The Environments of Significant Convective Events in the Western Mediterranean

ELISA TUDURÍ

Instituto Nacional de Meteorología, Centro Meteorológico de Baleares, and Grupo de Meteorología, Departament de Física, Universitat de les Illes Balears, Palma de Mallorca, Spain

CLEMENTE RAMIS

Grupo de Meteorología, Departament de Física, Universitat de les Illes Balears, Palma de Mallorca, Spain

(Manuscript received 28 November 1995, in final form 31 January 1997)

ABSTRACT

The environmental characteristics associated with 313 significant convective events in the western Mediterranean are investigated using radiosonde ascents made in Mallorca (Spain). The events are separated into five groups, based on the observed event (hail, heavy rain, ''dry'' storms, storms with heavy rain, and tornadoes). Classic stability indices, as well as values of convective available potential energy and helicity, are considered for each group. These traditional convective indices appear not to provide good guidance for discriminating environments associated with each group of events. In order to classify the environments, each sounding is defined by means of 34 variables that describe the thermal and humidity vertical structure, instability, precipitable water, and helicity. A cluster analysis shows that four different vertical structures appear. Each kind of event shows preference for the environments defined by a cluster. A simple method is presented for sounding classification using the four categories obtained from the cluster analysis. The method looks for the best correlation between a particular sounding and those defined by each cluster.

1. Introduction

The area of the Mediterranean Sea and surrounding lands comprising eastern Spain, southern France, Corsica, Sardinia, and northern Africa are known collectively as the western Mediterranean; see Fig. 1 (Meteorological Office 1954). In this area, significant convective weather events (Table 1) are not unusual, especially heavy rains, which are frequent during autumn, representing a climatic characteristic of the region. Most of the observational sites in eastern Spain have recorded precipitation amounts over 200 mm (7.9 in.) in 24 h (Font 1983). There are examples of even heavier rain; for example, on 15 November 1985 in Ibiza (Balearic Islands) rainfall amounts exceeding 200 mm (7.9 in.) were recorded in 2 h (Ramis et al. 1986), on 5 September 1987 in Mallorca (Balearic Islands) a comparable amount in a similar time was registered, and on 3 November 1987 in Gandia (Valencia region) more than 800 mm (31.5 in.) fell in 24 h. Hail events are much less frequent than heavy rain episodes and have their maximum frequency in winter, at least in the Balearic Islands (Gayá 1984), although the most damaging events tend to occur in summer and autumn. Very large hail is infrequent, but hailfalls with 6-cm- (2.4 in.) diameter hailstones were observed on 15 August 1954 and on 26 August 1968 in Mallorca (Miró-Granada 1969). Although few data on downbursts and tornadoes exist for the considered area, strong convective wind events of both types also have been observed. For example, the downbursts on 21 June 1984 and 9 September 1992, both in Menorca (Balearic Islands), and the tornadoes on 26 October 1991 in Mallorca, 8 October 1992 in Menorca, and the 31 August 1994 event in Catalonia (Fig. 2) have produced significant damage to buildings and trees.

Phenomenological and synoptic studies have been performed on particular heavy rain episodes using synoptic patterns derived from classic weather maps for the eastern Spanish coast (e.g., García-Dana et al. 1982; Ramis et al. 1986; Llasat 1987). More recently, objective diagnosis and mesoscale aspects have been considered for a case of heavy rain in Catalonia (Ramis et al. 1994). Heavy precipitation events have been related to the presence at low levels of a very warm and humid air mass over the western Mediterranean sea, generally located between the northern Africa and the eastern Spanish coasts, that is being advected toward Spain. At middle- and high-tropospheric levels, the flow over eastern Spain and the western Mediterranean Sea tends to be from the southwest, typically associated with a trough

Corresponding author address: Dr. Clemente Ramis, Grupo de Meteorología, Departament de Física, Facultat de Ciencies, Universitat de les Illes Balears, 07071 Palma, Spain.
E-mail: dfscrn0@ps.uib.es

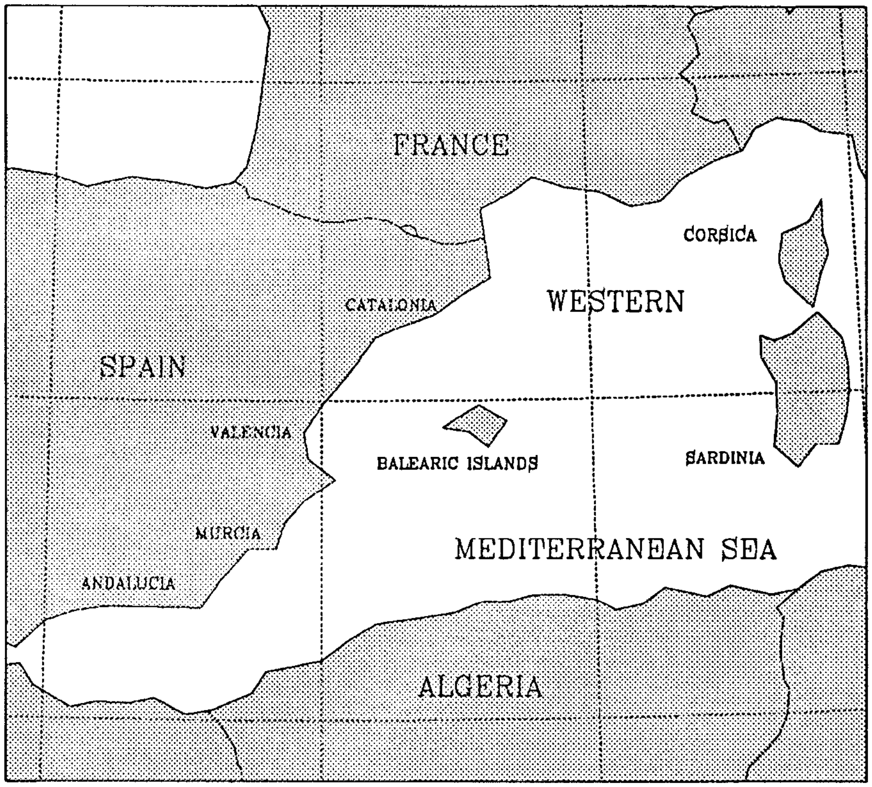

FIG. 1. The western Mediterranean region; some locations referenced in the text are indicated.

or low center located to the southwest or west of the Iberian peninsula. In such a synoptic situation, radiosonde observations made in the Balearic Islands (see Fig. 1) are usually representative of the thermodynamic environment in which the convection develops, since the Balearic Islands are within the warm and humid air mass. Llasat (1987) has shown, for example, that such data are useful for studying heavy rain situations in Catalonia. To our knowledge, only a single case study of a hail event has been made (Miró-Granada 1969) and also a short study of the environment, deduced from the radiosonde ascent in Mallorca, for the Menorca tornado (Gayá and Soliño 1993). The thermodynamic environmental characteristics in these cases seem to be very similar to that obtained for heavy rain events.

TABLE 1. Categories, criteria, symbols, and number of the considered events.

Event	Criteria	Symbol	Number of cases
Hail	Observation	Ha	77
Thunderstorm + heavy rain	Observation of thunder + rate > 30 mm h^{-1}, lasting at least 10 min	R1	114
Heavy rain	Rate > 30 mm h^{-1}, lasting at least 10 min	R2	52
Dry thunderstorm	Observation of thunder + rate < 30 mm h^{-1} + wind gust > 14 m s^{-1}	T	65
Tornado	Observation and damage	W	5
Fair weather	No significant event observed	F	23

It is known that convective events tend to favor situations having certain synoptic characteristics: upward large-scale velocity, convergence of water vapor at low levels, and instability, as well as subsynoptic (mesoscale) mechanisms for initiating convection (Doswell 1987). A diagnosis of such situations includes the study of static stability of air masses based on radiosonde data (using thermodynamic diagrams), which is a routine task in forecast offices. One of the most important aims of such an analysis is to determine the likelihood of convective development and estimate the strength of the convection and related meteorological phenomena within the scenario defined by the synoptic-scale meteorological situation. In this sense, stability indices have been developed in order to help forecasters estimate such characteristics quickly and easily.

One of the first developed, and arguably the most famous, was the Showalter index (Showalter 1953). Showalter presented threshold values of his index for different levels of convective development in the U.S. midwest, but these thresholds have been used widely around the world, sometimes without taking into account that they may be representative only for the spe-

FIG. 2. The tornado in Catalonia on 31 August 1994. (Courtesy of *Diario de Tarragona.*)

cific area in which they were developed. Other popular indices are the lifted index (Galway 1956), the K index (George 1960), the total totals index (Miller 1972), and the potential wet-bulb index (Pickup 1982).

Since the introduction of radar as a tool to monitor and observe thunderstorms, significant differences have been observed in the echoes depending on the degree of convective organization. The concepts of single-cell, multicell, and supercell convection were introduced (see, e.g., Weisman and Klemp 1986), and studies of the environments in which the different convective structures occur have been conducted. The vertical shear of the wind at low levels has been recognized as a very important factor for the development of supercells (see, e.g., Weisman and Klemp 1986; Doswell 1991).

Other studies have discussed the vertical structure of temperature and humidity data, indicating that the existence of a small temperature inversion at low levels and a dry layer at midlevels can be favorable for the development of strong winds. On the other hand the existence of a high humidity throughout the troposphere is favorable for the production of heavy convective rain. Three major temperature and humidity vertical structures have been identified over the United States and related to some significant weather events (see, e.g., Barnes and Newton 1986).

McCann (1994) has developed an index called WINDEX for estimating the strength of downbursts. WINDEX can be computed from environmental data measured by soundings, and McCann's results show it to correlate well with his observations.

Stability indices have been calculated by Ramis et al. (1986) and Llasat (1987) for several heavy rain episodes in eastern Spanish coast. However a more general study of the relation between well-known indices and convective weather has not been done. Moreover, no studies of the energetics and vertical wind structure have been performed for heavy rain events or when hail or strong winds from convective cells are present. This kind of information would be useful to local forecasters in the western Mediterranean; a related work suggesting this has been presented by Jacovides and Yonetani (1990) for the Cyprus area in the central Mediterranean.

A short discussion on the representation of real instability of the soundings in maritime sites like the Balearic Islands is needed. In warm seas, like the Mediterranean, the flux of water vapor is normally from the sea surface to the atmosphere. The capacity to modify

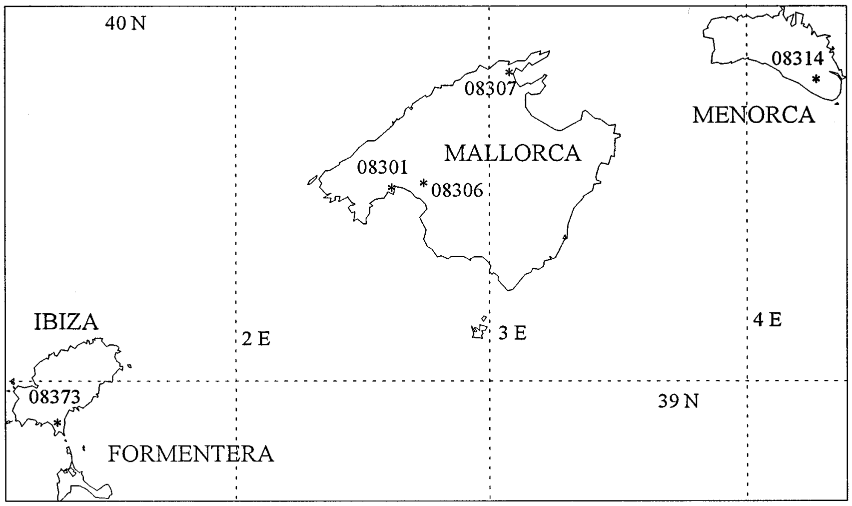

FIG. 3. The Balearic Islands; observatories used in the study are indicated by their WMO number. The Palma de Mallorca radiosonde station is located at station number 08301.

the airmass potential instability over such seas is very strong; air masses become convectively unstable if they reside over the sea long enough. Of course, the potential instability only develops if the meteorological situation becomes favorable. Indices that determine stability from low-level parcels often attain values (specially in summer) that indicate significant convection intensity but it does not occur. In contrast, sometimes when significant convection occurs, these indices do not show that fact.

Therefore, the first aim of this paper is to consider the forecasting value provided by the diverse stability indices, by convective available potential energy (CAPE, as a physical measurement of the instability), and by helicity (*H,* as an approximate measurement of the type and strength of temperature advection, as shown in appendix A) in different cases when significant convective events had occurred in the western Mediterranean. We also wish to develop techniques to identify the environments associated with each type of event. The identification of environments has been made by using a clustering technique. This objective technique allows for the classification of entities (environments in our case), defined by a set of variables, into groups so that the entities in each group are similar to one another but different from the entities to the other groups. Finally a simple classification scheme for individual soundings is presented.

It is known that thunderstorms develop within environments that exhibit high spatial and temporal variability (Doswell 1982; Brooks et al. 1994). Thus, all the conditions for a *proximity sounding,* as defined by Darkow (1969), are very difficult to achieve in this kind of study. The sounding balloons normally are released at a World Meteorological Organization (WMO) radiosonde station in the hours prior to 0000 and 1200 UTC, and the distance between two radiosonde stations in the western Mediterranean is over 350 km. In order to meet the conditions for a proximity sounding as closely as possible, we have restricted our study to the Balearic Islands, located approximately in the center of the western Mediterranean region.

2. Database

From the meteorological archives of the principal observatories (i.e., those that issue coded meteorological reports, as seen in Fig. 3), we have developed a 10-yr series of reports between 1984 and 1993. Among these data, we have chosen the days in which some significant convective event was observed, applying the criteria defined in Table 1 (which are a minor modification of the operational criteria used at the Regional Meteorological Center of the Balearic Islands). The selected days have been represented meteorologically using the Palma de Mallorca radiosonde ascent (08301 in Fig. 3) at the date and time nearest to the observed event. The method for constructing the database can be considered an *environment-to-circulation* assignment (Yarnal 1993) in which the aim is to associate atmospheric conditions with meteorological phenomena at the surface.

The event categories consider both single events (hail, heavy rain, ‘‘dry’’ storms, and tornadoes) and combined events (storms with heavy rain). A similar classification of convective events was done by Fuelberg and Biggar (1994). After a plot of the radiosonde ascents on thermodynamic diagrams and the elimination of 10 cases by evident errors in their codes, a total of 313 cases remained (Table 1). We added 23 soundings, randomly selected, corresponding to days of fair weather. Our purpose in introducing these fair weather days was to see if any significant differences exist between environments with or without significant convective phenomena. Thus, we have six groups of events altogether.

The distribution during the year of the 313 cases of

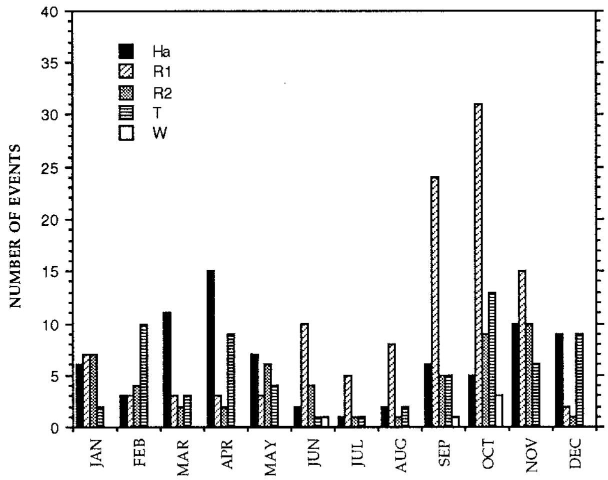

FIG. 4. Distribution of the events during the year.

significant convective events (Fig. 4) represents a short climatology of such events. The hail events are most frequent during the end of winter and the spring, whereas the thunderstorms producing heavy rain are much more frequent in autumn. Heavy rain from nonthundering clouds and dry thunderstorms also have their maximum frequency in autumn. The few tornadoes have been observed only in summer and autumn.

We have used the significant-level data in each sounding to obtain geopotential height, temperature, dewpoint, and wind at pressure surfaces from 1000 to 100 hPa in 10-hPa intervals, by interpolation using ln*p* as the vertical coordinate. The interpolated values constitute the database.

Because the events always occur at a distance less than 130 km from the sounding point (the distance to the farthest observation site from Palma), and the weather event occurs as much as 6 h before or after the nominal sounding times (0000 and 1200 UTC), the "Darkow" conditions are met approximately. From meteorological charts, we are reasonably certain that the soundings, for most of the considered cases, are representative of the air mass in which convection begins and develops.

3. Environment instability study

The sounding data values have been used to calculate some stability indices and other parameters related to the environment for each case.

a. Stability indices

We have considered the following stability indices:

1) lifted index (Galway 1956),

$$\mathrm{LI} = T_{500} - T'_{500};$$

2) lifted parcel (or lifted surface) index (see, e.g., Sanders 1986),

$$\mathrm{LP} = T_{500} - T''_{500};$$

3) Showalter index (Showalter 1953),

$$\mathrm{SI} = T_{500} - T'(850)_{500};$$

4) potential wet-bulb index (Pickup 1982),

$$\mathrm{PI} = \theta_{\mathrm{W500}} - \theta_{\mathrm{W850}};$$

5) K index (George 1960),

$$\mathrm{KI} = (T_{850} - T_{500}) + \mathrm{TD}_{850} - (T_{700} - \mathrm{TD}_{700});$$

and

6) total totals index (Miller 1972),

$$\mathrm{TT} = (T_{850} + \mathrm{TD}_{850}) - 2(T_{500});$$

where T_{500}, T_{700}, and T_{850} are the temperatures at 500, 700, and 850 hPa; T'_{500} is the temperature at 500 hPa of an air parcel representative of conditions in the lowest 100 hPa of the sounding lifted adiabatically up to its lifted condensation level (LCL) and further pseudoadiabatically up to 500 hPa; T''_{500} is the temperature at 500 hPa of the surface air parcel lifted in the same way; $T'(850)_{500}$ is the temperature that the 850-hPa parcel reaches at 500 hPa after lifting; θ_{W500} and θ_{W850} are the wet-bulb potential temperatures at 500- and 850-hPa levels, respectively; and, finally, TD_{700} and TD_{850} are the dewpoint temperatures at 700 and 850 hPa.

The first three stability indices are based on the temperature difference between a parcel lifted to 500 hPa and the observed temperature at that level. The other indices (KI, TT, PI), do not consider explicitly the evolution of any lifted parcel, but rather emphasize the temperature and humidity structure of the air mass.

7) For WINDEX (McCann 1994),

$$\mathrm{WI} = 5[H_M R_Q(\Gamma^2 - 30 + Q_L - 2Q_M)]^{0.5};$$

where H_M is the height of the zero temperature level in km, Q_L is the mean mixing ratio in the lowest 1 km, Q_M is the mixing ratio at the melting level, R_Q is $Q_L/12$ (but R_Q takes the value 1 if $Q_L/12 > 1$), and Γ is the lapse rate in degrees per kilometer from the surface to the melting level. This equation estimates the maximum wind gust in knots that a storm downdraft can develop when it reaches the surface.

8) CAPE: This parameter is a measure of the maximum energy available to a representative parcel rising vertically through an undisturbed environment. It represents a more physical measurement of the instability than that indicated by the indices, in part because it is not keyed to mandatory pressure levels. CAPE is calculated by (see Weisman and Klemp 1986)

$$\mathrm{CAPE} = g\int_{\mathrm{LFC}}^{\mathrm{EL}} \frac{\theta - \theta_e}{\theta_e}\,dz,$$

where LFC is the level of free convection and EL

TABLE 2. Mean (m) and standard deviation (sd) of the indices and precipitable water for the different categories of events.

Event index	Ha		R1		R2		T		W		F	
	m	sd	m	sd	m	sd	m	sd	m	sd	m	sd
LP	−0.1	4.1	−1.7	4.4	0.8	3.3	0.6	3.3	−6.0	3.2	3.9	7.5
LI	2.1	3.3	0.5	3.1	2.4	2.8	2.1	3.1	−2.0	2.6	6.5	5.3
SI	3.6	3.1	2.8	2.6	3.9	2.5	4.0	3.0	1.7	2.3	7.5	4.9
PI	1.1	1.7	0.5	1.7	1.3	1.5	1.3	1.6	−0.2	1.3	2.2	2.9
KI	19.0	10.3	23.9	9.4	21.7	10.2	19.8	8.7	25.8	4.4	2.8	17.7
TT	49.5	5.9	47.3	4.0	46.8	4.9	48.2	6.7	48.0	2.8	38.2	8.8
WI	24.8	10.8	26.6	15.5	20.1	12.7	23.0	11.2	39.8	9.6	17.2	16.6
PW	17.0	6.7	25.1	7.5	22.1	6.0	18.0	6.2	27.2	6.9	16.3	6.8

is the equilibrium level of the representative parcel, g is the acceleration due to gravity, θ is the potential temperature of the rising parcel, and θ_e is the potential temperature of the environment. On an energy-conserving thermodynamic diagram, CAPE is represented by the area between the environmental temperature profile and the lifted parcel temperature profile between its LFC and its EL.

This definition often refers to the surface parcel (see Doswell and Rasmussen 1994 for a discussion), but surface-based CAPE can be zero while some parcel at higher levels can have positive CAPE. We have calculated the CAPE for all the parcels from the ground to 500 hPa. We assigned to each sounding the CAPE of the *lowest* parcel that has a positive value since it represents the lowest parcel that can develop convection. Therefore, another parameter appears: the positive CAPE level (PCL), which is the pressure of the lowest parcel with positive CAPE. We have assigned the surface pressure to PCL when CAPE is zero for all the considered parcels.

9) Helicity (H) is defined by (Lilly 1986)

$$H(z) = -\int_{z_0}^{z} \mathbf{k}\cdot\left(\mathbf{V} \times \frac{\partial \mathbf{V}}{\partial z}\right) dz,$$

where $\mathbf{V}$ is the horizontal wind and $\partial\mathbf{V}/\partial z$ is the vertical shear of the horizontal wind. It can be interpreted as an approximate measure of the strength of temperature advection (see appendix A) and is equal to minus twice the signed area defined by the hodograph between z_0 and z. If the wind veers (backs) with height, H is positive (negative) and represents warm (cold) advection. We use $z - z_0 = 3$ km, where z_0 is the height of the pressure level indicated by PCL, looking for temperature advection just above the lowest parcel able to develop convection.

10) Precipitable water (PW): The precipitable water has been obtained by

$$\mathrm{PW} = g^{-1} \int_{p}^{p_0} q\, dp,$$

where q is the specific humidity and p and p_0 represent the pressure levels of the layer under consideration. We have used $p_0 =$ the surface pressure and $p = 500$ hPa.

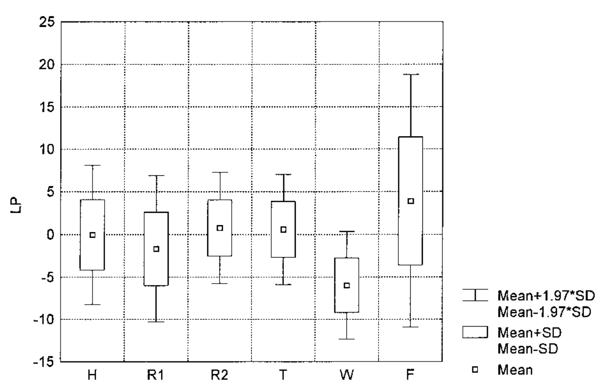

FIG. 5. Boxplot showing the distribution of lifted parcel index for each considered event.

b. Results

The preceding indices and parameters have been related to the six groups of considered events, with the results shown in Table 2 for the indices. Generally speaking, the differences in the means among the categories are not very large (relative to the dispersion). The lifted index reveals the largest differences in mean values among the various event categories. WINDEX, although designed to estimate the maximum wind gust from downbursts, appears to be related reasonably well to the intensity of the events, showing its smallest mean value for the fair weather events and its largest mean value for the tornado cases.

Figure 5 serves to illustrate the typical distribution of these indices, showing (for the LP index) the mean and standard deviation for the different events. The large standard deviation of these values is a well-known characteristic of indices, arising in part from issues of representativeness and perhaps also from changes in the atmospheric structure in the time between the nominal sounding time and the event.

Plots of CAPE versus H can be used to show the relation between stability and temperature advection for

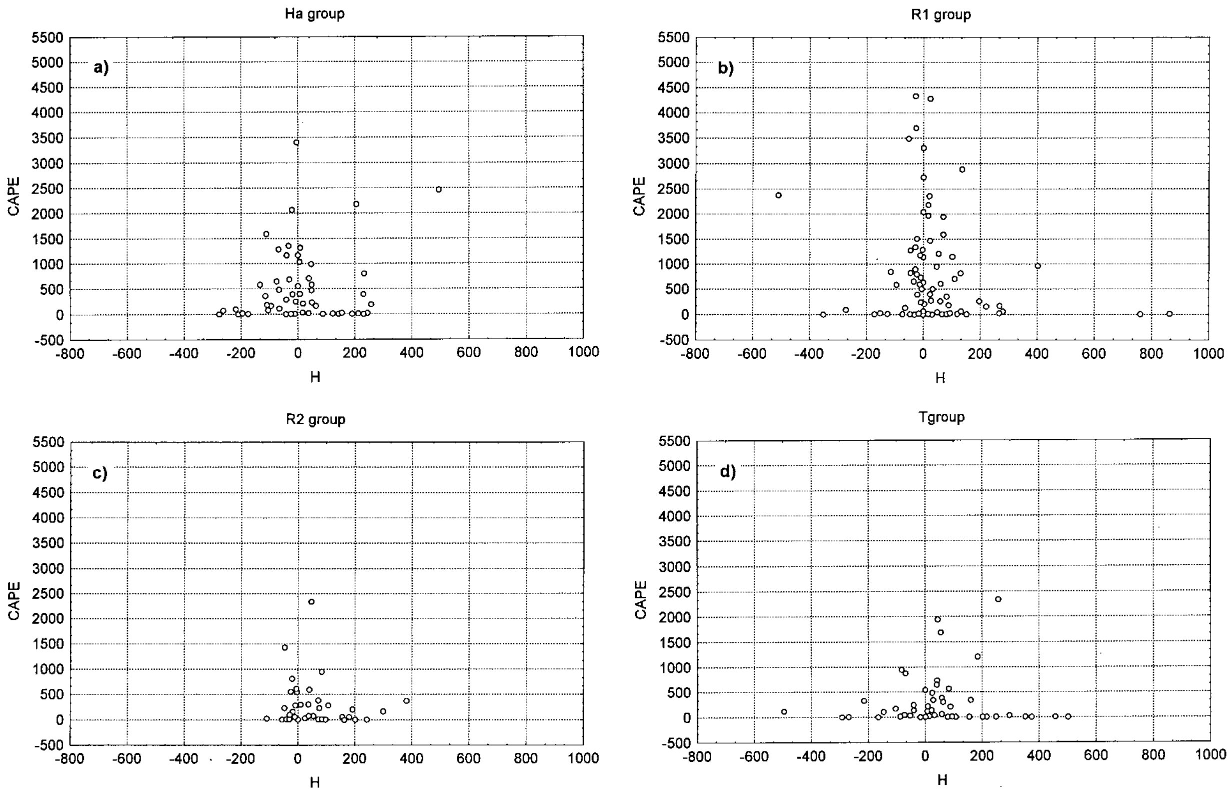

FIG. 6. Scatterplot of CAPE vs H for (a) Ha events, (b) R1 events, (c) R2 events, and (d) T events.

the events we have considered (Fig. 6). No clear insights can be observed among the different events.

4. Environment classification

The preceding look at diverse stability indices has suggested that many of them are inappropriate for the western Mediterranean weather events of interest. Basically, they have been developed to deal with different weather events in different geographical regions than ours. In order to look for environmental descriptors related to the events we have considered, we have chosen to follow a different strategy than developing indices or adapting index thresholds to suit local events and climatology. In particular, a *cluster analysis* has been applied to the soundings, using the *k*-means method (see appendix B). Cluster analysis has been widely used in climatology [e.g., Sumner et al. 1993, 1995; Gong and Richman 1995 (especially their appendix A, where an extensive list of applications of cluster analysis to different problems is included)], but as far we know, this methodology has not yet been applied to classify soundings. Our aim is to identify environments associated with different events in a manner similar to that used by Barnes and Newton (1986).

Prior to performing the cluster analysis, we have chosen to characterize the soundings by means of a set of variables that define the atmospheric state. This set comprises a group of 34 variables that describe the vertical temperature structure, the vertical distribution of humidity, the latent instability of the air mass, the temperature advection, and the water content. The temperature structure has been defined from 1000 to 100 hPa at 50-hPa intervals by means of the *difference of temperature between the temperature given by the sounding and the mean temperature deduced climatologically.* The mean temperature for each particular day has been calculated by linear interpolation between the monthly mean states (Ramis 1976), assigning the monthly mean values to the central day of the month. Soundings at 0000 and 1200 UTC have been considered separately. The object of this transformation is to eliminate the seasonal and diurnal cycles. The vertical distribution of humidity has been defined by the dewpoint depression calculated at 50-hPa intervals from 1000 up to 500 hPa. Latent instability has been represented by means of two parameters: CAPE and the PCL. The type and strength of temperature advection has been introduced by means of H referenced to the PCL parcel as described above. Finally, the water vapor content has been measured by means of the PW calculated from the surface to 500 hPa.

As the different variables have different units, they have been standardized in order to obtain nondimensional variables with zero mean and unit standard deviation (see appendix B). Then the *k*-means clustering

TABLE 3. Intraclusters distances (mean and standard deivation).

	Mean	Std dev
Cluster 1	0.6536	0.2084
Cluster 2	0.6471	0.1801
Cluster 3	0.9874	0.3707
Cluster 4	0.7003	0.2742

TABLE 4. Interclusters distances.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	0	—	—	—
Cluster 2	4.1735	0	—	—
Cluster 3	6.8690	5.7494	0	—
Cluster 4	4.3710	8.0841	9.7955	0

method has been applied to the 336 cases (313 significant convective event days, plus 23 fair weather days). When using the *k*-means method, the number of clusters has to be chosen *before* the clustering process, so some tests were made looking for a good separation. The tests have been made through the intercluster distances when two to six clusters were built. We have constructed plots of the minimum intercluster distance versus the number of clusters and the point where the curve exhibits a sudden flattening indicates the appropriate value for the number of clusters, following similar criteria as used by Thorndike (see Everitt 1980). As a result of the tests, we decided upon four clusters. Tables 3 and 4 show the intra- and intercluster distances for the four considered clusters.

In the *k*-means method, the clusters are formed around an estimated initial partition known as *seed points* (see appendix B). Some tests were made in order to estimate the influence of the initial seed points on the final clusters. The results show that the differences between the clusters using different seed points are indistinguishable.

Figure 7 shows the mean of each variable for the four clusters. A very clear separation exists between clusters within the thermal vertical distribution and precipitable water and less for the humidity vertical structure (except for cluster 4), stability parameters, and helicity. It can be seen that clusters 1 and 4 are separated by thermal vertical structure, humidity at low levels, helicity, and precipitable water, while clusters 2 and 3 are most clearly distinguished by the humidity vertical structure, helicity, and precipitable water.

Cluster 1 is associated with cases that deviate slightly from the mean values in temperature, but the humidity is relatively high in the troposphere, there are low CAPE values and slightly low PCL values, the temperature advection is slightly negative, and the PW is moderate. Cluster 2 represents events with warm air in the troposphere and somewhat cold air in the stratosphere, very high humidity in the troposphere, high CAPE and PCL values, a dominance of warm advection, and high values of PW. Cluster 3 is associated with cases having warm air in the troposphere and cold air in the stratosphere, very high dewpoint depression (low humidity) throughout the troposphere, relatively small values of CAPE, high values of PCL, cold advection, and relatively small values of PW. Cluster 4 connects cases with cold air in the troposphere (especially in the high troposphere), warm air in the stratosphere, moderate humidity in the troposphere (except in the layer from 900 to 700 hPa where it is high), small values of CAPE, low values for PCL, a dominance of cold advection, and small values for PW.

Table 5 shows the number of cases in each cluster and the number of the different events that include the clusters, as well as the percent of the events in each cluster. From Table 5 it can be seen that more than 40% of the Ha events are included in cluster 4, so hail pref-

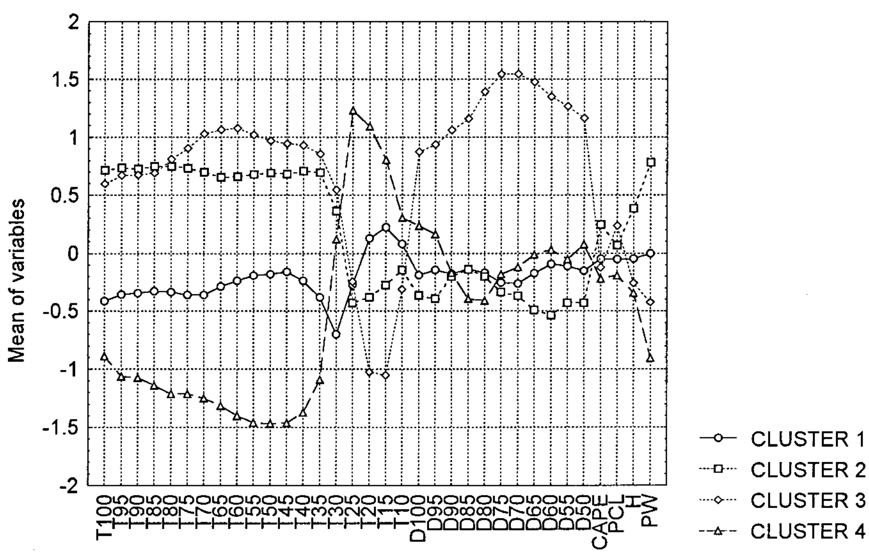

FIG. 7. Mean, standardized variables for each cluster.

TABLE 5. Number of cases in each cluster and frequency (in percent) for the different types of events.

Cluster	No. cases	Ha	(%)	R1	(%)	R2	(%)	T	(%)	W	(%)	F	(%)
1	107	22	(28)	43	(38)	17	(32)	21	(32)	2	(40)	2	(9)
2	109	12	(16)	53	(46)	23	(44)	17	(26)	3	(60)	1	(4)
3	49	10	(13)	7	(6)	6	(12)	6	(10)	0	(0)	20	(87)
4	71	33	(43)	11	(10)	6	(12)	21	(32)	0	(0)	0	(0)

erentially occurs with low-tropospheric and high-stratospheric temperatures, moderate humidity in the low troposphere, cold air advection, not very high instability, and relatively low water vapor content in the troposphere.[1] Also, more than 30% of the dry thunderstorms appear in this cluster. The R1 and R2 events are included preferentially in cluster 2 with the 45% of the cases. Heavy rain preferentially occurs with warm air in the troposphere, very high humidity, relatively large instability, warm air advection, and very high PW values in the lower troposphere. The W events also prefer this type of environment, although the number of cases may not be statistically significant. Finally, the F (fair weather) events strongly prefer the environment represented by cluster 3: high temperatures and very low humidity in the troposphere, moderate stability, moderate cold advection, and low values of PW.

As an example, Fig. 8 shows the distribution of the Ha and R1 events during the year and the number of events that correspond to each cluster. It can be seen that hail events (Fig. 8a) actually occur in environments represented by clusters 1 and 4 throughout the year, but hail events can also appear with environments represented by clusters 2 and 3. The R1 events (Fig. 8b) occur mainly during the autumn in environments represented principally by cluster 2.

5. Assigning a sounding to a cluster

A very simple method can be used to assign any individual sounding not included in the original database to a cluster. In addition to its simplicity, the method executes rapidly even on ordinary desktop computers.

The method can be summarized in the following steps.

1) Given a sounding, use the significant-level data to interpolate temperature, dewpoint depression, and wind at 50-hPa intervals from 1000 to 100 hPa.
2) Calculate the differences at each interpolated level between observed temperature and the mean temperature for the considered day. Consider the 0000 and 1200 UTC times separately.
3) Calculate CAPE, PCL, *H*, and PW.
4) Obtain the 34 variables to define the environment as temperature differences as indicated in (2) at 50-hPa interval from 1000 to 100 hPa, dewpoint depression at 50-hPa intervals between 1000 and 500 hPa, CAPE, PCL, *H*, and PW.
5) By using the mean and standard deviation for each variable obtained in the original database, standardize the 34 variables.
6) The correlation coefficients between these variables and the mean values of the variables for each one of the four clusters are calculated.

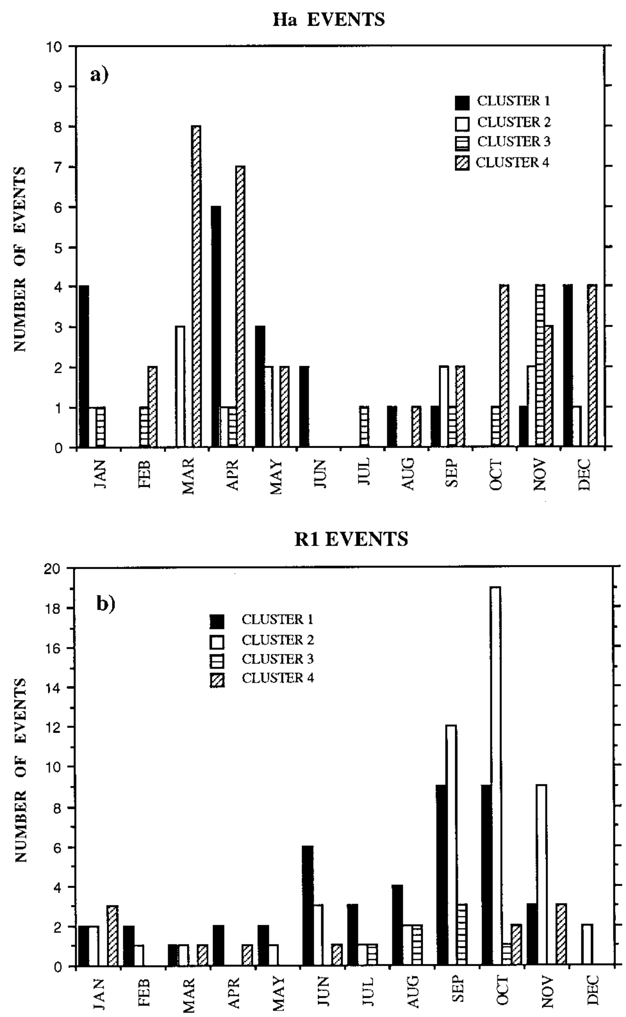

FIG. 8. Distribution during the year and corresponding clusters for (a) Ha events and (b) R1 events.

[1] Note that the observed hail typically is not large, so the processes associated with most of the hailfalls in the Balearic Islands may not be similar to the events that typically produce *large* hailstones in, for example, the United States.

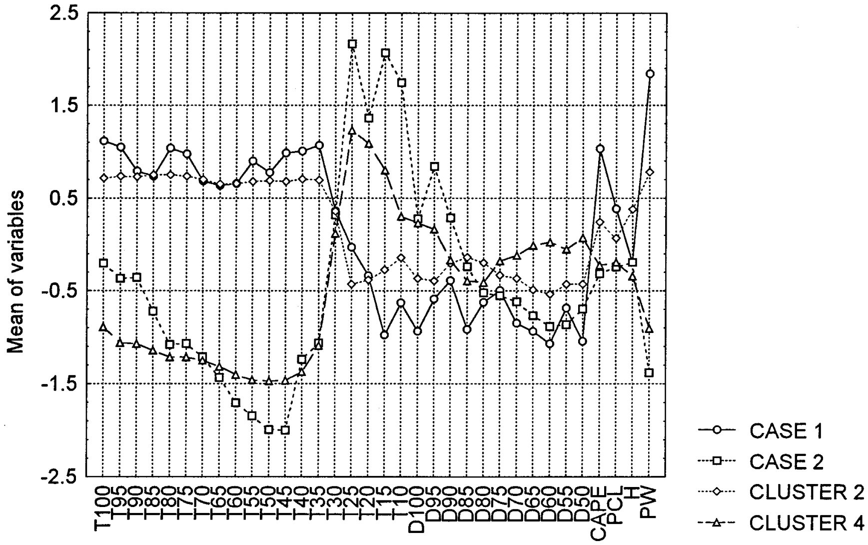

FIG. 9. Comparison of two soundings with clusters.

7) The sounding is assigned to the cluster with which it exhibits the highest correlation.

This analysis made in conjunction with the synoptic and mesoscale situation can suggest what events might be most probable with that environment.

As an example of the method, we present two soundings corresponding to 1994. Case 1 corresponds to an environment related to R1 events in the Balearic Islands during autumn, and case 2 corresponds to an Ha event during winter.

Comparison with the environment classification shows that case 1 has high correlation with cluster 2 (0.92) while the correlation with the other clusters is negative, in particular for cluster 4 (−0.72). Similarly, a high correlation coefficient between case 2 and cluster 4 (0.85) has been found while it has low correlation with cluster 1 and negative correlation with clusters 2 and 3 (−0.57 and −0.61, respectively). The method locates case 1 in the cluster that contains the highest proportion of R1 events, and case 2 is located in the cluster that contains the highest proportion of Ha events. Figure 9 shows the representation of the variables for cases 1 and 2 and clusters 2 and 4.

In order to do an evaluation of this technique, we have selected randomly 300 days during 1995. The meteorological reports from the observatories indicated in Fig. 3 show that a subset of 38 days presents some significant convective events as defined in Table 1. We have analyzed the soundings of these 38 days (as indicated in section 2) and we have added another subset of 11 randomly selected days corresponding to fine weather. The methodology indicated above has been applied to each of these 49 soundings (Table 6). Results show that the majority of Ha events occurred with environments represented by cluster 4, the R1 events with environments represented by cluster 2, the R2 events are included in clusters 1–3, the T events appeared principally in environments represented by cluster 2, and all cases of fine weather are included in cluster 3.

TABLE 6. Results of the evaluation of the method to assign a sounding to a cluster.

Event	No. of events	Cluster 1	Cluster 2	Cluster 3	Cluster 4
H	10	2	1	2	5
R1	13	3	8	0	2
R2	7	2	3	2	0
T	8	0	4	2	2
F	11	0	0	11	0

6. Conclusions and comments

A study of the characteristics of the environments, deduced from Mallorca (Spain) radiosonde ascents, related to significant convective events in the western Mediterranean has been made. The study of classic stability indices generally shows them to be poor parameters for separating the different classes of events considered. A study of the environment, including information deduced from the vertical distribution of temperature, humidity, and wind, would seem to be more appropriate than simply calculating indices.

A first effort to classify the environments in which significant convective events occur has been made. The soundings have been defined by means of 34 variables that include the vertical distribution of temperature, humidity, instability, helicity, and precipitable water. The *k*-means clustering method has been used to determine that four different environments appear. The different events show clear preferences for particular environments. Basically, hail events become more frequent with cold, humid air in the troposphere and cold advection, while heavy rain events occur with warm, humid air in all the troposphere and warm advection at low levels.

Given a sounding, a very simple method to assign the sounding to a cluster has been presented, consisting of finding the cluster to which the sounding is most highly correlated. It has been applied to 38 soundings corresponding to significant convective events and 11 fair weather days during 1995. The method locates the major number of soundings in the cluster that is most representative of the event associated with the sounding.

This study reveals that when convective indices are employed as forecasting aids outside the situations they were designed to describe, their forecast value can be relatively low. Of course, it is always possible to design new indices or to modify the threshold (critical) values of the old indices to fit new situations. However, the alternative methodology we have presented is very general and can be used to determine characteristic sounding structures associated with any type of event. It only requires the creation of a developmental dataset including 1) identification of cases where events occurred, 2) choosing a few cases where the events did *not* occur, and 3) soundings corresponding to the chosen cases. With such a dataset, following the methods we have described, it should be possible to develop sounding classifications suited to the specific weather events of interest and to the specific region of interest. We believe this approach will be an alternative method to developing indices.

The inclusion of dynamical synoptic forcings deduced from diagnostic studies, as variables to identify environments, probably can improve the discrimination of environments related to different kinds of events. Of course, the inclusion of synoptic forcings [e.g., divergence of **Q** vector (Hoskins and Pedder 1980) or convergence of water vapor at low levels] would need the use of secondary products from objective analysis performed jointly with the sounding.

Acknowledgments. The authors are grateful to the Instituto Nacional de Meteorología of Spain for providing, through the Meteorological Center of the Balearic Islands, all the data for carrying out this study. This work has been partially sponsored by the EC under Contract EV5V-CT94-0442 and CICYT Grants AMB95-1136-CE. We acknowledge Dr. Charles A. Doswell III (NSSL/NOAA), who is visiting the UIB under Grant DGICYT SAB95-0136, for his comments and help in the final version of the paper, and Prof. Sergio Alonso for his suggestions during the development of the work.

APPENDIX A

Helicity as a Measurement of Temperature Advection

Helicity is defined (Davies-Jones et al. 1990) as

$$H = -\int_{z_0}^{z} \mathbf{k}\cdot\left(\mathbf{V} \times \frac{\partial \mathbf{V}}{\partial z}\right) dz, \tag{A1}$$

where **V** is the horizontal wind and **k** the unit vertical vector. By using the hydrostatic equation, (A1) can be written in terms of pressure as a vertical coordinate:

$$H = -\int_{p_0}^{p} \mathbf{k}\cdot\left(\mathbf{V} \times \frac{\partial \mathbf{V}}{\partial p}\right) dp. \tag{A2}$$

Assuming that the measured wind in a sounding is well approximated by the geostrophic wind, the thermal wind is

$$\frac{\partial \mathbf{V}_g}{\partial p} = -\frac{R}{fp}\mathbf{k} \times \nabla_p T, \tag{A3}$$

where R is the specific constant of dry air, f the Coriolis parameter, p the pressure level, and T the temperature.

By using (A3), the expression (A2) can be written

$$H = -\int_{p_0}^{p} \frac{R}{fp}(-\mathbf{V}_g\cdot\nabla_p T)\, dp, \tag{A4}$$

which shows that ground-relative helicity is associated with the average geostrophic temperature advection in the layer bounded by p_0 and p.

The *storm-relative* helicity (SRH) is defined by

$$\mathrm{SRH} = -\int_{z_0}^{z} \mathbf{k}\cdot\left[(\mathbf{V} - \mathbf{c}) \times \frac{\partial \mathbf{V}}{\partial z}\right] dz, \tag{A5}$$

where **c** is the storm velocity. By using the same argument as above, the SRH is a measurement of the geostrophic temperature advection in a coordinate system moving with the storm.

Suppose that the actual wind is the geostrophic wind plus an ageostrophic component. Then in (A2),

$$\mathbf{k}\cdot\left(\mathbf{V} \times \frac{\partial \mathbf{V}}{\partial p}\right) = \mathbf{k}\cdot\left(\mathbf{V} \times \frac{\partial \mathbf{V}_g}{\partial p}\right) + \mathbf{k}\cdot\left(\mathbf{V} \times \frac{\partial \mathbf{V}_a}{\partial p}\right). \tag{A6}$$

It is easy to show that

$$\mathbf{k}\cdot\left(\mathbf{V} \times \frac{\partial \mathbf{V}_g}{\partial p}\right) = \frac{R}{fp}(-\mathbf{V}\cdot\nabla_p T), \tag{A7}$$

which is proportional to the temperature advection by the actual wind.

As the helicity is normally calculated at low levels where the momentum advection is weak (Carlson 1991), the isallobaric wind can be considered a good approximation of the low-level ageostrophic wind, and then

$$\mathbf{V}_a = -\frac{1}{f^2}\nabla_p\left(\frac{\partial \phi}{\partial t}\right), \tag{A8}$$

where ϕ is the geopotential ($=gz$). By using the hydrostatic equation, it is found that

$$\frac{\partial \mathbf{V}_a}{\partial p} = \frac{R}{f^2 p}\nabla_p\left(\frac{\partial T}{\partial t}\right), \tag{A9}$$

and then

$$\mathbf{k}\cdot\left(\mathbf{V}\times\frac{\partial \mathbf{V}_a}{\partial p}\right) = \frac{R}{f^2 p}\mathbf{k}\cdot\left[\mathbf{V}\times\nabla_p\left(\frac{\partial T}{\partial t}\right)\right]$$
$$= \frac{R}{f^2 p}\nabla_p\left(\frac{\partial T}{\partial t}\right)\cdot(\mathbf{k}\times\mathbf{V}), \quad \text{(A10)}$$

but $(\mathbf{k}\times\mathbf{V}) = -\mathbf{V}'$, where $\mathbf{V}'$ is the wind turned $\pi/2$ clockwise. Then

$$\mathbf{k}\cdot\left(\mathbf{V}\times\frac{\partial \mathbf{V}_a}{\partial p}\right) = \frac{R}{f^2 p}\left[-\mathbf{V}'\cdot\nabla_p\frac{\partial T}{\partial t}\right], \quad \text{(A11)}$$

and so the vertical variation of the isallobaric wind produces helicity when the rate of local temperature variation is greater to the left than to the right of the actual wind.

In (A7), estimating the wind as 10 m s^{-1} and the gradient of temperature as 4 K (100 km)$^{-1}$, the order of magnitude of (A7) is 10^{-2} J kg^{-1} Pa^{-1}, which for a layer between the surface and 700 hPa represents a helicity of about 300 m^2 s^{-2}. In order to obtain the same order of magnitude for the term represented by (A11), it is necessary that a differential heating rate between both sides of the actual wind of 10 K (6 h)$^{-1}$ in a distance of 100 km. Therefore, except in very special cases, the contribution to the helicity from the vertical variation of the isallobaric wind is much smaller than the contribution from the temperature advection.

APPENDIX B

K-Means Clustering Method

Consider the following problem: given a number of entities (N) described by a set of variables (p), separate the entities in a number of groups (k, named clusters) in such a manner that the entities in each group are similar to one another but different from the entities in the other groups. Any technique that allows objective solutions to this problem is known as cluster analysis.

The N entities can be considered as a set of N points in a p-dimensional space, and then the clustering technique can be viewed as a tool to look for point groupings (clusters) in the p-dimensional space. The first stage for applying a clustering technique is to establish some measurement of the relationships between the entities (points); the most frequently used is the Euclidean distance. Variables are usually standardized before calculating similarity or dissimilarity between entities to eliminate the scale effect and to reduce the variables to nondimensional ones. If the entity X_i is given by

$$X_i = (X_{i1}, X_{i2}, \ldots, X_{ip}) \qquad i = 1, \cdots, N,$$

the standardized variables will be

$$x_i = (x_{i1}, x_{i2}, \ldots, x_{ip}) \qquad i = 1, \cdots, N$$

with

$$x_{ij} = (X_{ij} - X_j)/s_j \qquad j = 1, \cdots, p,$$

where X_j is the mean and s_j the standard deviation of the jth variable. The Euclidean distance between two entities is then

$$d_{ij} = \left[\sum_{l=1}^{p}(x_{il} - x_{jl})^2\right]^{1/2},$$

where x_{il} is the value of the lth standardized variable for the ith entity.

Clustering techniques can be classified in two major groups: hierarchical and nonhierarchical methods. In hierarchical methods, the entities are not classified in groups in one step. It proceeds by successive mergers (agglomerative procedure) or successive divisions (division procedure) following some criteria of similitude or dissimilitude. For example, agglomerative methods start with as many clusters as entities and finish when all the entities are in the same cluster. The results are normally presented graphically as hierarchical trees or dendrograms, used to decide how many clusters are finally considered. Among hierarchical methods are found the *single linkage, complete linkage, centroid, median,* and *Ward's* methods.

In nonhierarchical methods, clusters are formed by the optimization of a clustering criterion. Typically, some initial partition of the entities is chosen and then cluster members are altered to obtain a new partition that reveals any natural structure. The number of clusters, k, is specified a priori. The method most widely used by which the initial partition is made is to estimate cluster centroids. These estimates are known as *seed points* and the clusters are formed around them. Between nonhierarchical methods are found the *k-means* and *nucleated agglomerative* methods.

The k-means procedure can be summarized in the following steps.

1) Specify k seed points as a set of centroids of k clusters (e.g., the first k entities). If some data structure is known a priori, this knowledge can be used to select the seed points.
2) Calculate the Euclidean distance between the next entity and the centroids. Assign the entity to the cluster having the nearest centroid.
3) Recompute the centroid of the cluster after the assignment.
4) Repeat steps 2 and 3 until all the entities have been located.
5) Calculate the distance between each entity and the centroids of the k clusters.
6) Reassign each entity to the nearest cluster (represented by its centroid).
7) Repeat steps 5 and 6 until no change in cluster members is observed.

A detailed explanation and discussion of the clustering methods can be found in Everitt (1980) and a more specific application to geophysical problems in Gong and Richman (1995). FORTRAN codes of different methods can be found in Anderberg (1973) and commercial programs such as STATISTICA or SPSS have algorithms that do cluster analysis.

REFERENCES

Andenberg, M. R., 1973: *Cluster Analysis for Applications.* Academic Press, 359 pp.

Barnes, S. L., and C. W. Newton, 1986: Thunderstorms in the synoptic setting. *Thunderstorm Morphology and Dynamics,* E. Kessler, Ed., University of Oklahoma Press, 75–112.

Brooks, H. E., C. A. Doswell III, and J. Cooper, 1994: On the environment of tornadic and nontornadic mesocyclones. *Wea. Forecasting,* **9,** 606–618.

Carlson, T. N., 1991: *Mid-Latitude Weather Systems.* Harper Collins Academic, 507 pp.

Darkow, G. L., 1969: An analysis of over sixty tornado proximity soundings. Preprints, *Sixth Conf. on Severe Local Storms,* Chicago, IL, Amer. Meteor. Soc., 107–111.

Davies-Jones, R., D. Burgess, and M. Foster, 1990: Test of helicity as a tornado forecast parameter. Preprints, *16th Conf. on Severe Local Storms,* Kananaskis Park, AB, Canada, Amer. Meteor. Soc., 588–592.

Doswell, C. A., III, 1982: The operational meteorology of convective weather, Vol. I, Operational mesoanalysis. NOAA Tech. Memo. NWS NSSFC-5, 158 pp. [Available from National Severe Weather Storms Forecast Center, Federal Building 601E, 12th Street, Kansas City, MO 64106-2877.]

——, 1987: The distinction between large-scale and mesoscale contribution to severe convection: A case study example. *Wea. Forecasting,* **2,** 3–16.

——, 1991: A review for forecasters on the application of hodographs to forecasting severe thunderstorms. *Natl. Wea. Dig.,* **16,** 2–16.

——, and E. N. Rasmussen, 1994: The effect of neglecting the virtual temperature correction on CAPE calculations. *Wea. Forecasting,* **9,** 619–623.

Everitt, B., 1980: *Cluster Analysis.* Halsted Press, 136 pp.

Font, I., 1983: *Climatology of Spain and Portugal* (in Spanish). Instituto Nacional de Meteorología, 296 pp.

Fuelberg, H. E., and D. G. Biggar, 1994: The preconvective environment of summer thunderstorms over the Florida Panhandle. *Wea. Forecasting,* **9,** 316–326.

Galway, J. G., 1956: The lifted index as a predictor of latent instability. *Bull. Amer. Meteor. Soc.,* **37,** 528–529.

García-Dana, F., R. Font, and A. Rivera, 1982: *Meteorological Situation During the Heavy Rain Event in the Eastern Zone of Spain During October-82* (in Spanish). Instituto Nacional de Meteorología, 80 pp.

Gayá, C., 1984: Climatology of the Balearic Islands: Meteors (in Spanish). Instituto Nacional de Meteorología Publ. A-71, 204 pp. [Available from Instituto Nacional de Meteorología, Apartado 285, 28071 Madrid, Spain.]

Gayá, M., and A. Soliño, 1993: Tornadoes and downburst in Menorca (in Spanish). *Rev. de Menorca,* **I,** 5–18.

George, J. J., 1960: *Weather Forecasting for Aeronautics.* Academic Press, 673 pp.

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

Hoskins, B. J., and M. A. Pedder, 1980: The diagnosis of middle latitude synoptic development. *Quart. J. Roy. Meteor. Soc.,* **106,** 707–719.

Jacovides, C. P., and T. Yonetani, 1990: An evaluation of stability indices for thunderstorm prediction in greater Cyprus. *Wea. Forecasting,* **5,** 559–569.

Lilly, D. K., 1986: The structure, energetics and propagation of rotating convective storms. Part II: Helicity and storm stabilization. *J. Atmos. Sci.,* **43,** 126–140.

Llasat, M. C., 1987: Heavy rain events in Catalonia: Genesis, evolution and mechanism (in Spanish). Ph.D. dissertation, University of Barcelona, 250 pp. [Available from Departament de Publicacions, Universidad de Barcelona, Avda. Diagonal 647, 08028 Barcelona, Spain.]

McCann, D. W., 1994: WINDEX—A new index for forecasting microburst potential. *Wea. Forecasting,* **9,** 532–541.

Meteorological Office, 1954: *Weather in the Mediterranean.* Vol. 1, U.K. Air Ministry, 362 pp.

Miller, J. A., 1972: Notes on analysis and severe storms forecasting procedures of the Air Force global weather central. AWS–USAF Tech. Rep. 200, Headquarters Air Weather Service, Scott Air Force Base, IL, 102 pp.

Miró-Granada, J., 1969: An exceptional hail event in Mallorca (in Spanish). *Rev. Soc. Hist. Natural de Baleares,* **XV,** 20–56.

Pickup, N. M., 1982: Consideration of the effect of 500-hPa cyclonity on the success of some thunderstorm forecasting techniques. *Meteor. Mag.,* **111,** 87–97.

Ramis, C., 1976: *Contribution to the Upper Air Climatology over Mallorca* (in Spanish). University of Barcelona Press, 60 pp.

——, A. Jansà, S. Alonso, and M. A. Heredia, 1986: Convection over the western Mediterranean. Synoptic study and remote observation (in Spanish). *Rev. Meteor.,* **N 7,** 59–82.

——, M. C. Llasat, A. Genovés, and A. Jansà, 1994: The October 1987 floods in Catalonia: Synoptic and mesoscale mechanism. *Meteor. Appl.,* **1,** 337–350.

Sanders, F., 1986: Temperatures of air parcels lifted from the surface: Background, application, and monograms. *Wea. Forecasting,* **1,** 190–205.

Showalter, A. K., 1953: A stability index for thunderstorm forecasting. *Bull. Amer. Meteor. Soc.,* **34,** 250–252.

Sumner, G., C. Ramis, and J. A. Guijarro, 1993: The spatial organization of daily rainfall over Mallorca, Spain. *Int. J. Climatol.,* **13,** 89–109.

——, J. A. Guijarro, and C. Ramis, 1995: The impact of surface circulation on significant daily rainfall patterns over Mallorca. *Int. J. Climatol.,* **15,** 673–696.

Weisman, M. L., and J. B. Klemp, 1986: Characteristics of isolated convective storms. *Mesoscale Meteorology and Forecasting,* P. S. Ray, Ed., Amer. Meteor. Soc., 331–358.

Yarnal, B., 1993: *Synoptic Climatology in Environmental Analysis.* Belhaven Press, 195 pp.