INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **19**: 557–578 (1999)

DAILY RAINFALL AFFINITY AREAS IN MEDITERRANEAN SPAIN

R. ROMERO[a], C. RAMIS[a,*], J.A. GUIJARRO[b] and G. SUMNER[c]
[a] *Meteorology Group, Departament de Física, Universitat de les Ilees Balears, Palma de Mallorca, Spain*
[b] *Instituto Nacional de Meteorología, Centro Meteorológico de Baleares, Palma de Mallorca, Spain*
[c] *Department of Geography, University of Wales, Lampeter, UK*

Received 16 February 1998
Revised 12 October 1998
Accepted 12 October 1998

ABSTRACT

A subdivision of the Spanish Mediterranean region into daily rainfall affinity areas was conducted using a pluviometric data base of 410 gauges for the period 1964–1993. Areas were derived using principal components analysis on the between-site correlation matrix. Two regionalizations defining 12 areas are offered: the first, non-overlapping, solution is derived by cluster analysis on the most important unrotated PCs extracted, while the second, with overlapping pluviometric regions, is obtained by mapping the obliquely rotated principal components. Both solutions are similar, and emphasize the dominant role exerted by the complex topography of the region on the spatial organization of rainfall through its interaction with the main rain-bearing airflows and systems. A higher-resolution regionalization comprising 20 non-overlapping areas is also included. South and southeast domains, where topography is very prominent and complex, as well as the Balearic Islands, contain most of the additional detail offered by this solution. Copyright © 1999 Royal Meteorological Society.

KEY WORDS: western Mediterranean region; rainfall organization; climate regionalization; principal components analysis; cluster analysis; Spain

1. INTRODUCTION

The western Mediterranean region possesses a quite complex topographic configuration. Important mountain ranges flank the western Mediterranean Sea (notably the Atlas Mountains, the Appennines, the Mediterranean coastal ranges of Spain, the Pyrenees, the Central Massif of France and the Alps), which almost encircle the western Mediterranean basin and largely isolate it from other regions. Topographic variability is especially notable in Mediterranean Spain, here defined as the administrative regions of Catalonia, Valencia, Murcia, Andalucía and the Balearic Islands, measuring approximately 1000 km in both north–south and east–west directions. In some areas, the terrain height exceeds 3000 m comparatively close to the sea (Figure 1).

As has been highlighted in two previous papers (Romero *et al*., 1998, 1999), Mediterranean Spain provides an interesting area for studies of the spatial and temporal variability of rainfall. There are three main reasons for this. First, its high relief produces accentuated rainfall contrasts between adjacent uplands and lowlands, and between areas with differences of exposure to humid, maritime winds. Mid-latitude Atlantic disturbances can be effective in the Pyrenees and in western and interior Andalucía, but are normally significantly altered and weakened when they reach the sheltered areas along the Mediterranean coast. These areas rely principally on humid Mediterranean flows, and on locally generated Mediterranean cyclones, such as those which may develop to the lee of the Atlas Mountains, the Pyrenees and the Alps (Reiter, 1975). Second, rainfalls are very often torrential in character, especially during the autumn. The closed topography around the western Mediterranean Sea and the high insolation received during the summer lead to high sea surface temperatures, thereby

* Correspondence to: Meteorology Group, Departament de Física, Universitat de les Iles Balears, Palma de Mallorca, Spain; e-mail: dfscrn0@eps.uib.es

CCC 0899–8418/99/050557–22$17.50

ensuring a considerable moisture content and low static stability in the overlying air. As a consequence, the Mediterranean air mass very often possesses convective instability during the autumn (Meteorological Office, 1962; Ramis, 1995). This convective instability is released as organized and effective convection over the coastal areas (Font, 1983; Riosalido, 1990) when synoptic scale forcing is favourable. Finally, the region's latitude imposes strong seasonal contrasts which are most pronounced in the southernmost areas. While western Andalucía is one of the wettest zones during winter, its summers are extremely dry.

Romero *et al.* (1998) produced a homogenized daily rainfall data base for the period 1964–1993 for the 410 stations shown in Figure 1 to illustrate average yearly and seasonal variation in rainfall amount, frequency and intensity, and the persistence or absence of the rainfall in Mediterranean Spain. Romero *et al.* (1999) used the same data base to derive, using principal components analysis (PCA) and cluster analysis (CA), the main spatial patterns controlling significant and torrential daily rainfalls in the area. The results indicated that both the extent of the region and the degree of exposure or shelter within it induced by the complex topography produced marked spatial organization in daily rainfalls.

Objective regionalizations of different climatic variables (typically rainfall and temperature), have been accomplished for many regions of the world. These regionalizations can be obtained using different methods, ranging from elementary linkage analysis (e.g. Sumner and Bonell (1990), for daily rainfall in Wales), to more complex multivariate statistical analysis techniques. For example, Bärring (1988) presents a regionalization of daily rainfall in Kenya by means of common factor analysis; White *et al.* (1991) apply PCA and different rotation algorithms to monthly Pennsylvanian precipitation data; Bonell and Sumner

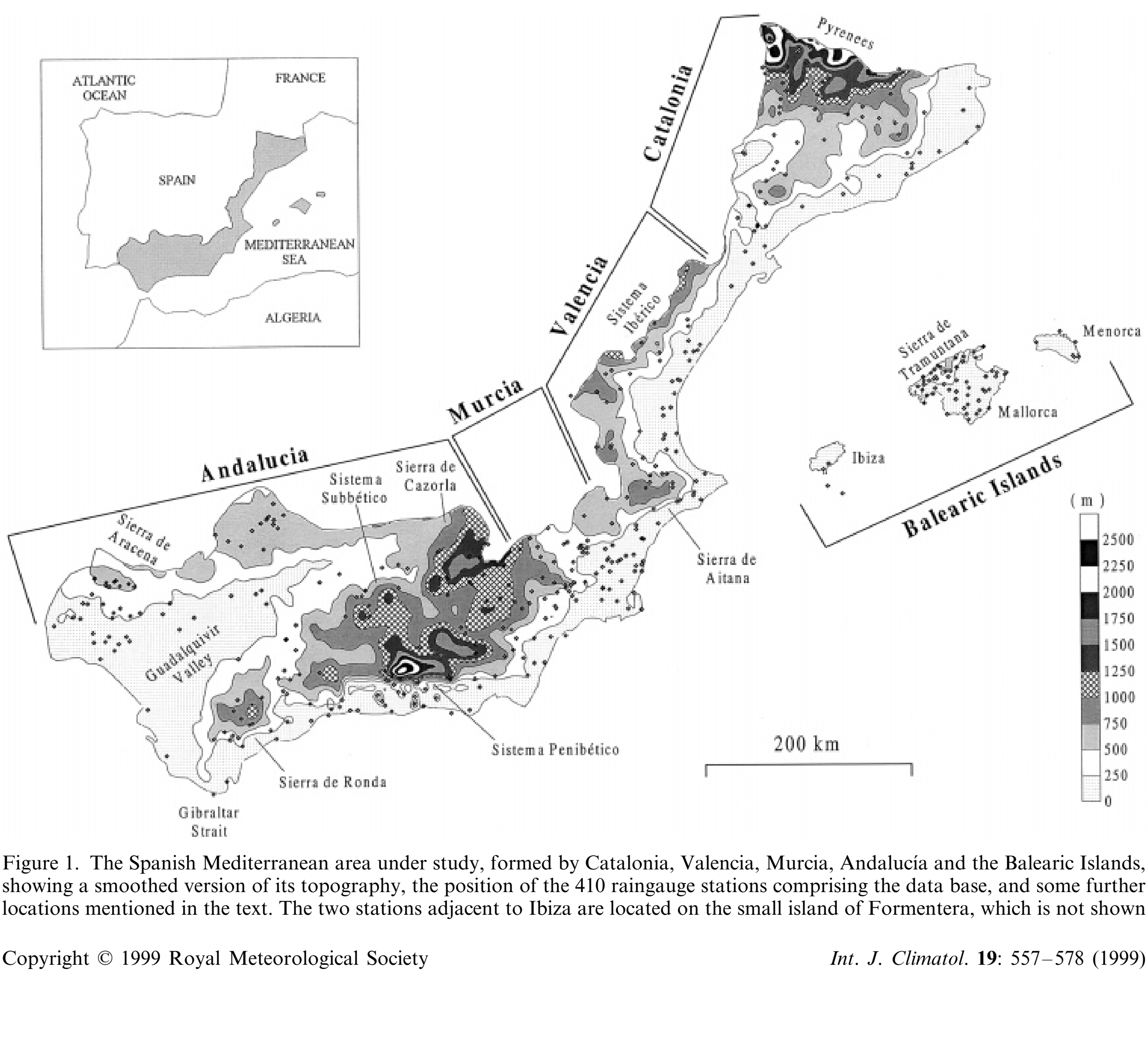

Figure 1. The Spanish Mediterranean area under study, formed by Catalonia, Valencia, Murcia, Andalucía and the Balearic Islands, showing a smoothed version of its topography, the position of the 410 raingauge stations comprising the data base, and some further locations mentioned in the text. The two stations adjacent to Ibiza are located on the small island of Formentera, which is not shown

(1992) use obliquely rotated principal components and three CA methods for daily precipitation in Wales; Gong and Richman (1995) test a large number of hierarchical and non-hierarchical CA methods for growing season precipitation data in North America east of the Rockies; and Fovell and Fovell (1993), determine climatic zones and subzones of the conterminous United States from monthly temperature and precipitation, after clustering the orthogonally rotated principal components of their dataset. Similar objective regionalizations do not exist for Mediterranean Spain as a whole, although studies for particular zones have been developed (e.g. Sumner *et al.* (1993) for Mallorca, and Periago *et al.* (1991) for Catalonia).

In this work, the 30-year data base of Romero *et al.* (1998) is made use of once again to analyze the spatial organization of daily rainfall over Mediterranean Spain. The analysis was carried out partly to improve understanding of the relevant physical processes, but more importantly, to provide an objective pluviometric regionalization of the area. Such a regionalization is potentially useful in providing a rational, climatological subdivision of Mediterranean Spain, and could be relevant for regional forecasting, which is more commonly carried out using administrative, rather than physically logical climatic or topographic, units.

2. DATA BASE AND METHODOLOGY

The data base comprises a complete, homogeneized daily rainfall series for 410 stations in Mediterranean Spain (Figure 1) for the period 1964–1993. The spatial coverage is generally satisfactory (and exceptionally good in Mallorca), although some gaps exist around the Guadalquivir valley, in mountainous areas of Andalucía, north of Valencia, and in some areas of Catalonia.

Initially, the available data consisted of 3366 records during a longer period, 1951–1995, provided by the Spanish Instituto Nacional de Meteorología (INM). Most of these stations possessed a poor temporal coverage. The final data base resulted from an initial selection of sites with a minimum of 1000 days (almost 3 years) availability (yielding 2842 stations), and a second selection searching for the longest sub-period with the highest number of stations with tolerable completion. This two-tier selection yielded 410 sites for the period 1964–1993, and involved the retention of all sites with a better than 90% daily availability. An iterative method using all the available information of the 2842 stations was then applied to fill any gaps in the 410-gauge series and for quality control. The details of the algorithm can be found in Romero *et al.* (1998).

In a semi-arid area such as Mediterranean Spain, there is a very high proportion of totally dry days at all sites. Since a regionalization of rainfall is being sought rather than dryness, the application of some method of screening the data seems appropriate. The present strategy has been to enter in the analysis only those days in which at least 5% of the stations registered more than 5 mm. This criterion yields a total of 3941 days out of the 10958 days contained in the 30 years of data. As discussed in Romero *et al.* (1999), the bulk of extreme events, generally convective, are effectively captured by the screening method. Obviously, very local important rainfalls produced by isolated convective cells are excluded, but these types of rainfalls are not what the study—which seeks by definition a generalization of affinity areas—is examining. Of the 3941 filtered days, 30.0% occurred in winter (December, January, February), 29.6% in spring (March, April, May), 26.8% in autumn (September, October, November) and only 13.6% in summer (June, July, August). In addition, to reduce positive skewness of the data $\log_{10}$-transformed rainfall (substituting values of 0.01 for zero entries) were used. With this transformation, histograms more closely approximated the normal distribution. However, it should be remembered that neither PCA nor CA explicitly demand that variables be normally distributed to operate correctly (Fovell and Fovell, 1993).

As in previous works (White *et al.*, 1991; Bonell and Sumner, 1992; Sumner *et al.*, 1993; Gong and Richman, 1995), the approach followed to accomplish the regionalization is first to subject the S-mode (site-by-site) correlation matrix to PCA. The PCA is designed not merely as a data reduction technique, but as a method which ensures that only the fundamental modes of variation of the data are considered.

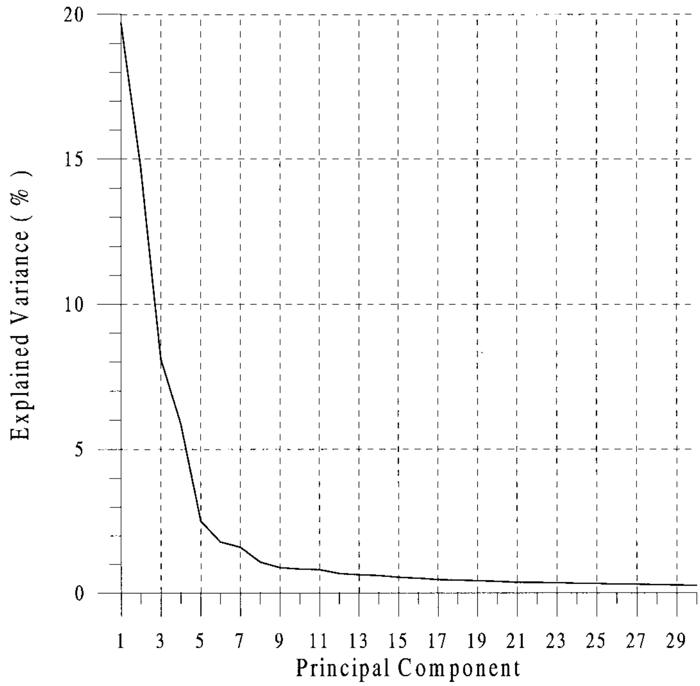

Figure 2. Principal components' scree plot. Only the first 30 PCs of the resulting 410 PCs are shown

Several criteria have been suggested for deciding how many PCs to retain in order to separate 'signal' from 'noise' (Jolliffe, 1986; Bärring, 1987; Preisendorfer, 1988), but a clear-cut number of PCs is certainly rare. In this study, the simple scree test of Cattell (1966a) was adopted. The scree plot of the present analysis (Figure 2) contains several breaks of the slope (at the fifth, eighth and less clearly at the twelfth PC). As shown in next section, maps of loadings for PCs 6–12 comprise subtle variations of pattern that can be associated with known modes of rainfall development, and the rotation algorithm applied on 12 PCs produced very simple structures. For these reasons, it was decided to retain 12 PCs, which account for 58.52% of the total variance (Table I). The level of explained variance seems to be acceptable, given the considerable number of rain gauges and the complex nature of the rainfall variable at the daily time scale.

Table I. Eigenvalues and percentages of explained and accumulated variance for the 12 extracted PCs

PC	Eigenvalue	Explained (%)	Accumulated (%)
1	80.79	19.71	19.71
2	60.23	14.69	34.40
3	33.18	8.09	42.49
4	24.14	5.89	48.38
5	10.27	2.51	50.89
6	7.30	1.78	52.67
7	6.53	1.59	54.26
8	4.40	1.07	55.33
9	3.60	0.88	56.21
10	3.40	0.83	57.04
11	3.32	0.81	57.85
12	2.76	0.67	58.52

The traditional hard regionalization, when regions do not overlap and must be irregularly shaped, is obtained by carrying out CA of the 410 sites based on the loadings attained by the 12 unrotated PCs (Bonell and Sumner, 1992; Sumner *et al*., 1993). That is, sites with similar loadings on the extracted PCs are clustered together. This approach aims to incorporate sites into one area with similar time distributions of rainfall, irrespective of overall amount, which depend critically on altitude. The CA technique used is the non-hierarchical k-means method (Anderberg, 1973), as implemented in the STATISTICA (1994) utility. The Euclidean distance is taken as the similarity index. Gong and Richman (1995) showed that non-hierarchical methods such as this out-performed hierarchical techniques.

A hard regionalization, although simple in concept, does not seem to be the choice that best preserves the underlying physics of precipitation generation and suppression. Causal mechanisms of precipitation are fuzzy and overlapping, and it was decided that the present regionalization should reflect this fact. For this reason, a solution consisting of overlapping regions was also considered. As in Gong and Richman (1995), this is accomplished by mapping the loadings of the rotated PCs, assuming that the rotation method can produce simple structures (Richman, 1986). In fact, some authors consider the application of rotation methods as being essentially an alternative to cluster analysis (Jolliffe, 1987). In general, oblique rotation is preferred over orthogonal rotation, since superior regional definition is obtained (White *et al*., 1991; Yarnal, 1993). VARIMAX (orthogonal) and OBLIMIN (oblique) (see Richman (1986) for a description) methods were used in this analysis. The latter produced strong simple structures (see section on Oblimin-rotated PCs), and was therefore selected to represent the overlapping regions.

3. PRINCIPAL COMPONENTS

3.1. Unrotated PCs

Figure 3 shows the plots of the 12 unrotated PCs, UNR-*i*, where *i* indicates the PC number. One of the objectives of performing PCA is to try to identify the main physical mechanisms of rainfall enhancement and suppression that explain the most important rainfall contrasts illustrated by the PCs. As commonly occurs, practically all loadings of the first PC (UNR-1), which accounts for 19.71% of the total variance (Table I), possess the same sign, with the highest values in the central part of the domain. This PC indicates that the first source of variance in the data is associated with the overall magnitude of rainfall events in the region as a whole. It is interesting to note, however, that as a consequence of the region's large areal extent, daily rainfalls affecting the entire the territory are rarely observed (see Figure 12, from Romero *et al*., 1999). This type of lack of correspondence between unrotated PCs and 'reality' is one of the disadvantages which led Richman (1986) to recommend the use of rotated PCs when physical interpretation is intended.

The other PCs exhibit an alternation of sign between the different zones, with, in general, increasingly complex structures as the PC number grows (Figure 3). These structures suggest the sequence of patterns demonstrated by Buell (1975), which are attributed by some authors (e.g. Buell, 1979; Richman, 1986) to a domain shape dependence of the unrotated PCs. Such a sequence is complicated in the present case, probably owing to the irregular shape of the domain. Richman (1986) cautions that as a consequence of domain shape dependence, the unrotated PC approach produces results that may not yield realistic modes of variation and should be looked upon with suspicion. From the present authors' point of view, each of the PCs shown in Figure 3 seems to comprise a combination of different, but known, modes of rainfall development. As a result, some of the individual PCs (UNR-5, UNR-6, UNR-8, UNR-11 and UNR-12) are hard to interpret. For the general aspects of the remainder PCs, however, plausible physical interpretations based on regional knowledge and experience, and some previous results can be found (Romero *et al*., 1998, 1999).

UNR-2, which accounts for 14.69% of the total variance, shows a marked gradation in a west–east direction. This is probably linked to one of the most fundamental climatological influences on rainfall in the region (Romero *et al*., 1999; Figure 12). Cold season low pressure systems located over the Atlantic

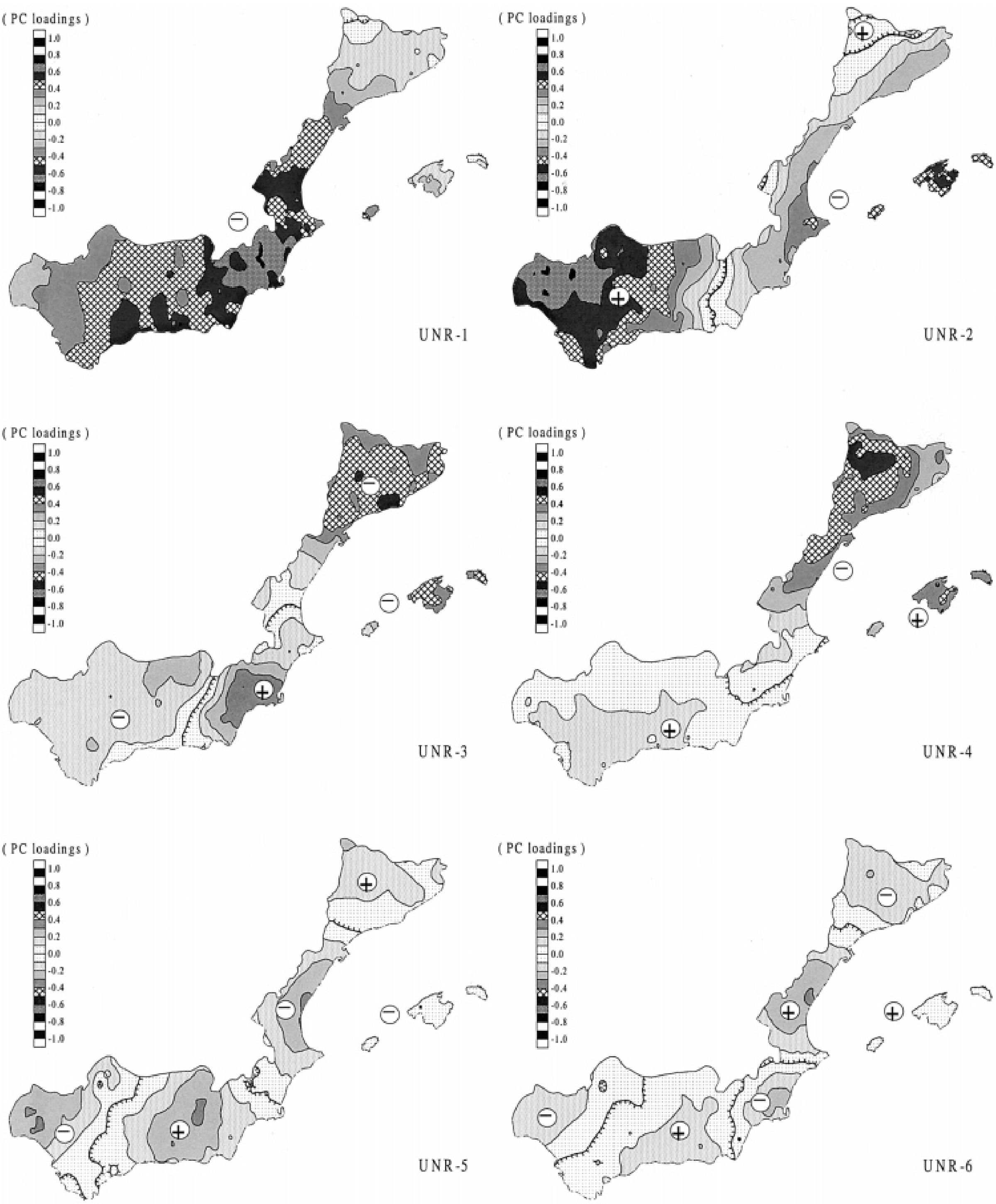

Figure 3. Spatial distributions of loadings for the first 12 unrotated PCs. Hachures along the zero isopleth point towards positive loadings; positive and negative zones are indicated

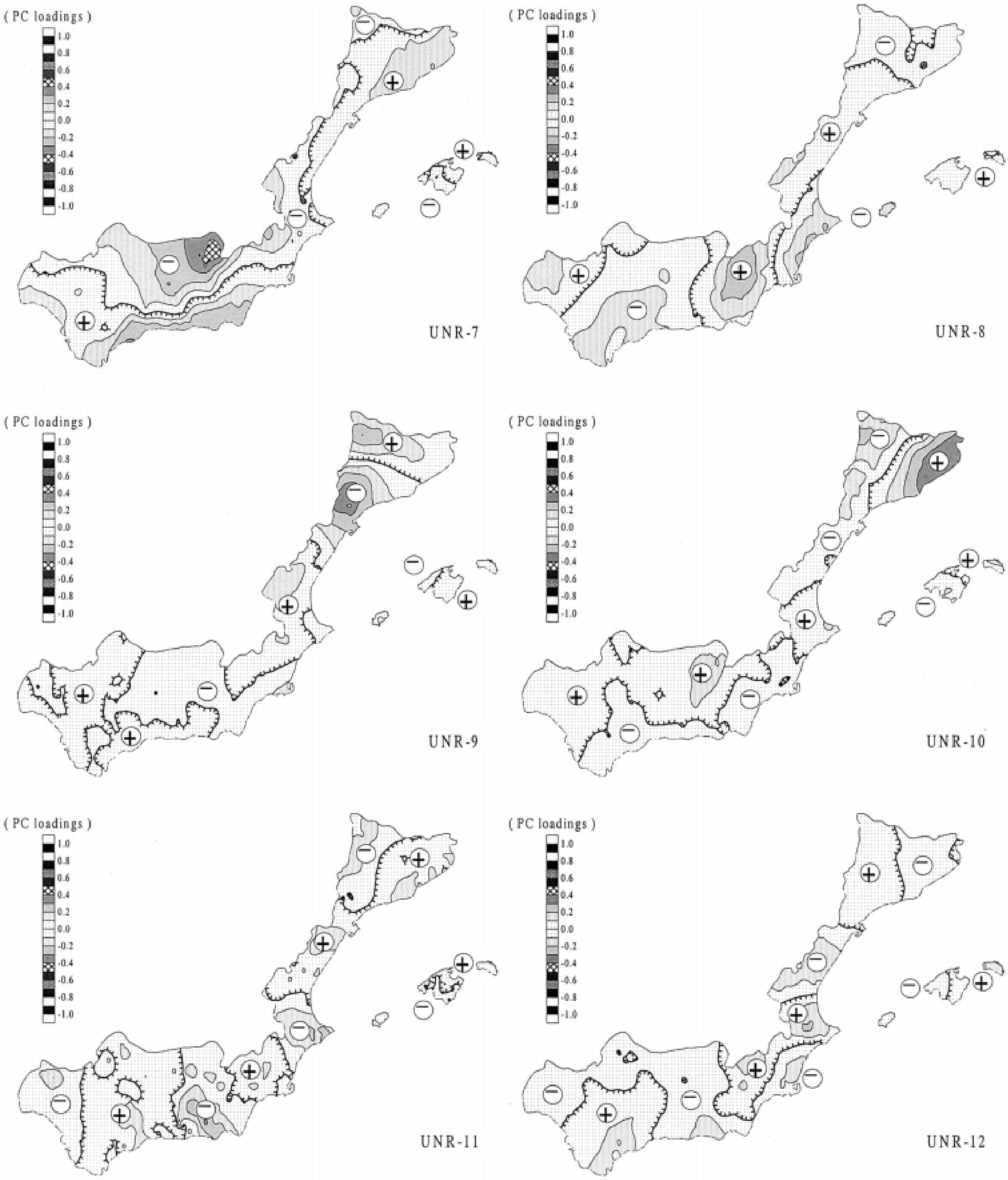

Figure 3 (*Continued*)

west of the Straits of Gibraltar (see Figure 1 for locations), are effective for western Andalucía, whereas the eastern regions are strongly sheltered by the marked topography of the whole Iberian Peninsula. Reverse sheltering action occurs when an atmospheric disturbance occurs east of the Straits of Gibraltar on the Mediterranean side. In this latter case, it is the eastern regions which are favoured by moist easterly flows.

UNR-3 (8.09% of the total variance) also shows a directional rainfall contrast, but in this case grading from northwest to southeast. Rainfall activity on Atlantic fronts approaching from higher latitudes and disturbances in the northern part of the Mediterranean basin would be consistent with this pattern. Under these circumstances, the southeast (around Murcia; Figure 1) is strongly sheltered. In fact, this part of Spain is the driest of all Mediterranean Spain, with annual rainfalls of less than 300 mm in some areas (Romero *et al.*, 1998).

UNR-4 (5.89% of the total variance) represents a contrast between the northern peninsular regions and the Balearic Islands. The structure found in the peninsular regions, with maxima in the Pyrenees and Sistema Ibérico (Figure 1), bears some resemblance to summer rainfall occurrence, which is dominated by convection triggered by the mountains and sustained by the Iberian thermal low (Alonso *et al.*, 1994). In contrast, the distribution of loadings over the Balearic Islands is similar to rainfall patterns experienced at times when low pressure centres occur to the east of these islands (Romero *et al.*, 1999), rather than the summer convection situation. UNR-7 (1.59% of the explained variance) exemplifies very well the mutual isolation between the interior and the southern coasts of Andalucía imposed by the southwest–northeast-aligned topographic ridges. In particular, for northwesterly winds associated with the passage of cold fronts over the Iberian Peninsula, Sierra de Cazorla and Sistema Subbético mountains (Figure 1) actively stimulate rainfall generation, whereas the southern coasts are, in contrast, well sheltered (Romero *et al.*, 1999). UNR-9 illustrates contrasts between the lowlands of southern Catalonia (one of the driest zones of the whole region), and the adjacent mountains of the Pyrenees and the Sistema Ibérico (Figure 1). UNR-10 shows a transition between eastern and western Catalonia: Atlantic fronts become effective only over the Pyrenees, but cyclones may often develop east of Catalonia (e.g. in the Gulf of Lyons), and in contrast, these favour rainfall in northeastern Catalonia.

The other PCs exhibit some loading patterns which may be logically linked to topography, for example the two maxima of opposite sign in UNR-6 north and south of Sierra de Aitana (Figure 1), and the contrast between lowlands and uplands in the southeast by UNR-8. In general however, the overall magnitude of these PC loadings, either positive or negative, is too small for clear interpretation.

3.2. Oblimin-rotated PCs

The purpose of rotating extracted PCs is to obtain a solution characterized by its simple structure (Richman, 1986). As Cattell (1966b) recommends, the degree of simplicity of the structure must be checked by inspecting the pairwise plots of the rotated loadings. Such plots involve the selection of pairs of PCs as the x- and y-axes, and plotting their PC loadings. A simple structure demands that the points (the 410 sites in this S-mode analysis) are concentrated within the hyperplane of one PC or near the origin, i.e. within the hyperplanes of both PCs (Richman, 1986).

Figure 4 shows selected examples of pairwise plots for combinations of the first three unrotated PCs. All of these plots define very weak simple structures, since there is a large number of 'complex variables' (those with large loading coefficients on both PCs). Figure 5 shows the analogous plots for the Oblimin-rotated PCs. As can be seen, rotation clearly succeeds in yielding strong simple structures. Only the pair OBL-2/OBL-5 (not shown) defines a moderate simple structure according to the convention suggested by Richman (1986).

Figure 6 shows the loadings distribution for the Oblimin-rotated PCs. Compared with the non-rotated solution (Figure 3), we find a clearly superior regional definition for the rotated solution, with the maxima in Figure 6 more closely associated with the topography of the region. The spatial distribution of PC loadings for OBL-1 closely matches the lowlands of the southeast (Figure 1). OBL-2, with its maximum over around Sierra de Aracena, corresponds to the Atlantic sector of Andalucía, west of the Straits of Gibraltar. Positive loadings of OBL-3 are strongest over the Balearic Islands, with the greatest loadings over Mallorca. OBL-4 practically contours the topography of this eastern edge of the Pyrenees. OBL-5 is quite extensive and represents the south-facing areas of south-central Andalucía, with the highest loadings found east of Sierra de Ronda along the coast. OBL-6 loads in particular over the east-facing lands of Valencia, north of the Sierra de Aitana. OBL-7 is associated with the mountains in the interior of

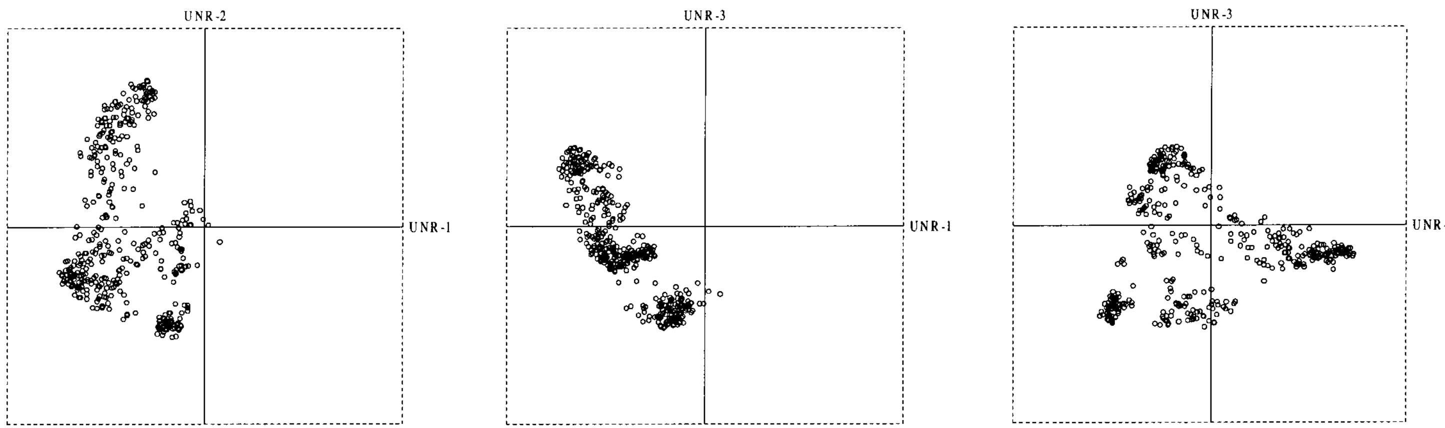

Figure 4. Selected examples of graphical pairwise plots for unrotated PCs

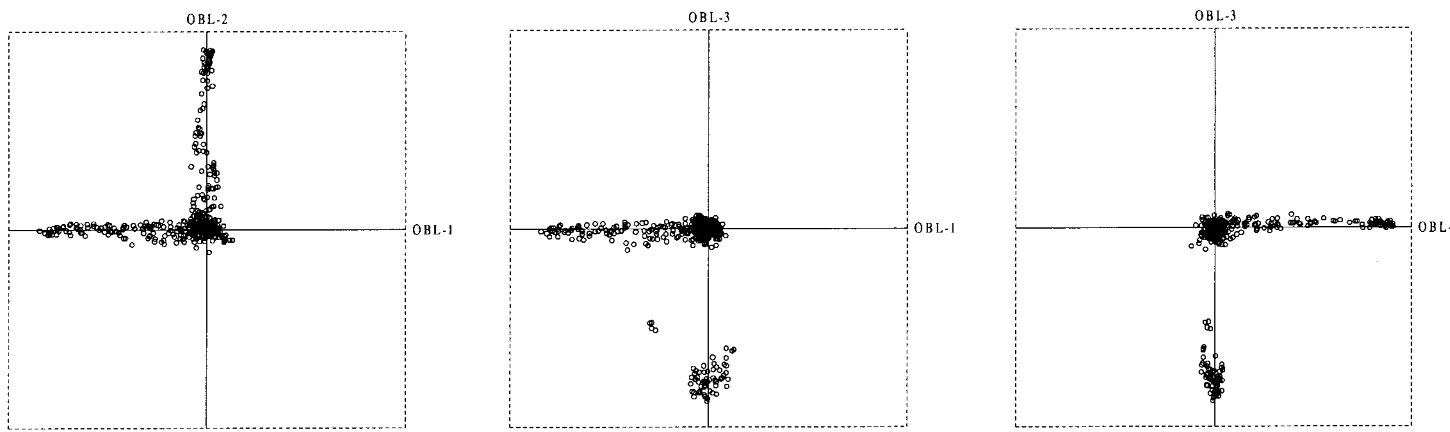

Figure 5. As in Figure 4 but for Oblimin-rotated PCs

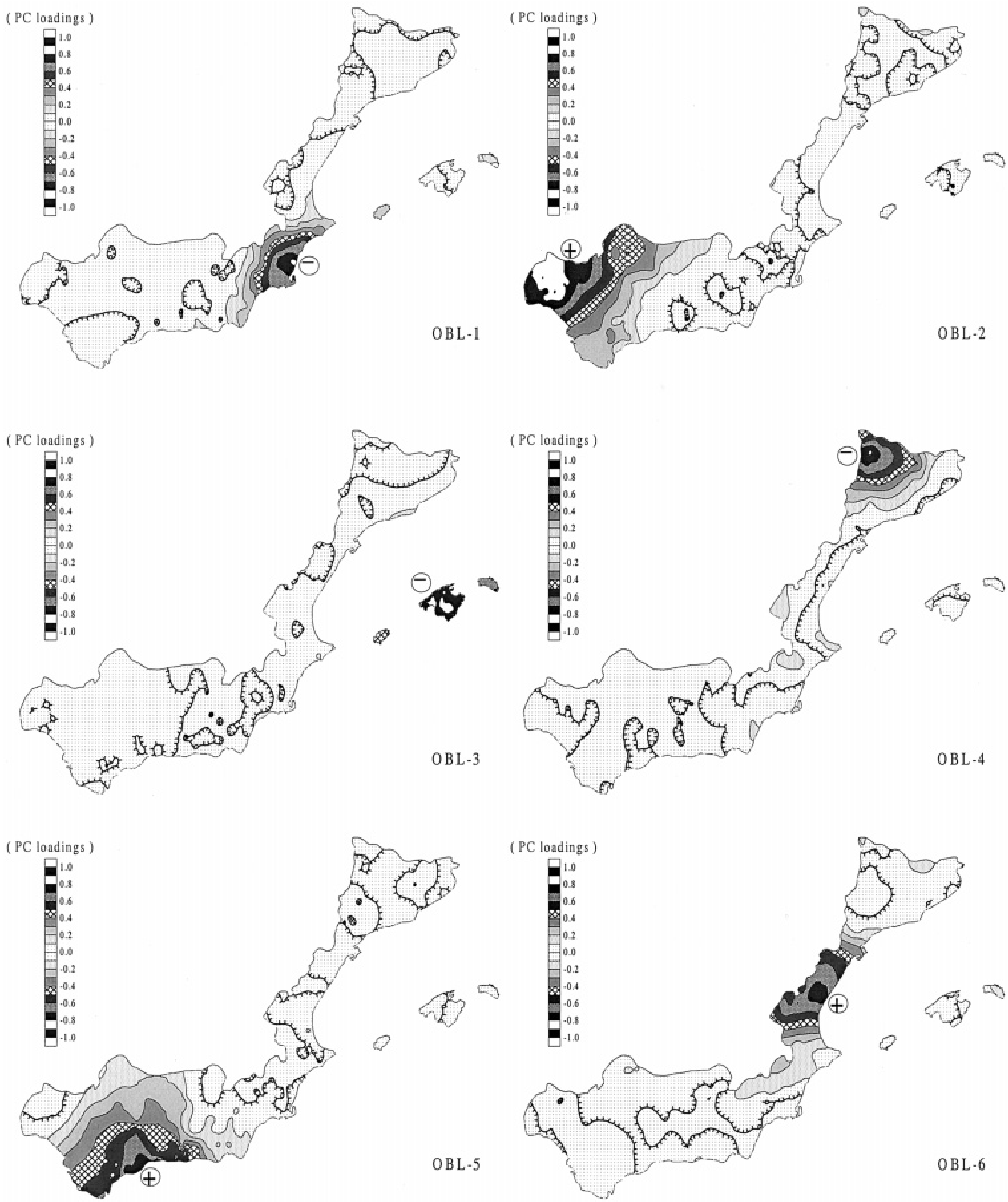

Figure 6. Spatial distributions of loadings for the 12 Oblimin-rotated PCs. Hachures along the zero isopleth point toward positive loadings; positive and negative zones are indicated

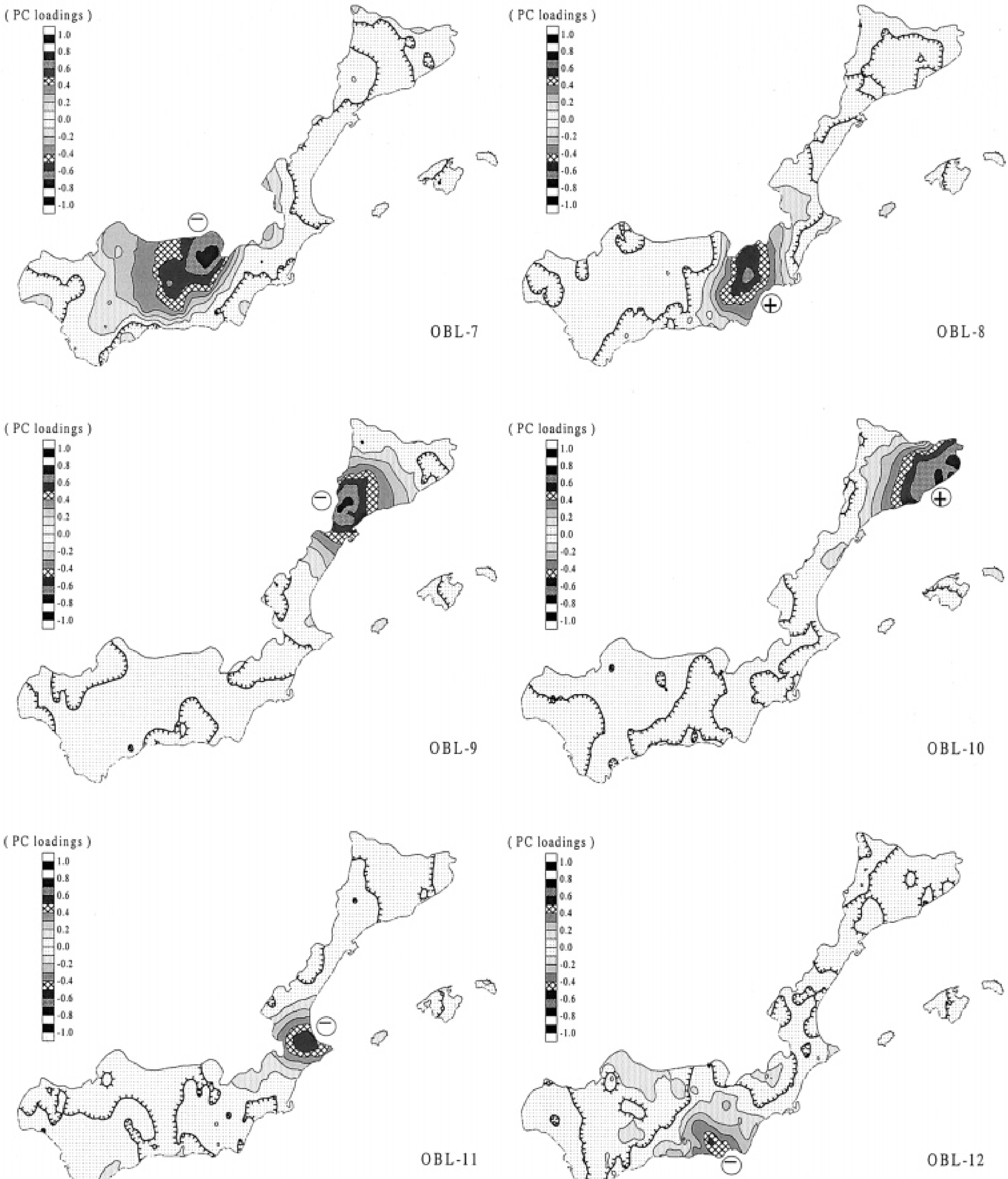

Figure 6 (*Continued*)

Andalucía, with their maximum over the Sierra de Cazorla. High OBL-8 loadings concentrate over the east-facing slopes of eastern Andalucía and Murcia. OBL-9 reaches its maximum approximately over the lower-lying areas of the interior of Catalonia. OBL-10 is concentrated over the coastal areas of Catalonia, east of the Pyrenees, and their immediate lower-lying hinterland. OBL-11 coincides with the well-exposed parts of the Sierra de Aitana, and OBL-12 is contiguous to OBL-5, with highest loadings over the south-facing lands of eastern Andalucía.

It seems, therefore, that the rotated solution provides a more precise sub-division of the region, thus making much better physical sense. Each of the subzones outlined in Figure 6 corresponds to a geographical area where a distinct set of mechanisms act either to enhance or suppress rainfall activity. These mechanisms are basically related to the configuration and position of rainfall-generating meteorological conditions and weather systems, or to the interactions (both enhancing and suppressing) between the topography and the rain-bearing airflows. These processes are summarized in Table II (see section on overlapping affinity areas). Note that a certain enhancing or suppressing process may be common to more than one region. What produces an apparently well-defined regionalization in the present case, is that for any two given regions among those outlined in Figure 6, there is at least one physical process that is relevant for only one of the regions, or that acts in the opposite sense in both regions. For example, both the OBL-1 and OBL-11 regions may experience rainfall enhancement at times of east–southeasterly winds. Equally, however, OBL-11 may also reflect enhancement during northeasterly winds, while the southeast (OBL-1) is sheltered from this wind direction. For details of the relief in these areas, see Figure 1.

4. REGIONALIZATIONS

4.1. *Hard affinity areas*

As indicated in the data base and methodology section, the hard regionalization for Mediterranean Spain is obtained by subjecting the 410 stations to CA (k-means method) using as input variables the loadings attained on the 12 unrotated PCs. It should be noted that the maps of Oblimin-rotated PC loadings shown in Figure 6 have already provided one regionalization solution, and here the suggestion made by Jolliffe (1987) that the application of CA to non-rotated PCs provides an alternative methodology is followed.

The final decision on the most appropriate number of clusters to generate is still an unresolved problem (Everitt, 1979), prone to subjectivity and generally dictated by experience. Several 'objective' methods have been suggested to detect a number of viable solutions depending on the level of detail desired. Among the 30 methods evaluated by Milligan and Cooper (1985), the best performance was attained by the *pseudo-F* test (Calinski and Harabasz, 1974). Thus, this test is considered here as guidance. The *pseudo-F* variable is formulated as:

$$\textit{pseudo-}F = \frac{A}{W}\frac{n-k}{k-1}, \tag{1}$$

where A and W are the among- and within-cluster sum of squares respectively, n is the number of objects (410 stations), and k is the number of clusters.

Figure 7 shows the results of the *pseudo-F* test. Local peaks in the test indicate an appropriate number of clusters (SAS Institute, 1985). A first clear maximum is found for six major affinity areas. This choice produces a very reasonable subdivision of Mediterranean Spain (Figure 8). One of the areas is formed by western Andalucía, with its eastern boundary following the eastern edge of the Guadalquivir Valley and crossing over Sierra de Ronda (Figure 1). We could refer to this area as the Atlantic zone, since its rainfalls critically depend on Atlantic flows. The second affinity area is formed by the rest of Andalucía, with the exception of the east-facing lands of its easternmost part. This resembles more a transition zone, with strong internal differences between the north and south induced by its prominent topography. Specifically, whereas Atlantic flows channeled through the Straits of Gibraltar and winds from the Mediterranean may be effective in the south, the north relies on Atlantic flows channeled along the Guadalquivir Valley or rain along fronts invading from the northwest. The third affinity area is formed

Table II. Summary of the 12 daily precipitation affinity areas for Mediterranean Spain (refer to Figure 11)

Region	Rainfall mechanisms
1. Atlantic sector of Andalucía (west of the Straits of Gibraltar, including Guadalquivir Valley and western Sierra de Ronda)	Atlantic depressions west of the Iberian Peninsula inducing southerly to westerly humid winds over western Andalucía. Sheltered by the high mountains of Andalucía for Mediterranean flows.
2. South-central Andalucía (around the Straits of Gibraltar and Sierra de Ronda, west of Sistema Subbético and Sistema Penibético)	Atlantic disturbances providing southwesterly to westerly flows over Andalucía. Rainfall enhancement by the Sierra de Ronda for northwest winds and Mediterranean flows.
3. Mountainous interior of Andalucía (around Sistema Subbético and Sierra de Cazorla)	Rainfall generation by northwest winds, principally linked to the passage of cold fronts over the Iberian Peninsula. Mountain-induced rainfalls for Atlantic flows channeled along the Guadalquivir Valley.
4. South-facing coastal lands of eastern Andalucía (around Sistema Penibético)	Rainfall development by southwesterly flows passing across the Straits of Gibraltar, and by southeasterly flows. Sheltering by the mountains of Andalucía.
5. East-facing slopes of eastern Andalucía and Murcia (to the east of Sierra de Cazorla and Sistema Penibético)	Substantial rainfalls for easterly-component flows, normally associated with well-marked troughs at upper levels over western Spain. Sheltering for other flows
6. Southeast lands (south to Sierra de Aitana) and Ibiza (Balearic Islands)	The most arid zone of the region. Strongly sheltered for all flows, except for southeasterly to easterly flows.
7. Central Valencia (around Sierra de Aitana)	Topographical enhancement, with the most substantial contribution for easterly or northeasterly winds.
8. East-facing lands of northern Valencia (north of Sierra de Aitana)	Rainfall generation for on-shore flows over the Gulf of Valencia, provided by south Mediterranean lows. Summer convection triggered by Sistema Ibérico. Rainfall sheltering by the Sistema Ibérico.
9. North Valencia–South Catalonia (between Sistema Ibérico and the Pyrenees)	Sheltering by the Pyrenees and Sistema Ibérico for most of the flows. Rainfalls for on-shore flows.
10. North interior of Catalonia (around the Pyrenees)	Topographical enhancement by the Pyrenees. Large contribution from Atlantic fronts circulating over northern Spain, warm season convection, and southerly moist flows associated with big Atlantic lows.
11. Eastern Catalonia (east of the Pyrenees)	Rainfall generation for on-shore flows, cold fronts from the north, easterly to northerly flows induced by depressions in the north of the Mediterranean basin and warm season convection. Sheltered by the Pyrenees.
12. Balearic Islands	Rainfall generation for a wide variety of flows. Rainfalls for cyclones to the east of the Balearics.

by the east-facing lands of Andalucía, Murcia and South Valencia, and the fourth affinity area is the rest of Valencia north of Sierra de Aitana. These areas should be called the Mediterranean zones, since they depend essentially on the Mediterranean easterly moist flows for much of their rainfall. The fifth and sixth affinity areas comprise Catalonia and the Balearics, respectively. Although these zones are also exposed to potentially rain-bearing easterly flows, they are also much more subject to rainfall associated with higher latitude disturbances, and are, in addition, directly affected by the cyclones which frequently develop in the northern part of the western Mediterranean Sea.

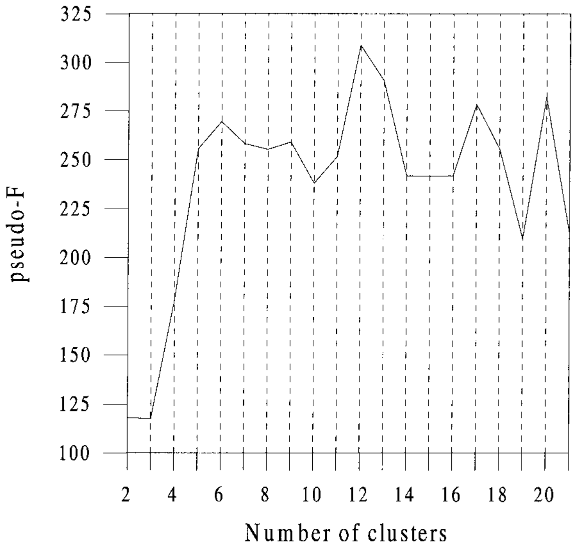

Figure 7. *Pseudo-F* test for several solutions of the hard regionalization

Figure 8. Hard regionalization of Mediterranean Spain based on six daily rainfall affinity areas

However, the simple six-cluster solution provides too coarse a regionalization to be useful. It omits some important spatial contrasts in rainfall which are known to be associated with regional topography. Therefore, it was decided to generate 12 affinity areas, which are shown in Figure 9. The 12-cluster solution is strongly supported by the *pseudo-F* test shown in Figure 7; the absolute maximum of the test is found at this solution, indicating a maximization of inter-cluster variability relative to intra-cluster variability when 12 groups are generated. Moreover, the solution is considered to be an appropriate compromise between the level of spatial resolution initially anticipated at the start of the project and the interpretation of affinity areas based solely on those main aspects of topography and synoptic features of atmospheric circulation (see next section). Conspicuously, the division into 12 affinity areas facilitates direct comparison with the overlapping solution developed in the next section, which also consists of 12 affinity areas (as many as there are PCs retained).

Solutions comprising 17 and 20 clusters may also be invoked from Figure 7. In general terms, the five additional areas in the former case (regionalization not shown) result from a separation of regions 3, 4, 5 and 12 (Figure 9) in eastern and western parts, and the isolation from region 2 of a few stations located along the coast and in the Guadalquivir Valley. The 20 affinity areas arising for the second case are shown in Figure 10. Comparing with Figure 9, it is observed that partitions of Valencia and Catalonia regions remain unchanged. However, the subdivision in Murcia and Andalucía (especially in its central and eastern parts where topography is very complex) becomes appreciably modified. In addition, the Balearic Islands are now divided into three affinity areas.

Figure 9. Daily rainfall affinity areas in Mediterranean Spain, resulting from the hard regionalization (12 clusters)

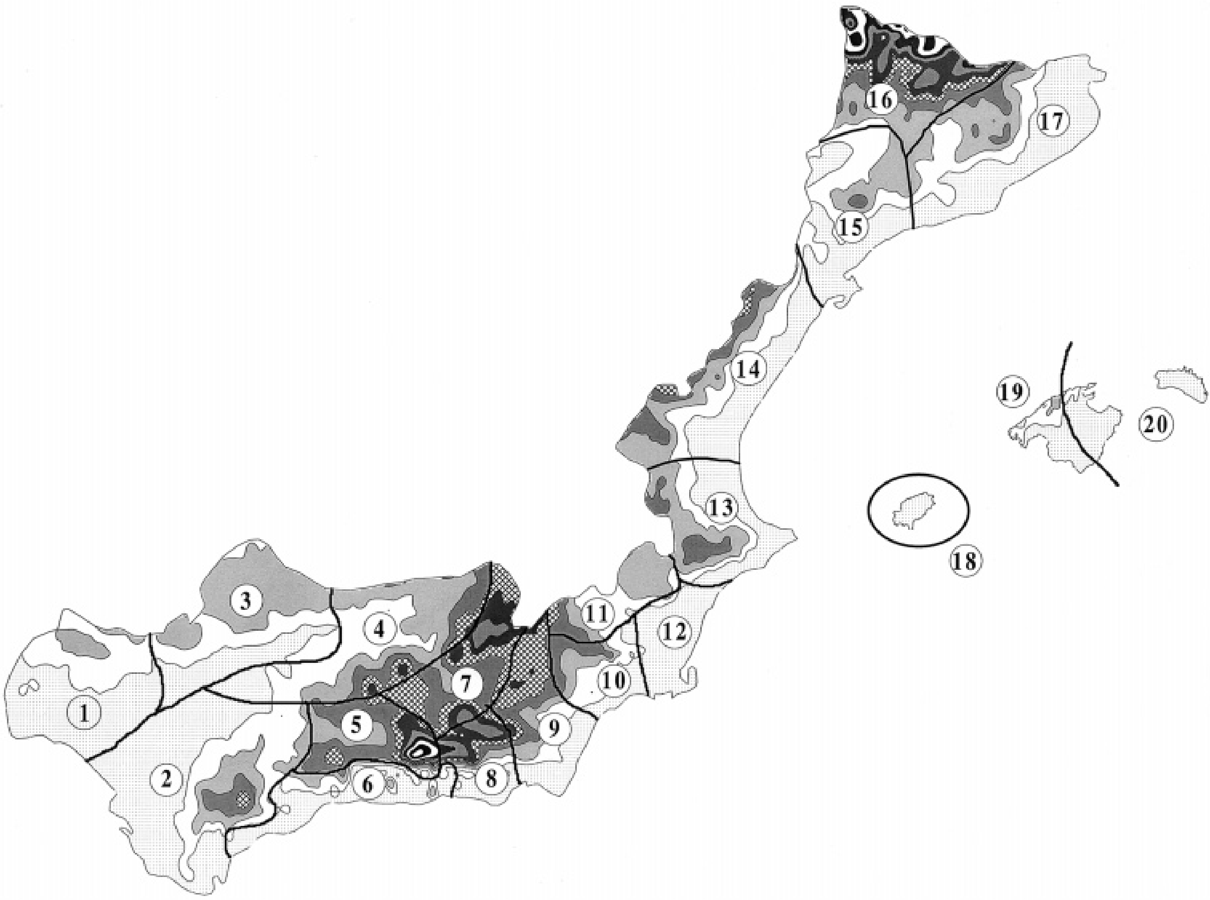

Figure 10. Hard regionalization of Mediterranean Spain based on 20 daily rainfall affinity areas

For Mediterranean Spain as a whole, the solution comprising 20 affinity areas provides an interesting regionalization. However, it contains too much detail for the purposes of linking the affinity areas with synoptic aspects of the weather systems. Some of the resulting areas are rather small and therefore mesoscale effects, such as the modification of airflows by local topography, should be considered for its interpretation. This is beyond the scope of this study.

4.2. Overlapping affinity areas

Figure 11 shows the 12 overlapping affinity areas obtained by drawing isopleths of the 0.25 loadings of the Oblimin-rotated PCs. There is a close correspondence between these resulting areas and those obtained for the 12-cluster solution shown in Figure 9. This confirms the appropriateness of the Oblimin method as an alternative to rigid (and therefore unphysical) cluster analysis methods. Apart from the nature of the boundaries, the only notable difference between the hard and the overlapping regionalizations is that region 4 appears more confined to the eastern part of the southern coasts of Andalucía for the latter case. In general, however, the two results may be regarded as essentially equivalent. Thus, the detailed discussion on the emergent overlapping affinity areas which follows is also valid for the hard regionalization model.

As noted in in the section on Oblimin-rotated PCs, similar rainfall-producing mechanisms may be effective in more than one affinity area. Accordingly, using the 11 typical daily rainfall patterns derived by Romero *et al.* (1999) (in the form of rainfall composites for 'similar' days; see Figure 12), some of these patterns may simultaneously embrace several affinity areas. Specifically, pattern S1 of Figure 12 comprises rainfall restricted largely to region 1; pattern S2 rainfall in regions 1, 2, 3, 4, 10 and 11; S3, rainfall in regions 2 and 3; S4 shows the most substantial rainfalls in regions 2 and 3; S5 combines regions 5, 6 and 7; S6 includes regions 6, 7 and 12; S7, regions 7, 8, 9 and 11; S8, regions 10 and 11; S9, almost exclusively region 10; S10, rainfalls in region 12 and to a lesser extent in region 11; and for S11 rainfall is practically exclusively confined to region 12.

This regionalization of Mediterranean Spain is summarized in Table II, which includes a list of relevant meteorological situations and physical processes responsible for rainfall activation, enhancement or suppression, in the derived affinity areas. These physical factors are necessarily linked to the importance of the sea/land boundary, and the topography of Mediterranean Spain and the surrounding regions, and the attributed operative processes are based on previous findings (in Romero *et al.*, 1998, 1999). In particular, an association of meteorological patterns with the rainfall composites derived in Romero *et al.* (1999) has been used. Such association was based on visual inspection of a sample of meteorological maps (as given by ECMWF 'gridded analyses'), for each of the clusters of days derived therein. As noted in Romero *et al.* (1999), such an approach constitutes only a first approximation. An objective classification of the synoptic and sub-synoptic circulation patterns associated with the typical rainfall distributions is the objective of a subsequent study.

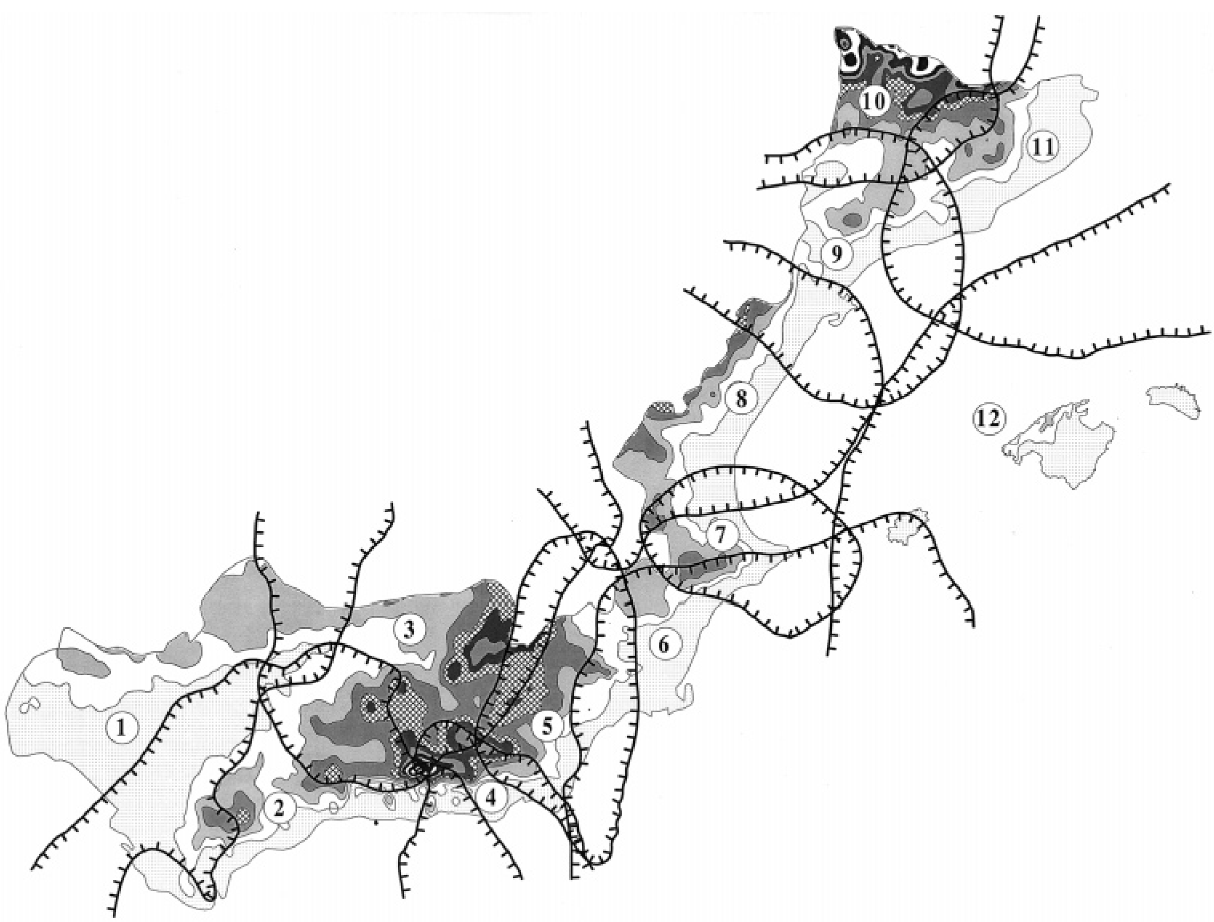

Figure 11. Daily rainfall affinity areas in Mediterranean Spain, resulting from the overlapping regionalization (12 clusters)

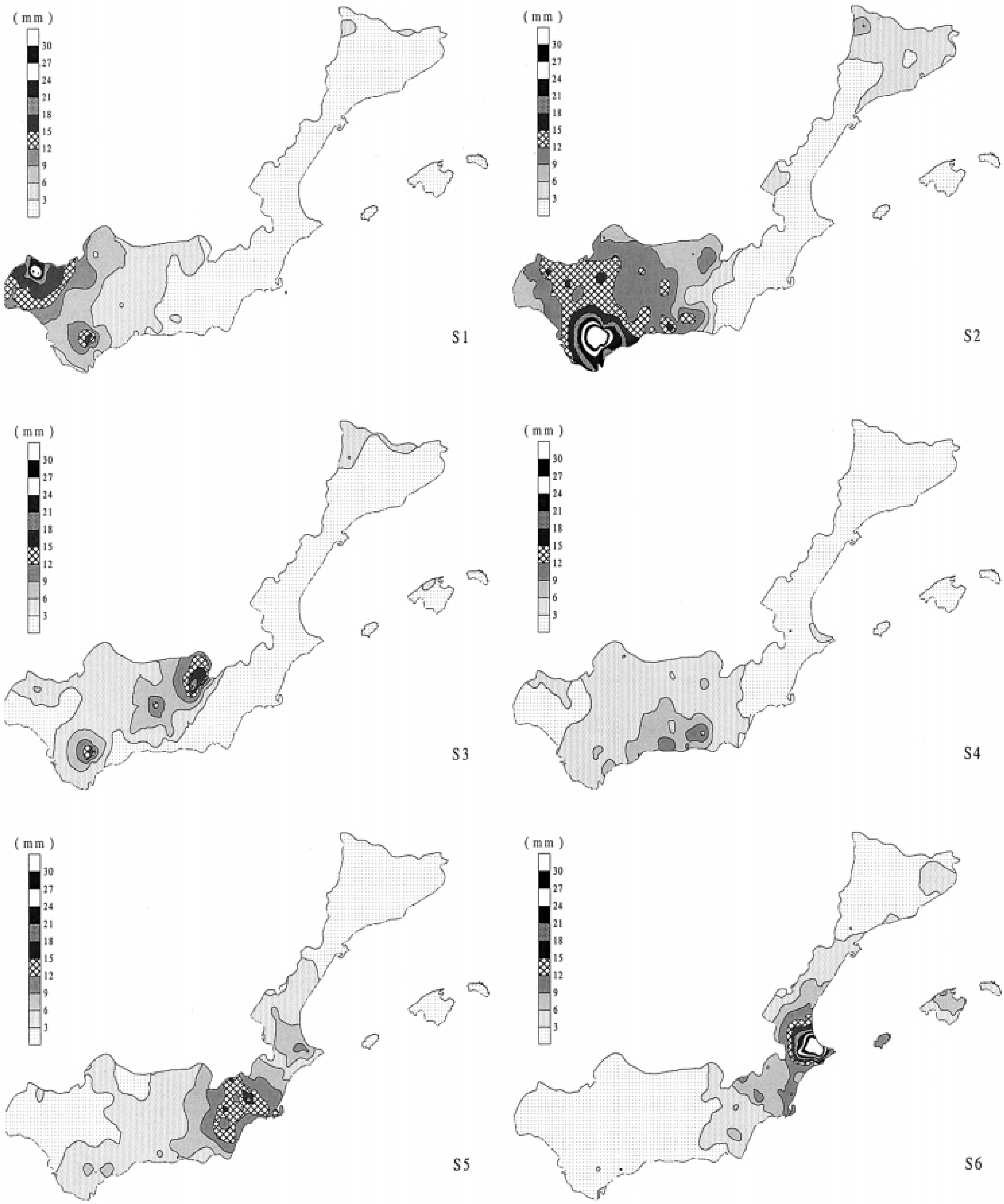

Figure 12. Daily rainfall composites for the 11 typical patterns of significant rainfall in Mediterranean Spain derived by Romero *et al.* (1999)

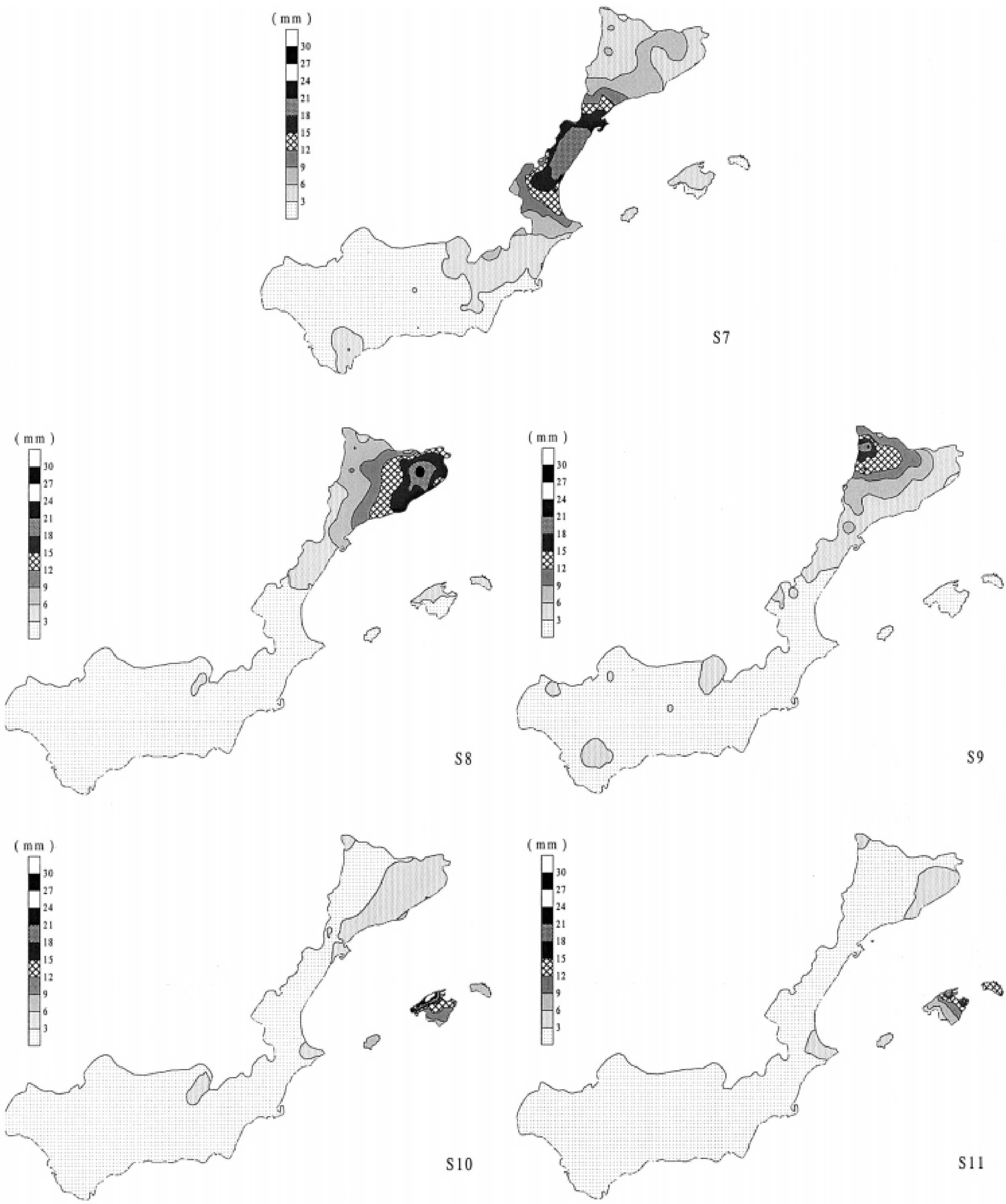

Figure 12 (*Continued*)

5. CONCLUSIONS

An objective regionalization of Mediterranean Spain for daily rainfall has not previously been obtained until now. This has been in spite of the potential value of such an assessment to studies of regional climatic variability and to improvements in regional public weather forecasting. This has been partially owing to the lack of availability of integrated data sets in format appropriate for detailed statistical analysis. In Romero *et al*. (1998), a daily rainfall data base for the 1964–1993 period was compiled for the Mediterranean regions of Spain, based on complete and homogeneous series at 410 regularly distributed raingauge stations. Such a data set resolves reasonably well the main topographical units of the region, and was found to be suitable for studies on the spatial and temporal variability of rainfall (Romero *et al*., 1998, 1999).

As part of a general objective devoted to the characterization of the precipitation of Mediterranean Spain, some of this work has focused on the time distribution of rainfalls. Considering only significant daily rainfalls (3941 during 1964–1993), the particular objective in this paper was to derive a sufficiently detailed regionalization for the area by clustering the stations of the data base with similar temporal precipitation series. Two approaches were followed: one producing the traditional hard regionalization through the application of cluster analysis to extracted S-mode principal components analysis; and the other producing a more physically consistent overlapping regionalization obtained simply by the rotation of extracted principal components (Gong and Richman, 1995). From both approaches, 12 daily rainfall affinity areas were derived, fulfilling the desired level of detail given the extent and complexity of the region, and yet permitting a simple interpretation in terms of the most important mechanisms of rainfall development and suppression that operate in the region.

Both regionalizations (Figures 9 and 11) are basically equivalent. They reflect the strong spatial organization of rainfalls as a result of the accentuated geographical contrasts of exposure to both Atlantic and Mediterranean rain-bearing flows. The obtained regionalization provides an important foundation for many further mesoscale climatic studies in Mediterranean Spain. Once the detail of the fundamental atmospheric patterns and the processes which generate, enhance or suppress precipitation in each of the affinity areas are identified, new predictive tools will be available for the regional forecaster in Mediterranean Spain. Such a classification can also help to assess the impact of future changes of the general and regional atmospheric circulation on the rainfall regimes in the region, although if such circulation changes are pronounced, the derived affinity areas may, of course, no longer apply.

ACKNOWLEDGEMENTS

Raw precipitation data were provided by the Instituto Nacional de Meteorología (INM) of Spain. This study has been sponsored by CICYT grant CLI95-1846.

REFERENCES

Alonso, S., Portela, A. and Ramis, C. 1994. 'First considerations on the structure and development of the Iberian thermal low', *Ann. Geophys*., **12**, 457–468.

Anderberg, M.R. 1973. *Cluster Analysis for Applications*, Academic Press, New York, 359 pp.

Bärring, L. 1987. 'Spatial patterns of daily rainfall in central Kenya: application of principal component analysis, common factor analysis and spatial correlation', *J. Climatol*., **7**, 267–290.

Bärring, L. 1988. 'Regionalization of daily rainfall in Kenya by means of common factor analysis', *J. Climatol*., **8**, 371–390.

Bonell, M. and Sumner, G.N. 1992. 'Autumn–winter daily precipitation areas in Wales, 1982/83 to 1986/87', *Int. J. Climatol*., **12**, 77–102.

Buell, C.E. 1975. 'The topography of empirical orthogonal functions', *Preprints Fourth Conference on Probability and Statistics in Atmospheric Science*, Tallahassee, Florida, American Meteorological Society, pp. 188–193.

Buell, C.E. 1979. 'On the physical interpretation of empirical orthogonal functions', *Preprints Sixth Conference on Probability and Statistics in Atmospheric Science*, Banff, Alberta, American Meteorological Society, pp. 112–117.

Calinski, R.B. and Harabasz, J. 1974. 'A dendrite method for cluster analysis', *Commun. Stat*., **3**, 1–27.

Cattell, R.B. 1966a. 'The scree test for the number of PCs', *Mult. Behav. Res*., **1**, 245–276.

Cattell, R.B. 1966b. *Handbook of Multivariate Experimental Psychology*, Rand McNally, Chicago, IL.

Everitt, B. 1979. 'Unsolved problems in cluster analysis', *Biometrics*, **35**, 169–181.

Font, I. 1983. *Climatología de España y Portugal*, Instituto Nacional de Meteorología, Apartado 285, 28071, Madrid, 296 pp.

Fovell, R.G. and Fovell, M.C. 1993. 'Climatic zones of the conterminous United States defined using cluster analysis', *J. Clim*., **6**, 2103–2135.

Gong, X. and Richman, M.B. 1995. 'On the application of cluster analysis to growing season precipitation data in North America east of the Rockies', *J. Clim*., **8**, 897–931.

Jolliffe, I.T. 1986. *Principal Component Analysis*, Springer, New York, 271 pp.
Jolliffe, I.T. 1987. 'Rotation of principal components: some comments', *J. Climatol.*, **7**, 507–510.
Meteorological Office, 1962. *Weather in the Mediterranean*, Vol 1, Air Ministry, 362 pp.
Milligan, G.W. and Cooper, M.C. 1985. 'An examination of procedures for determining the number of clusters in a data set', *Psychometrika*, **50**, 159–179.
Periago, M.C., Lana, X., Serra, C. and Fernández Mills, G. 1991. 'Precipitation regionalization: an application using a meteorological network in Catalonia (NE Spain)' , *Int. J. Climatol.*, **11**, 529–543.
Preisendorfer, R. W. 1988. *Principal Component Analysis in Meteorology and Oceanography*, Elsevier, New York, 425 pp.
Ramis, C. 1995. 'Las observaciones de la atmósfera fibre en Mallorca: una breve historia y algunos resultados', *Rev. Ciènc.*, **17**, 41–58.
Reiter, E.R. 1975. *Handbook for Forecasters in the Mediterranean*, Naval Postgraduate School, Monterey, California.
Richman, M.B. 1986. 'Rotation of principal components', *J. Climatol.*, **6**, 293–335.
Riosalido, R. 1990. 'Characterization of mesoscale convective systems by satellite pictures during PREVIMET MEDITERRANEO-89 (in Spanish)', *Segundo Simposio Nacional de Predicción*, Instituto Nacional de Meteorología, Apartado 285, 28071, Madrid, pp. 135–148.
Romero, R., Guijarro, J.A., Ramis, C. and Alonso, S. 1998. 'A 30 year (1964–1993) daily rainfall data base for the Spanish Mediterranean regions: first exploratory study', *Int. J. Climatol.*, **18**, 541–560.
Romero, R., Ramis, C. and Guijarro, J.A. 1999. 'Daily rainfall patterns in the Spanish Mediterranean area: an objective classification', *Int. J. Climatol.* (in press).
SAS Institute, 1985. *SAS User's Guide: Statistics*, SAS Institute, 959 pp.
STATISTICA, 1994. *STATISTICA for Windows (Volume III): Statistics II*, StatSoft Inc., Tulsa OK.
Sumner, G. and Bonell, M. 1990. 'The spatial organisation of daily rainfall in Wales during autumn–winter', *Weather*, **45**, 174–183.
Sumner, G., Ramis, C. and Guijarro, J.A. 1993. 'The spatial organization of daily rainfall over Mallorca', *Int. J. Climatol.*, **13**, 89–109.
White, D., Richman, M. and Yarnal, B. 1991. 'Climate regionalization and rotation of principal components', *Int. J. Climatol.*, **11**, 1–25.
Yarnal, B. 1993. *Synoptic Climatology in Environmental Analysis*, Belhaven Press, London, 195 pp.

