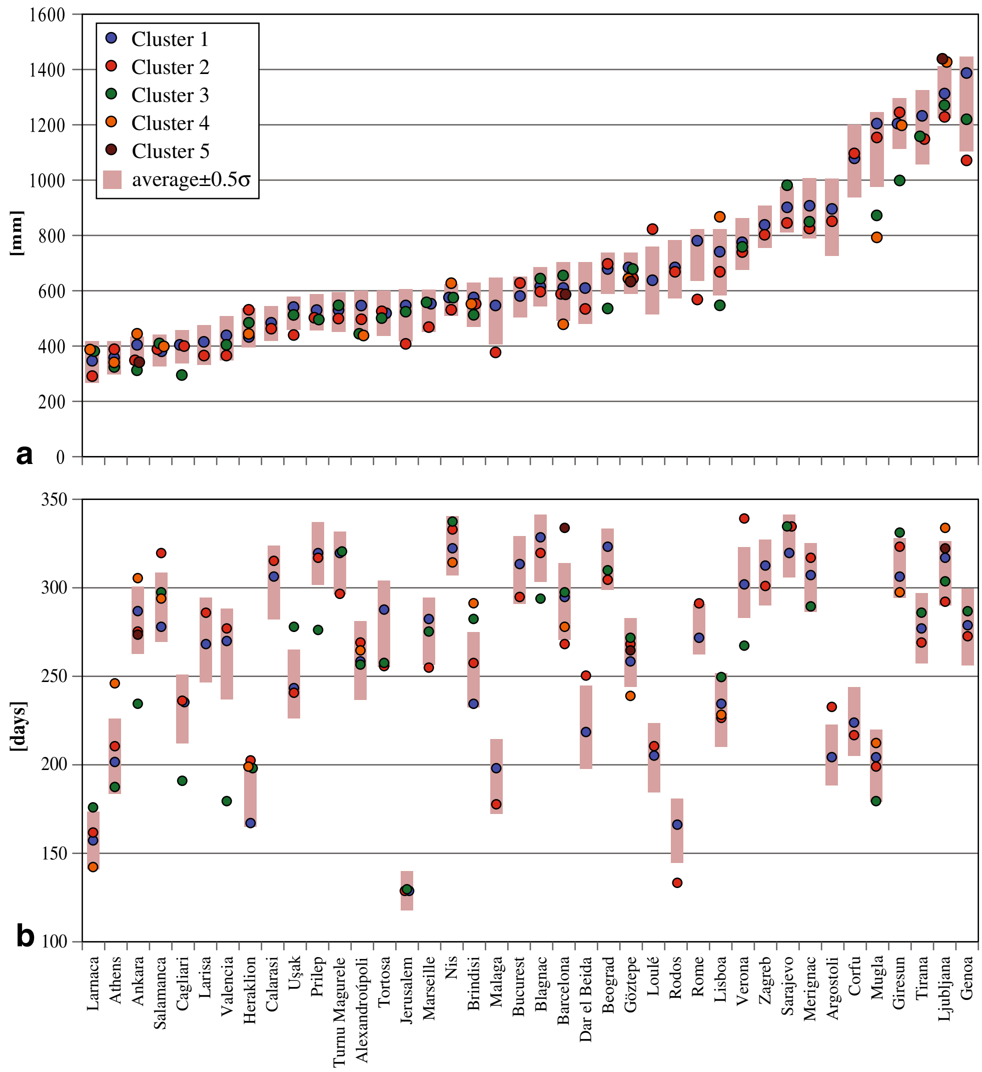

Fig. 8 The averages TOTAL± 0.5σ for the entire dataset and the clusters' average TOTAL (**a**) and the average RSL ±0.5σ for the entire dataset and the clusters' average RSL (**b**). Stations are sorted according to the average TOTAL in an ascending order

grouped in Jerusalem with only 87% (47 out of 54 years). Therefore, it seems that the limitation of having at least two clusters and not more than 5 is justified.

Stations with a high clustered percentage of years grouped into only two or three clusters indicate a relatively predictable accumulation courses. Hence, it is possible to assume that, for example, the uncertainty related with the intraseasonal distribution in Bucurest (96.0% of the years clustered in only two clusters) is lower than in Ljubljana (87.5% in five clusters). Another way of expressing the uncertainty related to accumulation rates can be derived from the percentage of years grouped into cluster 1 (highest probability). The highest percentage of years grouped in cluster 1 was found in Bucurest (89.3%); hence, only 10.7% of the years have a different distribution, while in Brindisi, only 33.3% of the years were clustered in cluster 1, meaning that two thirds of the years had different distributions.

The clustered annual distributions in all stations are presented in Fig. 6. The shadow represents a year period from the SAD in each station (Table 1) until SAD+364. Generally, it can be observed that the AP courses have three different shapes:

1. Accumulation rates having straight oblique lines are found mainly in stations from the northern part of the study area, e.g., Zagreb, Sarajevo, Merignac, Beograd, and Ljubljana. This shape is formed when precipitation occurs throughout the entire year, and therefore, the accumulation rate is constant all year long.
2. An “S” shape accumulation rate is found mainly in the southern Mediterranean basin, e.g., Jerusalem, Larnaca, Heraklion, Rodos, and Muğla. This shape is formed when, at the beginning of the year, the accumulation rate is very slow; at some point, it increases during the mid year, and at the end of the year, it slows down again. Most of these stations demonstrate two distinguishable seasons: a dry period during the summer months, mainly between May and September, and a concentrated rainy period during the rest of the year.

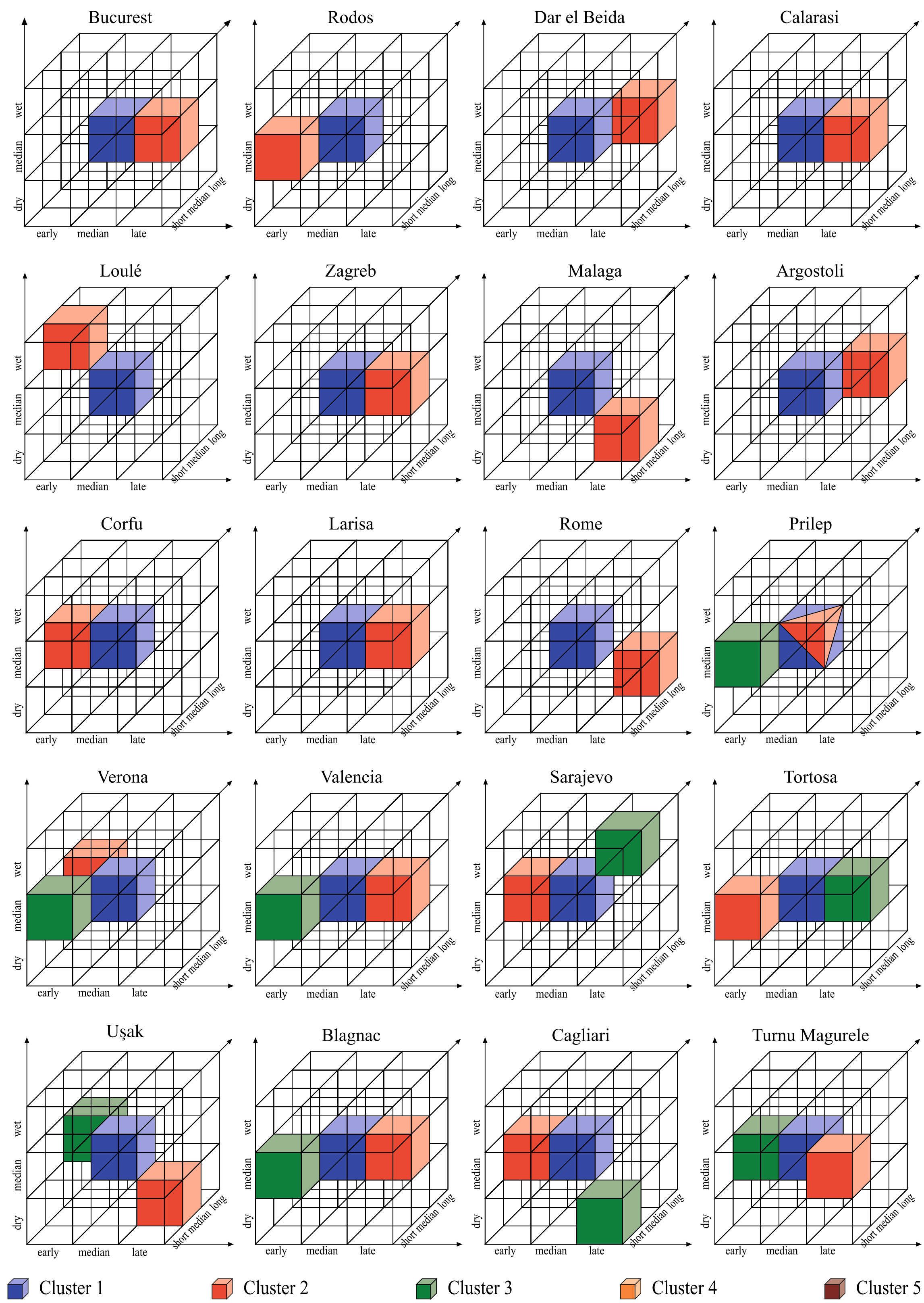

Fig. 9 A schematic illustration for the definitions of the clusters according to the average TOTAL, RSL, and DAP in all stations. Stations are sorted according to the number of classifications and Table 2

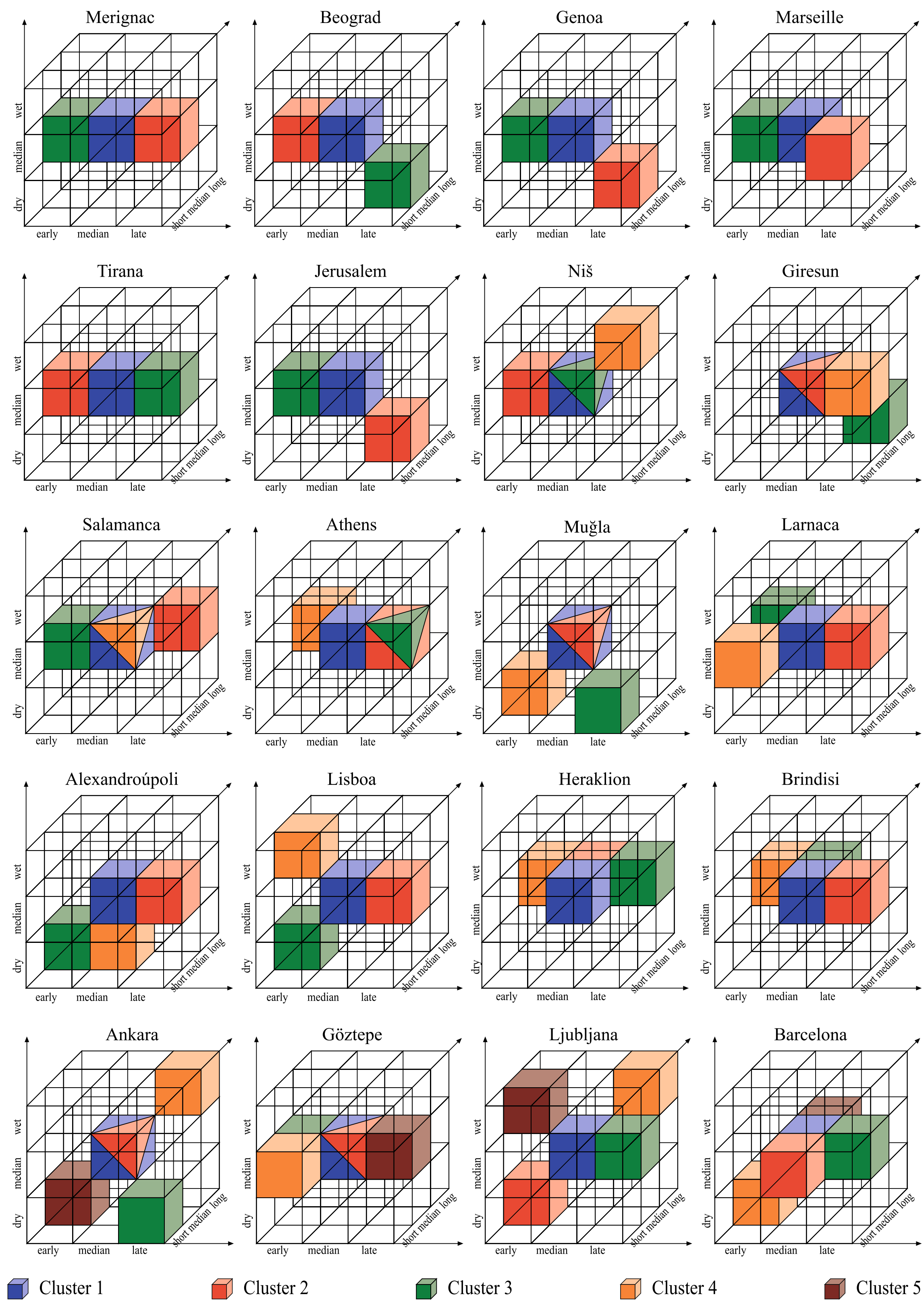

Fig. 9 (continued)

Table 3 Characteristics of all clusters in all stations

Cluster 1										
TOTAL	Dry			Median			Wet			
RSL \ DAP	Early	Median	Late	Early	Median	Late	Early	Median	Late	
Short										
Median					41					
Long										
Cluster 2										
TOTAL	Dry			Median			Wet			
RSL \ DAP	Early	Median	Late	Early	Median	Late	Early	Median	Late	
Short				2	2	2				
Median	1		4	7	5	11	1			
Long			1	1	1	3				
Cluster 3										
TOTAL	Dry			Median			Wet			
RSL \ MSD	Early	Median	Late	Early	Median	Late	Early	Median	Late	
Short			3	4						
Median	2		2	7	1	5			1	
Long			1	2	1	1				
Cluster 4										
TOTAL	Dry			Median			Wet			
RSL \ DAP	Early	Median	Late	Early	Median	Late	Early	Median	Late	
Short				2						
Median	2	1			1	1	2		1	
Long				3				1	1	
Cluster 5										
TOTAL	Dry			Median			Wet			
RSL \ DAP	Early	Median	Late	Early	Median	Late	Early	Median	Late	
Short	1									
Median	1					1	1			
Long					1					
Clusters summary										
TOTAL	Dry			Median			Wet			Sum
RSL \ DAP	Early	Median	Late	Early	Median	Late	Early	Median	Late	
Short	1		3	8	2	2				16
Median	6	1	6	14	48	18	4		2	99
Long			2	6	3	4		1	1	17
Sum	7	1	11	28	53	24	4	1	3	132
	19			105			8			
	40				54				38	

3. Crossed courses (occur in clusters 3, 4, or 5), which indicate inconsistent intraseasonal rates and timing. For example, a season starts slow and late, and at its end, it accumulates faster and ends early, or vice versa (e.g., Niš, Alexandroú-poli, Salamanca, Lisboa, Brindisi, and Barcelona).

Figure 7 presents the median periods and the $DAP_{(25\%, 50\%, 75\%)}$ of every cluster. The objective of this illustration is to point out the differences in the AP timing. However, since, for example, $DAP_{(25\%)}$ accumulates as early as May in some stations while in others this occurs around November, the comparison between the stations can only be within each group of stations having the same SAD, as presented in Reiser and Kutiel (2008). The stations' order is therefore based on their SAD and according to the >number of clusters in an ascending order. All DAPs of all clusters 1 were within the median period, and so these clusters were classified as median in all stations. The largest departures were identified in three stations located in the eastern coast of the Iberian Peninsula. Departures of clusters 2 and 3 in Barcelona (81 days) and Tortosa (66 days), respectively, from the median period of $DAP_{(25\%)}$ (Fig. 5a) and in cluster 3, in Valencia, a 105-day departure from the median period of $DAP_{(75\%)}$ (Fig. 5c).

Figure 8 illustrates the average TOTAL$\pm 0.5\sigma$ range for the entire dataset and the averages RSL$\pm 0.5\sigma$ in all stations. In addition, the average TOTAL and average RSL of every cluster are presented. The stations are sorted according to the TOTAL in an ascending order, where Larnaca is the driest and Genoa the wettest. One noticeable fact is that there is no correlation between the TOTAL and the RSL. Hence, a station can be dry, and its rainy season will not necessarily be short and vice versa. Furthermore, it can be noticed in Fig. 8a that the departures from the average TOTAL are mostly drier than the average range.

The combined definitions of the three parameters (DAPs, TOTAL, and RSL) in all stations are presented in Fig. 9. Stations are sorted according to Table 2 and then according to the number of classifications, since in eight stations, there were two clusters characterized with the same classification in an ascending order. In these cases, the differences between the two types of seasons could not be described accurately by using only the parameters, which describe the timing, quantity, and length. The differences that caused the separation of the relevant years into two clusters in the clustering process were evident in the rate of the annual course as demonstrated in Fig. 6. One example for such a case can be seen in Ankara, where clusters 1 and 2 were both defined as median for all three parameters. The clustered annual courses as presented in Fig. 6 indicate that cluster 1 has a more moderate and constant accumulation, while the course in cluster 2 accumulates in a slower rate at the beginning of the season (up to 25%) and in a faster rate at the second half of the season. In general, most of these cases are crossed courses.

According to Fig. 9, *all* clusters 1 (blue) in all stations are identical and were classified as Median for all three parameters, DAPs, TOTAL, and RSL. Other similarities in the classification types that were found:

1. In Bucurest, Zagreb, Calarasi, and Larisa, clusters 2 are median for the TOTAL and for the RSL and late for the DAPs.
2. In Dar el Beida and Argostoli, clusters 2 are median for the TOTAL, long for the RSL, and median again for the DAPs.
3. In Turnu Magurele and Marseille, three identical clusters were found. Clusters 2 were median for the TOTAL, short, and late; clusters 3 were median in both TOTAL and RSL and Early DAPs.
4. In Merignac and Tirana, three identical clusters were found; yet, their probabilities were different, as can be seen in Table 2. Whereas in Merignac, the probability of a median TOTAL, median RSL, and early DAPs distribution was only 8.5% (Cluster 3), it stands on 22.0% in Tirana (cluster 2). On the other hand, the probabilities for a late distribution (TOTAL and RSL are both median) was almost similar, with 13.6% in Merignac (Cluster 2) and 14.6% in Tirana (Cluster 3).
5. In Beograd, Genoa, and Jerusalem, three identical clusters were found, though the probabilities were different as can be seen in Table 2. In Beograd, the probability for a median TOTAL, median RSL, and early DAPs classification was 22.9% (Cluster 2), while only 10.0% for a dry and median RSL and late DAPs (cluster 3). In Genoa and Jerusalem, on the other hand, the probability for median TOTAL, median RSL, and early DAPs (cluster 3) was only 10.6% and 7.4%, respectively, whereas for a dry and median RSL and late DAPs (cluster 2), the probability was 29.8% and 33.3%, respectively (Table 2).
6. Three identical clusters were found in Valencia, Tortosa, and Blagnac. The probability for a median cluster (cluster 1) was the highest in Valencia (among these three stations), with 80.9%, and the smallest was in Blagnac (74.6%). In Valencia and Blagnac, the probability for a median TOTAL, median RSL, and late DAPs type was similar (10.3% and 10.2%, respectively); yet,

Table 4 Characteristics of all non-clustered years in all stations

<table>
<tr><th colspan="11">Non clustered years</th></tr>
<tr><td>TOTAL</td><td colspan="3">Dry</td><td colspan="3">Median</td><td colspan="3">Wet</td><td rowspan="2">Sum</td></tr>
<tr><td>RSL DAP</td><td>Early</td><td>Median</td><td>Late</td><td>Early</td><td>Median</td><td>Late</td><td>Early</td><td>Median</td><td>Late</td></tr>
<tr><td>Short</td><td>12</td><td></td><td>22</td><td>4</td><td></td><td>14</td><td>6</td><td></td><td>7</td><td>65</td></tr>
<tr><td>Median</td><td>10</td><td>1</td><td>10</td><td>9</td><td></td><td>7</td><td></td><td></td><td>5</td><td>42</td></tr>
<tr><td>Long</td><td>3</td><td></td><td>12</td><td>10</td><td></td><td>11</td><td>2</td><td></td><td>4</td><td>42</td></tr>
<tr><td rowspan="3">Sum</td><td>25</td><td>1</td><td>44</td><td>23</td><td></td><td>32</td><td>8</td><td></td><td>16</td><td rowspan="3">149</td></tr>
<tr><td colspan="3">70</td><td colspan="3">55</td><td colspan="3">24</td></tr>
<tr><td>56</td><td colspan="3"></td><td>1</td><td colspan="3"></td><td>92</td></tr>
</table>

in Tortosa, it was only 6.7%. Finally, the probability for a median TOTAL, short RSL, and early DAPs type was the smallest in Valencia (7.4%) and the highest in Tortosa (13.3%).

Table 3 summarizes all classifications in all stations. As aforementioned, all clusters 1 presented a median character of the three parameters. Clusters 2 had more variability in three classifications with 13 features. Most clusters (34 out of 41) had a median TOTAL character, 16 of them were late and ten were early.

Out of a total of 132 clusters found in all 41 stations, 99 were median in their RSL, 16 were short, and 17 had a long classification. There were 19 dry courses and only eight wet, while most of the clusters had a Median rainfall quantities (105). The intraseasonal distribution, represented by three DAPs, had a relatively even distribution in the various types. Forty clusters classified as early, 38 as late, and 54 had a course median.

There are 149 years (in all stations together) out of a total of 2,303 analyzed years (6.5%), which were not clustered. The non-clustered years from all stations were classified as well, using the same method, and the results are presented in Table 4. Most of these years were late (92 out of 149), almost half of the cases had a dry character (70 years) and 65 were short. These extreme years will be further studied.

There were four types of courses, which did not appear in any of the clusters or non-clustered years (marked with a gray shadow in Table 4). All four types are median in their timing and:

1. Wet and short
2. Wet and median (in length)

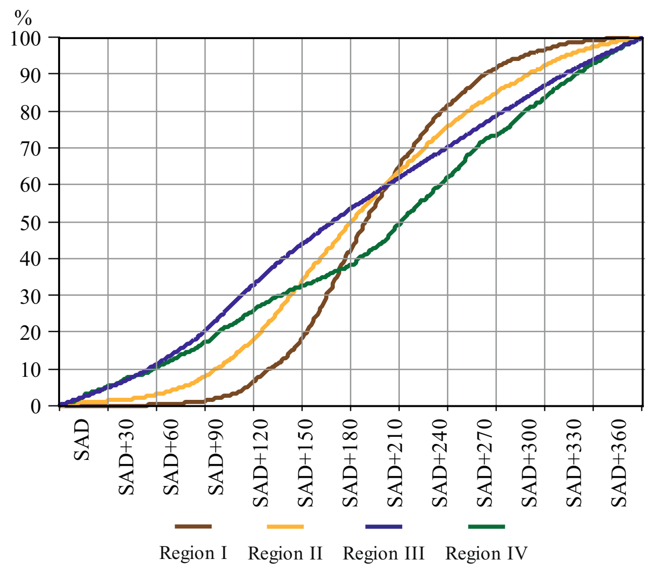

Fig. 11 The averaged AP annual courses in the four regions presented in Fig. 10

3. Dry and short
4. Dry and long.

Only two cases were classified as dry with median RSL and DAPs: cluster 4 in Alexandroúpoli and one non-clustered year, 1972, in Beograd.

The spatial distribution of the stations according to their annual course of cluster 1 is presented in Fig. 10. Figure 11 presents the average annual courses of cluster 1 in the four regions. Each line represents the average course of all stations in the relevant region. The horizontal axis relates to a year period, from the SAD until SAD+ 364.

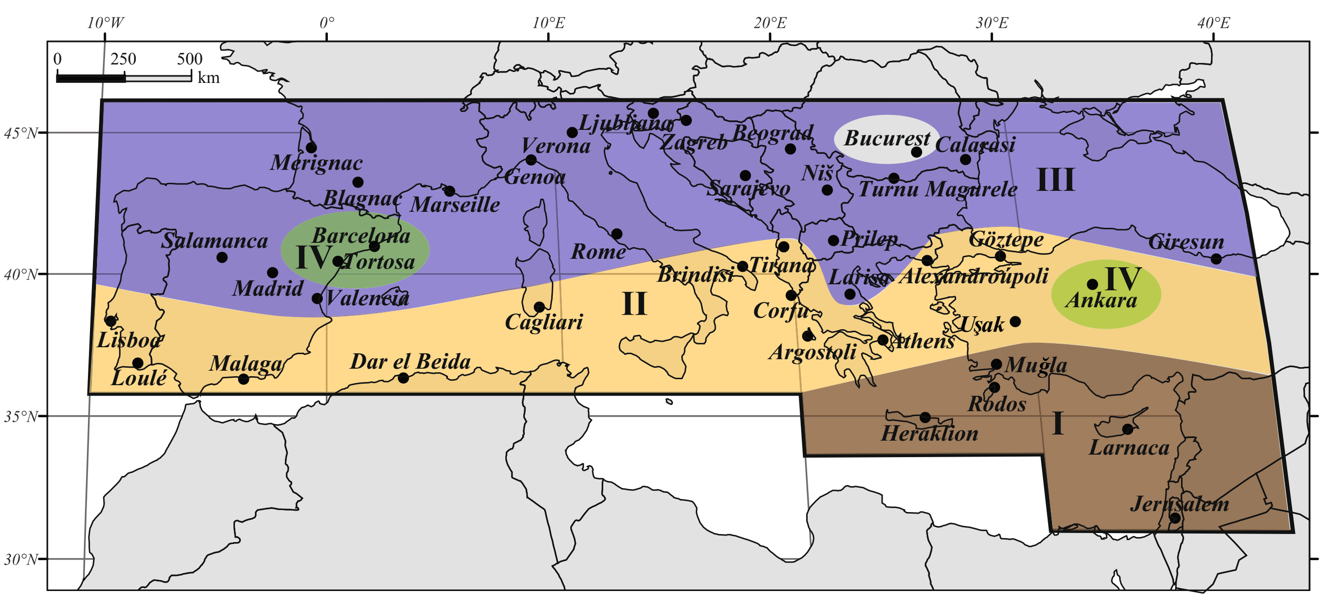

Fig. 10 Spatial distribution of the annual courses of custer 1

The study area was divided into four regions:

- Region I in the Southeastern Mediterranean, comprising five stations. It presents an “S” shape distribution with a long dry period of approximately 3 months in the beginning of the year and about two additional dry months at its end, leaving a very intense rainy season, comprising of about 6–7 months.
- Region II in the central basin is comprised of 13 stations. These stations have also an “S” shape distribution, though less severe than in region I. This means that the rainy season is longer, and the dry period ends earlier and starts later at the end of the rainy season.
- Region III in the Northern part of the basin is comprised of 19 stations. This region presents an oblique line, which means that precipitation occurs throughout the entire year with no real definable dry or wet periods.
- Region IV is comprised of three stations: Barcelona, Tortosa, and Ankara. In this region, like in region III, rainfall occurs throughout the year but at a slower rate for the first half of the rainy season and a faster one in the second half.

Bucurest is the only station that was not clustered in any region.

5 Conclusions

The main conclusions of the present study can be summarized as follows.

The use of CA, as a tool for identifying the temporal (rather than spatial) behavior of the rainfall along the year, was proven to be satisfactory. It enables to detect and group various years into clusters with similar accumulation courses.

When applied to the Mediterranean basin, stations were characterized by two to five different types of seasons with various probabilities. Furthermore, each cluster in each station was characterized also by its TOTAL and its RSL. The main findings are:

1. The most frequent type (cluster 1) was median in all three parameters, in all stations with probabilities ranging from 33.3% to 89.3%.
2. The probability for a year to be dry is higher than wet. Likewise, the probability of a year to be long is higher than to be short.
3. The highest probability for *non-clustered years* is to be characterized as dry, short, and late (14.8%, 22 years out of 149). These years were excluded from the representations and are considered as extreme events. A further analysis of these years is necessary in order to provide more information about extreme rainfall variability.
4. A spatial distribution of the most frequent AP courses reveals that, in the northern part of the Mediterranean basin, there is a constant accumulation rate throughout the entire year, whereas the southern part is characterized by an intense accumulation rate and a long dry period.

Acknowledgments We wish to thank Mrs. Noga Yoselevich from the department of Geography and Environmental Studies, University of Haifa for her help with producing the figures presented in this paper and to Dr. Yakov Aviad for helping us to create the RUEM and solving every technical problems in a most elegant way.

References

García-Oliva F, Ezcurra E, Galicia L (1991) Pattern of rainfall distribution in the central Pacific coast of Mexico. Geogr Ann 73:179–186

Lázaro R, Rodrigo FS, Gutirrez L, Domingo F, Puigdefáfragas J (2001) Analysis of a 30 year rainfall record in semi-arid SE Spain for implications on vegetation. J Arid Environ 48:373–395

Logue JJ (1984) Regional variations in the annual cycle of rainfall in Ireland as revealed by principal component analysis. J Climatol 4:597–607

Maheras P (1988) Changes in precipitation conditions in the western Mediterranean over the last century. J Climatol 8:179–189

Norrant C, Douguédroit A (2005) Monthly and daily precipitation trends in the Mediterranean (1950–2000). Theor Appl Climatol 83:89–106

Paz S, Kutiel H (2003) Rainfall regime uncertainty (RRU) in an eastern Mediterranean region—a methodological approach. Isr J Earth-Sci 52:47–63

Ramos MC (2001) Rainfall distribution patterns and their change over time in a Mediterranean area. Theor Appl Climatol 69:163–170

Reiser H, Kutiel H (2008) Rainfall uncertainty in the Mediterranean: Definition of the rainy season—a methodological approach. Theor Appl Climatol 94:35–49

Reiser H, Kutiel H (2009a) Rainfall uncertainty in the Mediterranean: definitions of the daily rainfall threshold (DRT) and the rainy season length (RSL). Theor Appl Climatol 97:151–162

Reiser H, Kutiel H (2009b) Rainfall uncertainty in the Mediterranean: Dryness Distribution. Theor Appl Climatol. doi:10.1007/s00704-009-0163-4

Rodrigo FS (2002) Changes in climate variability and seasonal rainfall extremes: a case study from San Fernando (Spain), 1821–2000. Theor Appl Climatol 72:193–207

Rodriguez-Puebla C, Encinas AH, Nieto S, Garmendia J (1998) Spatial and temporal patterns of annual precipitation variability over the Iberian Peninsula. Int J Climatol 18:299–316

Romero R, Guijarro AJ, Ramis C, Alonso S (1998) A 30 years (1964–1993) daily rainfall data base for the Spanish Mediterranean regions: first exploratory study. Int J Climatol 18:541–560

Tennant WJ, Hewitson BC (2002) Intra-seasonal rainfall characteristics and their importance to the seasonal prediction problem. Int J Climatol 22:1033–1048

Tomozeiu R, Lazzeri M, Cacciamani C (2002) Precipitation fluctuations during the winter season from 1960 to 1995 over Emilia–Romagana, Italy. Theor Appl Climatol 72:221–229

Türkeş M (1996) Spatial and temporal analysis of annual rainfall variations in Turkey. Int J Climatol 16:1057–1076