ELSEVIER

Atmospheric Research 86 (2007) 173 – 193

ATMOSPHERIC RESEARCH

www.elsevier.com/locate/atmos

Improvement of the analog forecasting method by using local thermodynamic data. Application to autumn precipitation in Catalonia

J. Gibergans-Báguena [a,*], M.C. Llasat [b]

[a] Department of Applied Mathematics III, Polytechnic University of Catalonia, Barcelona, Spain
[b] Department of Astronomy and Meteorology, University of Barcelona, Spain

Received 14 June 2006; received in revised form 13 April 2007; accepted 13 April 2007

Abstract

The objective of this paper is to present the improvement of quantitative forecasting of daily rainfall in Catalonia (NE Spain) from an analogues technique, taking into account synoptic and local data. This method is based on an analogues sorting technique: meteorological situations similar to the current one, in terms of 700 and 1000 hPa geopotential fields at 00 UTC, complemented with the inclusion of some thermodynamic parameters extracted from an historical data file. Thermodynamic analysis acts as a highly discriminating feature for situations in which the synoptic situation fails to explain either atmospheric phenomena or rainfall distribution. This is the case in heavy rainfall situations, where the existence of instability and high water vapor content is essential. With the objective of including these vertical thermodynamic features, information provided by the Palma de Mallorca radiosounding (Spain) has been used. Previously, a selection of the most discriminating thermodynamic parameters for the daily rainfall was made, and then the analogues technique applied to them. Finally, three analog forecasting methods were applied for the quantitative daily rainfall forecasting in Catalonia. The first one is based on analogies from geopotential fields to synoptic scale; the second one is exclusively based on the search of similarity from local thermodynamic information and the third method combines the other two methods. The results show that this last method provides a substantial improvement of quantitative rainfall estimation.

Keywords: Quantitative forecasting; Analog; Daily rainfall; Radiosounding; Catalonia; Spain

1. Introduction

Forecasting precipitation over a period of two or three days is a response to very important needs of a socio-economic kind. In particular, forecasting heavy rainfall that could potentially lead to catastrophes for people and material goods attracts special attention: a good estimate of the precise time they will occur, how intense they will be and their location is needed. However, precipitation is more difficult to forecast than other variables such as temperature or wind, mainly because of three factors:

1) The hydrodynamic and thermodynamic mechanisms that govern this variable bring into play additional microphysical phenomena that are less well known, and therefore they are more difficult to model than temperature and wind (Romero et al., 1997).

* Corresponding author.
E-mail addresses: jose.gibergans@upc.edu, jose.gibergans@upc.es (J. Gibergans-Báguena), carmell@am.ub.es (M.C. Llasat).

0169-8095/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.atmosres.2007.04.002

2) The scale of some mechanisms involved in the precipitation processes is smaller than the grids commonly used by numerical forecasting models. Furthermore, it is necessary to parameterize the usual grids to represent these kinds of phenomena, and especially the extreme cases, as happens with convection.
3) Finally, precipitation in the Mediterranean often is closely linked to orography through the upslope advection of moisture (Doswell et al., 1996; Romero et al., 1997). However, besides this upslope advection, orography can also help to maintain the stationarity of the convective systems, trigger the potential instability or create some mesoscale structures that modifies the pressure field in surface (Rigo and Llasat, 2007; Mariani et al., 2005; Riosalido et al., 1997; Ramis et al., 1994).

Nowadays, the best numerical weather prediction models are increasing their skill of quantitative precipitation forecasts (Roebber et al., 2004), but, for rainfall forecasting from one to three days, and taking into account the limitations exposed previously, statistical techniques can play a major supporting role. Indeed, recent analyses made using Limited Area Models have shown completely different results when they have been applied to the same heavy rainfall event (Mariani et al., 2005). The advantage of these statistical methods is that they give information about the probability distribution, which is especially important to civil protection and water management agencies. Recently, some operational weather services in Europe have implemented the analog forecasting method in operational tools. This is the case of the INM, National Institute of Meteorology in Spain (as it can be consulted on its website http://www.inm.es/web/infmet/predi/preci.html), and Regione ARPA Emilia-Romagna in Italy (Obled et al., 2002; Diomede et al., 2006). This technique consists of:

1) Condensing the available information in large data sets (geopotential fields or other variables at various pressure levels) using principal components analysis (PCA),
2) Selecting analogous situations from an historical file using a similarity criterion,
3) Drawing up a probabilistic rainfall forecast on the basis of the precipitations observed during such analogous situations.

This technique replaces the subjective procedure carried out by a person who makes a precipitation forecast using the geopotential and surface maps and then relates them to the rainfall distribution in the region based on his experience. This subjective procedure was already applied before the 50's and the analog forecasting method from an objective analysis was drawn up during the 1970s by Duband (1970, 1974, 1980, 1981). Afterwards, it has been applied by the operational forecasting section of the Direction Technique Générale d'Électricité de France in conjunction with the results offered by Météo-France. Efforts have been made since then — with greater or less success — to develop the method (Radinovic, 1975; Gutzler and Shukla, 1984; Toth, 1989; Guilbaud, 1997; Obled et al., 2002; Bontron, 2004).

The objective of this paper is to show the improvement on the traditional analogues method by using local thermodynamic information in order to do the quantitative precipitation forecasting (QPF). Thermodynamic analysis of the atmosphere allows obtaining information such as the possible existence of instability or auspicious environment for convection, the water vapor content as well as the possible formation and development of clouds, and the existence of wind shear and jet streak. Besides mesoscale analysis, thermodynamic analysis is a useful discriminating feature in situations in which the synoptic situation does not allow a complete explanation of either the atmospheric phenomena or their distribution. In the case of heavy rainfalls, where the existence of conditional instability and high water vapor content are essential preconditions (Doswell et al., 1996; Schultz et al., 2000), the thermodynamic analysis can provide very useful information. Much work has been carried out on this topic as it relates to the Mediterranean area (Ramis et al., 1994, 1995; Llasat et al., 1996; Barret and Cheng, 1996; Boni et al., 1996; Doswell et al., 1998; Jansà et al., 2001a,b; Alpert et al., 2002). The methodology developed here has been applied over Catalonia (North East of the Iberian Peninsula, Fig. 1). This region is characterized by a high frequency of heavy rains, which sometimes give rise to catastrophic flash-floods (Llasat et al., 2003) or extended flooding that can even affect the South of France or the North of Italy (Llasat et al., 1999). The similar heavy rainfall conditions present over the just mentioned areas (Gulf of Genova surroundings) suggest an extension of the methodology shown in this paper to that region (Llasat, 2004).

This paper begins in Section 1 with a brief introduction of the situation and matter of interest. A detailed description of the data used is provided in Section 2. Data quality control and selection of the most suitable thermodynamic descriptors are discussed and shown in Section 2, which also includes the thermodynamic analysis. Section 3 contains the presentation of the analogues methods used, and particularly of method improved with information from the Palma de Mallorca radiosounding. A methodology for the forecast verification is given in

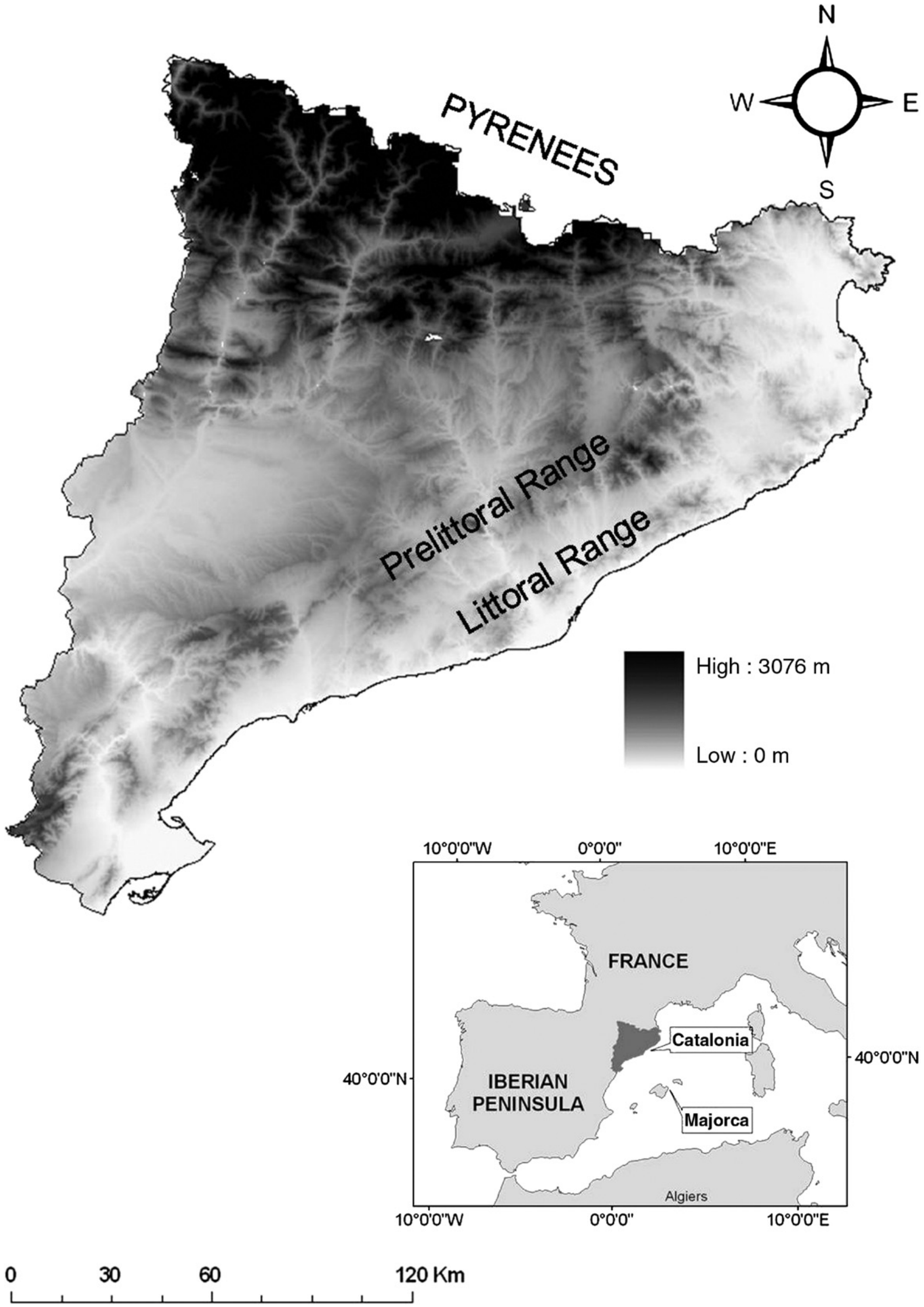

Fig. 1. (a) Location of Catalonia. (b) Enlarged map showing major orographical features and the location of several places mentioned in the text.

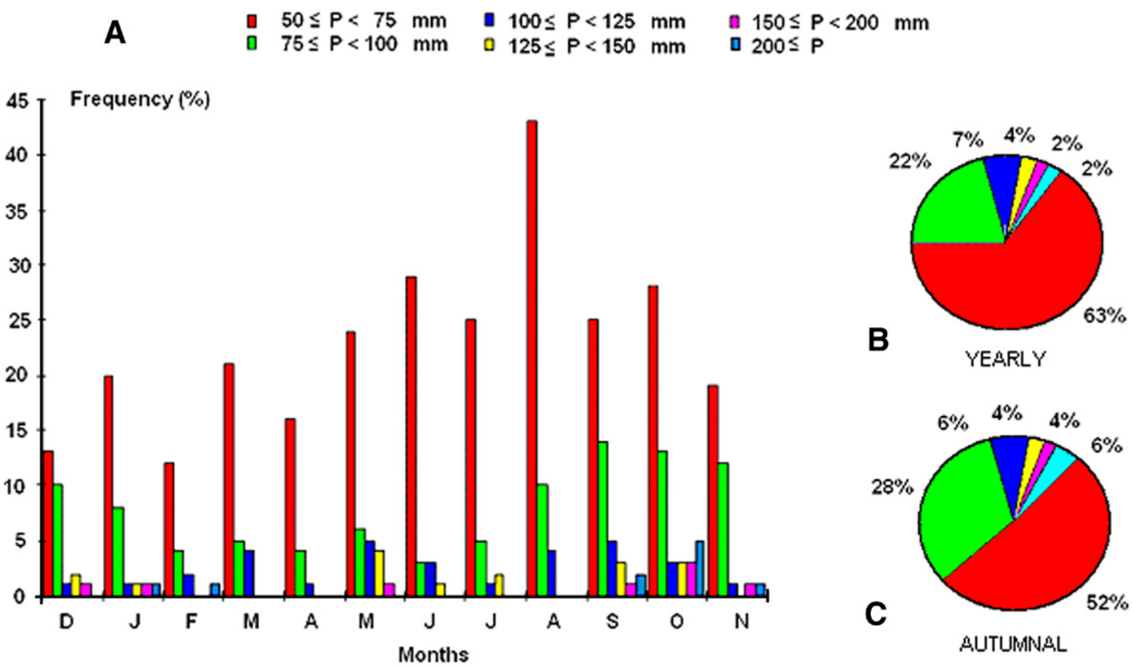

Fig. 2. A) Distribution of daily rainfall superior to 50 mm/24-h in Catalonia (1975–1990) over the course of the year. The sector diagrams show the yearly (B) and autumnal (C) distribution of these precipitations.

Section 4. Below, Section 5 provides a discussion of the results obtained for the fall season (the months of September, October and November). Finally, the paper finishes with the most important conclusions and some proposals of future work.

2. Description and analysis of the database

2.1. Daily pluviometry. Pluviometric regionalization

Catalonia is located on the NE corner of the Iberian Peninsula (Fig. 1). Its coastline runs along the NE–SW direction and its main orographical features are a coastal mountain range with some peaks exceeding 1000 m and the Pyrenees, which lie roughly from W to E along the northern border of Spain. The main rivers are the Ebro and its tributaries, but most floods affect the minor ones, such as Llobregat, Besós and Fluvià rivers, which flow roughly from the coastal mountain range to the Mediterranean sea.

As a rule, minor floods are recorded almost every year somewhere at the coastal plains. These are caused by short-lived showers with moderate (over 1 mm/min) or high (above 3 mm/min) rainfall rates, which occur mostly in late summer or in early fall. However, catastrophic events causing serious damage to property and even casualties are recorded from time to time (Llasat and Puigcerver, 1994). Under favourable conditions, the above-mentioned topographical features can act as triggering and bolstering atmospheric convection factors (Romero et al., 1997). Autumn and spring are the seasons of the year with the maximum accumulated precipitation (Fig. 2). Fall (September–October–November) is characterized as being the season of the year with the highest percentage of precipitation exceeding 75 mm in 24 h (Gibergans-Báguena, 1993). In this season the return periods of this precipitation threshold are less than 10 years in some zones of Catalonia. Taking into account the different features of rainfall over Catalonia along the year, the regionalization can vary from one season to another (Lana et al., 1995). In this paper, special attention has been paid to the fall season.

Table 1
Variance and total accumulated variance (%) explained by the first 7 principal components from the daily rainfall for the fall

Principal component	Eigenvalue	Variance (%)	Accumulate variance (%)
1	18.28898	40.6	40.6
2	4.42664	9.8	50.5
3	2.75503	6.1	56.6
4	1.92402	4.3	60.9
5	1.74075	3.9	64.7
6	1.31455	2.9	67.7
7	1.14201	2.5	70.2

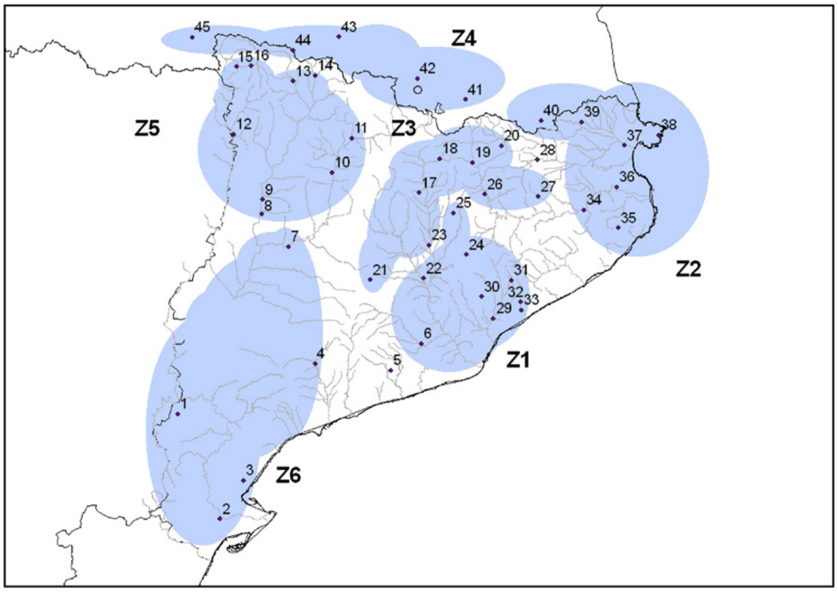

Fig. 3. Fall pluviometric regionalization for Catalonia. Table 6 contains the names of this pluviometric stations.

The initial information at our disposal consisted of daily precipitation series from 45 pluviometric stations in Catalonia and the South of France for the period 1970–1983. A previous quality control focusing on the spatial correlation was applied. After that control, gaps were less than 5% and they were filled in by means of a correlation method taking into account the closest and best correlated stations ($r > 0.6$).

Given that the daily precipitation in the study region varies widely from one station to another (Lana et al., 1995), we must work with mean values representative of sufficiently homogeneous zones. A first step, therefore, is to undertake a pluviometric regionalization of Catalonia and a grouping of the stations into homogeneous zones in relation to the available period. Then, we created the pluviometric groups applying the following methodology from Bonell and Summer (1992), which consists in:

1) Principal Component Analysis (PC Analysis) of the daily precipitation data, in order to remove possible correlations between variables;
2) Cluster analysis on the basis of observations in the space of the principal components.

Table 1 shows the variation explained by the first 7 principal components (criterion of Kaiser, 1958). The first two components explain 50% of the variation, while all 7 together explain over 70%, indicating that it was possible to reduce the 45 variables to only 7. Fig. 3 shows the results obtained after the cluster analysis, choosing six pluviometric regions that include the Pyrenean stations on the Spanish and French sides. Table 2 shows the main characteristics of every group. The Zone 4 is observed to be the one that records the largest accumulated precipitation in the fall (247.3 mm); however, the zone in which the highest precipitation values in 24 h are reached is Zone 2 (north-east coastal plains).

2.2. Geopotential fields at 700 hPa, 1000 hPa and 700/1000 hPa thickness

Geopotential data were available at 00 UTC from 37 radiosonde stations in central and Western Europe over the period 1953–1990 (Fig. 4 and Table 3). These data were provided by Eléctricité de France (EDF). Because this paper investigates the improvement of the analog forecasting method usually applied by previous authors (Duband, 1981; Guilbaud, 1997), the geopotential fields selected were the same as those used by these authors:

Table 2
Pluviometric characteristics for the studied zones in the fall

	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6
Fall mean (mm)	185.0	213.5	195.5	247.3	188.4	173.8
Fall percentage rainy days (%)	24.9	24.0	23.6	34.7	31.0	19.2
Fall daily mean (mm)	8.2	9.8	9.3	7.8	6.7	9.9
Fall maximum daily mean (mm)	81.3	111.2	116.1	120.8	108.2	62.4
Fall daily absolute maximum (mm)	185.9	430.0	271.8	240.0	169.5	187.2

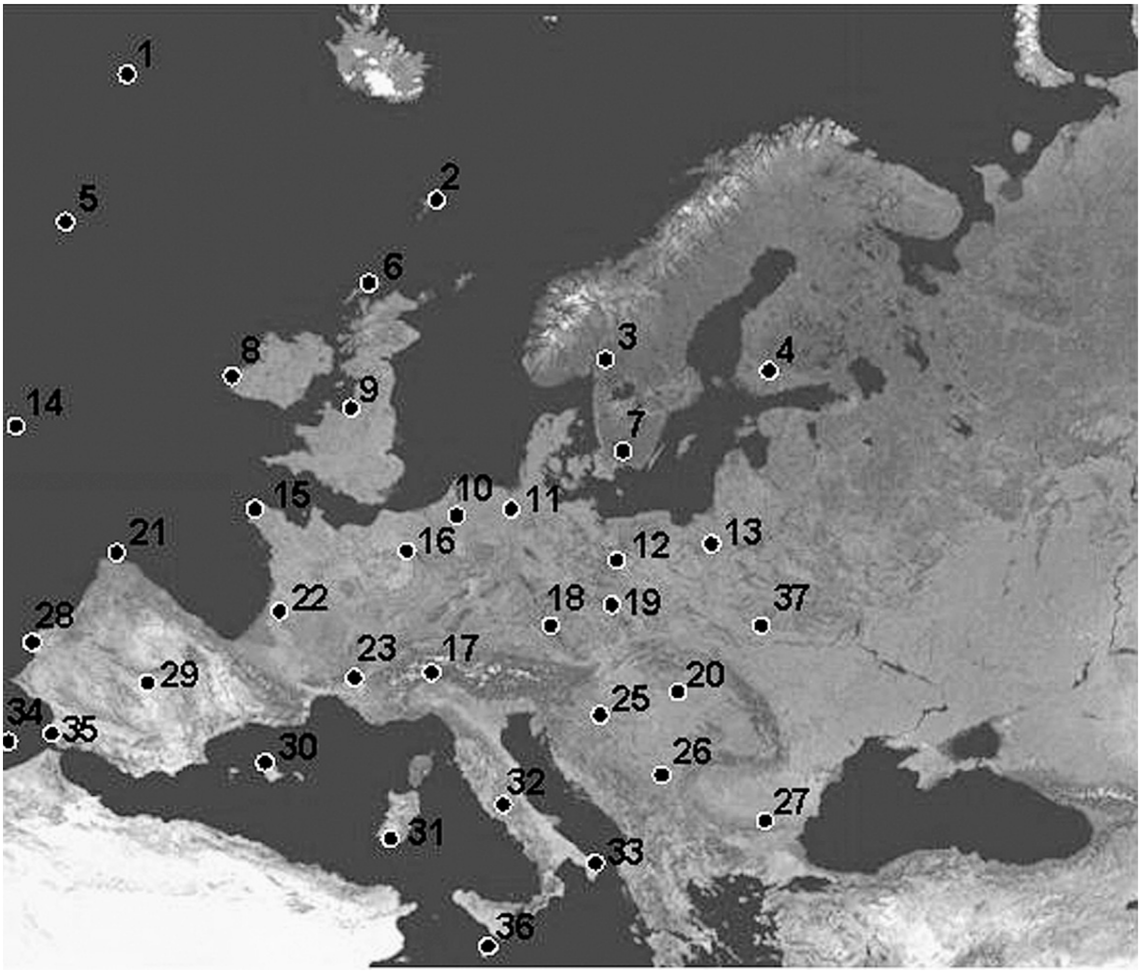

Fig. 4. Radiosonde station network. Table 3 contains the names of this radiosonde stations.

700 and 1000 hPa. These fields were selected owing to the great importance of the role that lower and middle troposphere plays in rainfall processes in the zone of study, and particularly in the zones of highest precipitation intensity (Duband, 1981; Llasat et al., 1997; Tudurí and Ramis, 1997).

The information was condensed by carrying out a Principal Component Analysis, saving in a file the first 6 principal components (PCs) of the geopotential at 700 hPa, the first 6 PCs of the geopotential at 1000 hPa and the first 6 PCs of thickness 700/1000 hPa, which will be indicated as Z^{700}, Z^{1000} and E, respectively (Duband, 1980, 1981).

2.3. Palma de Mallorca thermodynamic data

Although the city of Barcelona now has a radiosonde station, this station was not used because sufficiently lengthy data series were not yet available for it. Nevertheless previous studies have shown the suitability of the Palma de Mallorca radiosounding for studying and monitoring the Mediterranean meteorological systems that usually affect Catalonia (Llasat, 1987; Llasat and Puigcerver, 1992; Ramis et al., 1995). Local thermodynamic data have been shown to be useful for the diagnosis of heavy rainfalls and severe weather in Catalonia and surrounding region (Llasat, 1987; Tudurí and Ramis, 1997; Doswell et al., 1998). The information used as a point of departure was that made up of the notable points from the Palma de Mallorca radiosoundings corresponding to 00 and 12 UTC for the period 1975–1989. A quality control was carried out on the radiosoundings, selecting those that showed they fulfilled certain requirements regarding both the quality and the quantity of information provided (Llasat et al., 1997).

Table 3
Radiosonde stations network

1	Point I
2	Torshaw'n
3	Oslo
4	Jokioinen
5	Point J
6	Stornoway
7	Kestrup
8	Valentia
9	Aughton
10	De Bilt
11	Emden
12	Lindenberg
13	Logionowo
14	Point K
15	Brest
16	Trappes
17	Payerne
18	Munich
19	Praga
20	Budapest
21	La Coruña
22	Bordeaux
23	Nîmes
24	Milan
25	Zagreb
26	Belgrado
27	Bucarest
28	Lisboa
29	Madrid
30	Palma
31	Cagliari
32	Roma
33	Brindisi
34	Funchall
35	Gibraltar
36	Malta
37	Poprad Tatry

Numbers correspond to its location in the map of Fig. 4.

The quality control of radiosoundings has three steps:

1. *Mistaken data: the different radiosounding were discretized taking 10 hPa intervals*. We calculated the mean (μ) and standard deviation (σ) of temperature for each interval. The distributions obtained in each one of the intervals were, however, different and generally far-removed from a normal distribution. We have accepted as valid those observations situated in the interval defined by $\mu \pm 5\sigma$. This interval contains at least 96% of the observations of this variable in this level according to Chebyshev inequality (Mood et al., 1974). Only the values that fail this criterion are eliminated, but not the whole radiosounding.
2. *Regarding the number of observations and the minimal pressure* reached by the radiosounding (temperature and humidity data), if the last level is below 500 hPa or the number of values fewer than 4, the radiosounding is eliminated.
3. *Regarding the number of observations and minimum pressure level for which wind data are available,* in this case, those radiosoundings that miss that kind of data are not eliminated. However they have been identified as “not good radiosounding”.

Application of this quality control meant that from 3991 observations at 00 UTC and 3947 at 12 UTC, 3450 and 3451 at 00 and 12 UTC respectively were finally used due to their being deemed reliable (Fig. 5 shows this process). In the end, all the thermodynamic parameters and variables from 2393 and 2265 radiosoundings at 00 and 12 UTC respectively were taken into account.

In order to have detailed information about the troposphere in the area of study, 33 thermodynamic parameters were calculated. The first 30 are based exclusively on temperature and humidity data, while the last three include wind. The calculated parameters are:

- Relative temperature (quotient between the temperature and the average temperature in this level): surface, 850 hPa, 700 hPa and 500 hPa.
- Potential temperature: surface, 850 hPa, 700 hPa and 500 hPa.
- Potential equivalent temperature: surface, 850 hPa, 700 hPa and 500 hPa.
- Equivalent potential temperature difference: surface — 950 hPa, 950–850 hPa, 850–700 hPa and 700–500 hPa, using Bolton's (1980) expression.
- Precipitable water mass (PWM) between: surface — 850 hPa, 850–700 hPa, 700–500 hPa. In order to distinguish between its value in the low troposphere and medium troposphere it is also calculated between 850 and 500 hPa and, finally, in order to know the total water content between surface and 500 hPa.
- Isozero pressure: level at which the temperature is 0 °C.
- Convective Available Potential Energy, CAPE, following Weisman and Klemp's (1986) definition.
- Positive energy or energy released by the particle during its ascent, $E>0$.
- Maximum vertical velocity, VMAX. Even though directly related to CAPE, it is useful to know its magnitude.
- Showalther index (Showalter, 1953), SI.
- Lifted index (Galway, 1956; Fujita et al., 1970), LI.
- K index (George, 1960), KI.
- Total totals index (Miller et al., 1972), TTI.

- Wet bulb potential temperature index (Pickup, 1982), PI.
- Severe Weather Advisory Trend (Miller et al., 1971), SWEAT.
- Vertical shear, SHR, defined as the difference between the wind averaged with the density of the 600 m below and the wind of the 500 m below (Weisman and Klemp, 1986).
- Bulk Richardson number (Weisman and Klemp, 1986), BRI, which provides a relationship between CAPE and vertical shear.

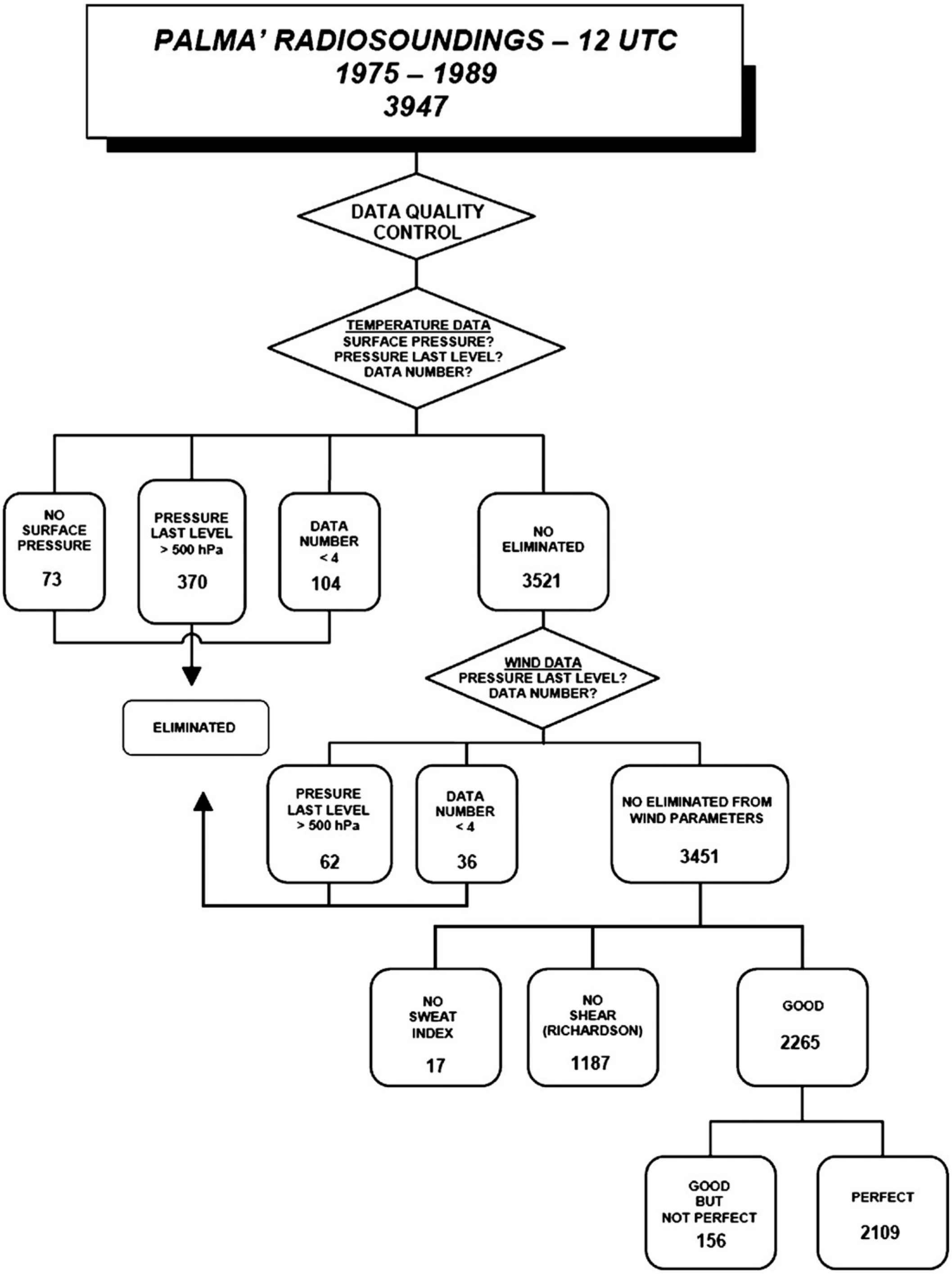

Fig. 5. Quality control schema and results for Palma de Mallorca radiosounding data at 12 UTC for the period 1975–1989.

With the objective to compress as much information as possible, a Discriminant Analysis with these 33 variables was carried out (Timn, 2002; Wackernagel, 2003). The results of this analysis allowed the selection of 7 variables (further comments are included in Section 5 of this paper).

3. Methodology followed for rainfall forecasting

Two states of the atmosphere are deemed "analogues" when there is a resemblance between them (Lorenz, 1969). It would therefore be reasonable to expect that the meteorological phenomena associated with such states would also bear a resemblance to each other. Clearly, however, a large collection of analogues situations can have very different phenomena associated with them, although each of these phenomena will have a certain probability of occurring, so we may speak of a function of distribution of probability, of mean, quantiles, etc.

The principle of the various methods used in this paper can be summed up in two steps:

1) "Selection" or "search" of analogous days to a particular day (test day) according to selection criteria.
2) "Estimation" of mean precipitation for the test day for each study zones, on the basis of a cumulative distribution function of the rainfall obtained from the analogous days, calculated by interpolation the quantiles of 20, 60 and 90%.

For the period 1975–1983 (inclusive), the radiosounding data, rainfall data and pressure fields were all available. In order to validate the results, a 70% of the observations have been selected by a random process and the remaining 30% has been used to validate the methodology.

The various methods drawn up and tested, which are described below, differ from each other essentially in the different proposed ways of searching the analogous days.

3.1. Method A

In this method, selection of the analogues is carried out by means of two selection criteria. This is the classical method (Duband, 1981; Guilbaud, 1997). The first one, referred to as "criterion of proximity", lies in the space of the first 6 Principal Components (PC) of the geopotential field at 700 hPa at 00 UTC; whereas the second one, referred to as a "criterion of correlation", lies in the space of the 6 PC of the geopotential at 700 hPa, 6 PC of the geopotential at 1000 hPa and 1 PC from the 700/1000 hPa thickness, all of them at 00 UTC. The methodology is the following:

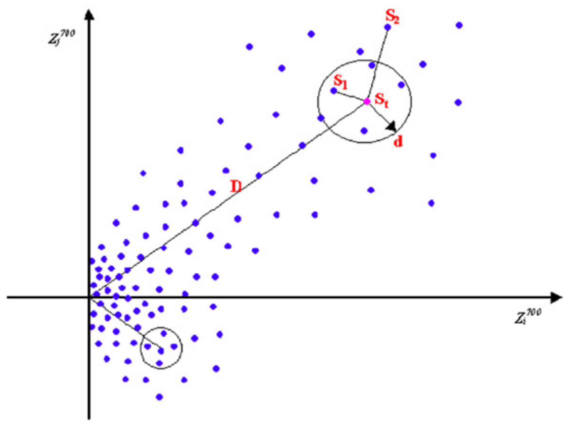

Fig. 6. Proximity criterion: a day (S_1) is considered as analogous if it is within *N*-ball centred in test day (S_t).

3.1.1. "Criterion of proximity"

Situations in the space of the *N* principal components that lie in the interior of an *N*-ball of radius *d* (a function of the distance from the origin to the test situation, *t*) are considered to be analogous if (Fig. 6):

$$\sum_{j=1}^{N} (Z_{ij} - Z_{tj})^2 < d^2 \quad (1)$$

where Z_{ij} is the value of the variable Z_j on the *i*-th day, and Z_{tj} is the value of that variable on the test day under study.

For situations far from the origin, the number of analogs decreases. In order to be able to have a similar number of analogous situations for all the searched days, the radius of the *N*-ball has to be increased.

In order to calculate the distance "*d*" we proceeded as follows. Let *D* be the distance to the origin of the test day, so that:

$$D^2 = \sum_{i=1}^{N} (Z_i)^2. \quad (2)$$

The Probability Density Function (PDF) of the distances D^2 follows a χ^2 law with *N* degrees of freedom, since Z_1, Z_2,..., Z_N are assumed to be independent and identically distributed normal random variables (Muirhead, 1982).

For a law of χ^2 with N degrees of freedom, probability density function is given by the following expression:

$$f(x) = \frac{x^{(N/2)-1} e^{-x/2}}{2^{N/2}\Gamma\left(\frac{N}{2}\right)}, \text{ with } \Gamma(N) = (N-1)! \tag{3}$$

For certain values of N, this PDF corresponds to a Gamma function. Since, in our case the number of principal components at 700 hPa is N=6, this function is expressed as:

$$f(x) = \frac{x^2 e^{-x/2}}{16} \tag{4}$$

and PDF corresponds to a Gamma function of parameters λ=3 and ρ=2.

All the quantiles of this PDF have a typical deviation σ_D proportional to the size of the sample, m, that is, to the number of analogous situations for each test day.

For example, if we take m=25, then:

$$\sigma_D = \frac{1}{f(x)}\sqrt{\frac{p(1-p)}{m}} = \frac{1}{f(x)}\sqrt{\frac{p(1-p)}{25}} \tag{5}$$

in which p is the probability calculated by interpolation based on a table of the Gamma function with 6 degrees of freedom.

For a number of variables between 6 and 8, Duband (1981) proposes the following expression:

$$d^2 = \omega^2\sigma_D^2 \Rightarrow d = \omega\sqrt{\frac{1}{f(x)}\sqrt{\frac{p(1-p)}{25}}} \tag{6}$$

in which ω adjusts the radius of the N-ball to have a number of analogous situations between 10 and 50.

3.1.2. "Criterion of correlation"

This criterion requires the days selected to have a high coefficient of correlation with the test day for the 13 PC as a whole: 6 PC of the geopotential at 700 hPa, 6 PC of the geopotential at 1000 hPa and 1 PC from the 700/1000 hPa thickness. The coefficient of linear correlation, r, is calculated between the vector

$$(Z_{i1}^{700}, Z_{i2}^{700}, \ldots, Z_{i6}^{700}, Z_{i1}^{1000}, Z_{i2}^{1000}, \ldots, Z_{i6}^{1000}, E_{i1}) \quad \text{of day } i$$

and the vector

$$(Z_{t1}^{700}, Z_{t2}^{700}, \ldots, Z_{t6}^{700}, Z_{t1}^{1000}, Z_{t2}^{1000}, \ldots, Z_{t6}^{1000}, E_{t1}) \text{ of the test day, } t$$

and we retain as analogous the day i for which the following requirements are satisfied (Duband, 1981):

$$u^2 = \frac{d^2}{r^2} < 6 \text{ and } r^2 > 0.25. \tag{7}$$

Finally, the days chosen as analogous to the test day are in agreement with it regarding to general circulation of the atmosphere.

3.2. Method B

Method B is the same as Method A, except that it uses exclusively the information provided by the radiosounding that, as explained in section above, is representative of the rainfall situations of the zone of study. When working with Method B, we only take into account the "criterion of proximity" applied to the selected thermodynamic predictors, calculated on the basis of the Palma de Mallorca radiosounding at 12 UTC, in the selection of the analogous days.

In this case, the radiosounding selected is that of Palma de Mallorca. In Method B, in the selection of the analogous days, account was taken only of the "criterion of proximity" applied to the selected thermodynamic predictors calculated on the basis of the Palma de Mallorca radiosounding at 12 UTC.

3.3. Method C

This third method considers into account both the information at the synoptic scale, provided by the principal components of the geopotential fields at 700, 1000 and 700/1000 hPa at 00 UTC, and the information at a local scale, provided by the thermodynamic variables calculated on the basis of the radiosoundings at 12 UTC in Palma de Mallorca.

In this case, the search of analogous days is carried out by using the following stages:

1) Firstly, for each test day, the best 50 analogues are taken by applying the selection criteria of Method A.
2) Secondly, the 25 closest days are taken from the analogous selected in the previous step, using the selection criterion of Method B.

4. Methodology followed to evaluated the forecast

The objective of forecast verification is to determine its quality. In order to carry out this verification it is necessary to relate the values obtained by a specific forecasting

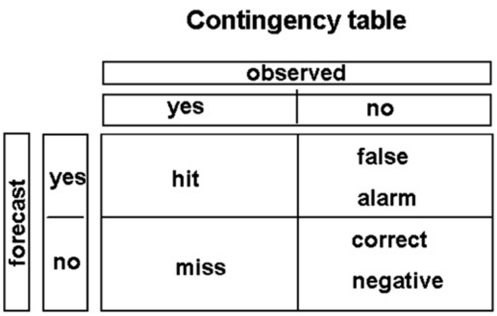

Fig. 7. Representation of the contingency table used to calculate hit rate(HR) and false-alarm rate (FAR).

method with those actually observed. By means of a suitable statistical method, indices are determined, which help to estimate the quality of the forecast so as to decide how useful it might be for users (Dobryshman, 1972; Daan, 1984, 1985; Murphy, 1985, 1987, 1990; Stanski, 1989). We present below the indices used.

4.1. Forecast of yes or no occurrence of a phenomenon

The simplest forecast is a categorical *yes* or *no* forecast of occurrence of a phenomenon such as *rain*, *hail*, etc. Given N_a analogues corresponding to day D, of which in N_p cases an occurrence of the phenomenon does arise and where P_c is the climatological percentage of days on which the phenomenon does occur, we have:

- If $N_p \leq N_a P_c \Rightarrow$ Forecast: Non-occurrence of the phenomenon
- If $N_p > N_a P_c \Rightarrow$ Forecast: Occurrence of the phenomenon

In order to assess the level of agreement between forecastings and observations, different measures of accuracy have been calculated as follows: once it has been obtained a forecasting answer (occurrence or non-occurrence of the phenomenon), we show the frequency of "yes" and "no" forecasts and observations in a contingency table (Fig. 7). A perfect forecast system would produce only hits and correct negatives, and no misses or false alarms.

Forecast accuracy can be measured from the contingency table by the hit rate (HR) and the false-alarm rate (FAR) as given by:

$$\mathrm{HR} = \frac{\text{hits}}{\text{hits} + \text{misses}}$$

$$\mathrm{FAR} = \frac{\text{false alarm}}{\text{false alarm} + \text{ correct negativeness}}.$$

Despite having the advantage of being very simple, deterministic and dichotomous type forecasts are of very limited practical utility, for often what is of interest is to gain a quantitative idea of the rainfall and its associated probability and not simply the categorical prediction of rain occurrence or non-occurrence. That is the reason why a probabilistic forecasting method was developed in different intervals of precipitation. Once the probabilistic forecasts are obtained for each category of precipitation, different thresholds are fixed in order to provide the dichotomous rain or non-rain prediction. At the moment we will focus on the exceedance of 0.1 mm in 24 h, since this is the threshold suggested by the National Institute of Meteorology in Spain (INM) to distinguish the rain occurrence or non-occurrence (more details about precipitation categories can be found in Section 4.2 in the current paper).

The verification method carried out on the predictions mentioned in paragraph above is relative operating characteristic (ROC) curves (Swets, 1973; Mason, 1982; Mason and Graham, 1999). Actually, 6 sets of ROC curves have been plotted, one for each pluviometric zone. ROC curves provide, in general terms, a comparison between hit rate (HR) and false-alarm rate (FAR), which respectively indicate the proportion of events for which a warning was provided correctly, and the proportion of non-events for which a warning was provided incorrectly. The hit rate is sometimes known as the Probability of Detection (POD), or prefigurance (Doswell et al., 1990; Wilks, 1995), and provides an estimate of the probability that an event will be forewarned. The false-alarm rate estimates the probability that for a non-event a warning will be provided incorrectly. The same way in the case of the deterministic forecast, a 2 × 2 contingency table can show the frequency of "yes" and "no" forecasts and observations, but in the particular case of probabilistic forecasts, the contingency table should have as many rows as there are forecast probability categories, not simply two (as in the case of dichotomous forecasts). Specifically, the thresholds for converting our quantitative probabilistic forecast to a dichotomous one will be 10%, 20%, 30%… up to 90%, so that we get 9 pairs of (FAR, HR). All the pairs (FAR, HR) obtained are plotted as an ensemble points on a diagram to form the ROC curve. One of the advantages of using ROC curves is that we can plot in the same graph the curves obtained from different forecasting systems (as methods A, B and C we are working with) and also make a comparison of the results obtained for each system. The global measure of forecast accuracy can be defined as the area under the ROC curve (AUC), so this value will be included for every forecast method in the ROC plots in Fig. 8. In general, for skilful forecast systems, the ROC curve

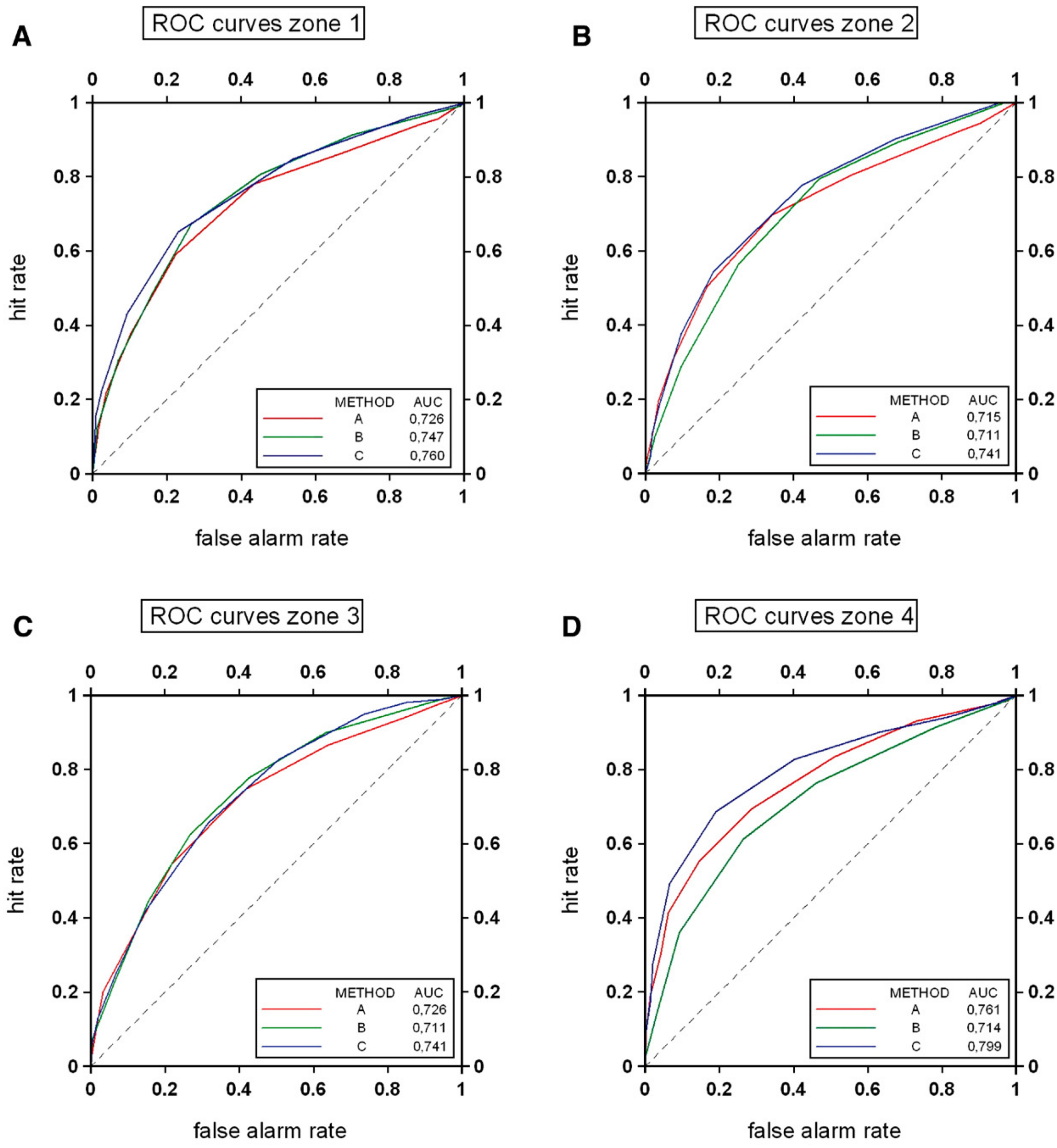

Fig. 8. ROC curves corresponding to the forecast rain/no rain for the six pluviometric zones of Catalonia by Analog Forecasting Methods (A, B and C), for the period 1975–1989. (A) ROC curve for Zone 1, (B) ROC curve for Zone 2, (C) ROC curve for Zone 3, (D) ROC curve for Zone 4, (E) ROC curve for Zone 5 and (F) ROC curve for Zone 6.

bends toward the top left, where hit rates are larger than false-alarm rates, and the total area under the curve is then greater than 0.5. When the curve lies close to the diagonal, the forecast system does not provide any useful information, and the area beneath the curve is approximately 0.5. If the curve lies below the line, negative skill is indicated. The reader is referred to Atger (2001) and Pellerin et al. (2003) for detailed explanations on computation and interpretation of the ROC.

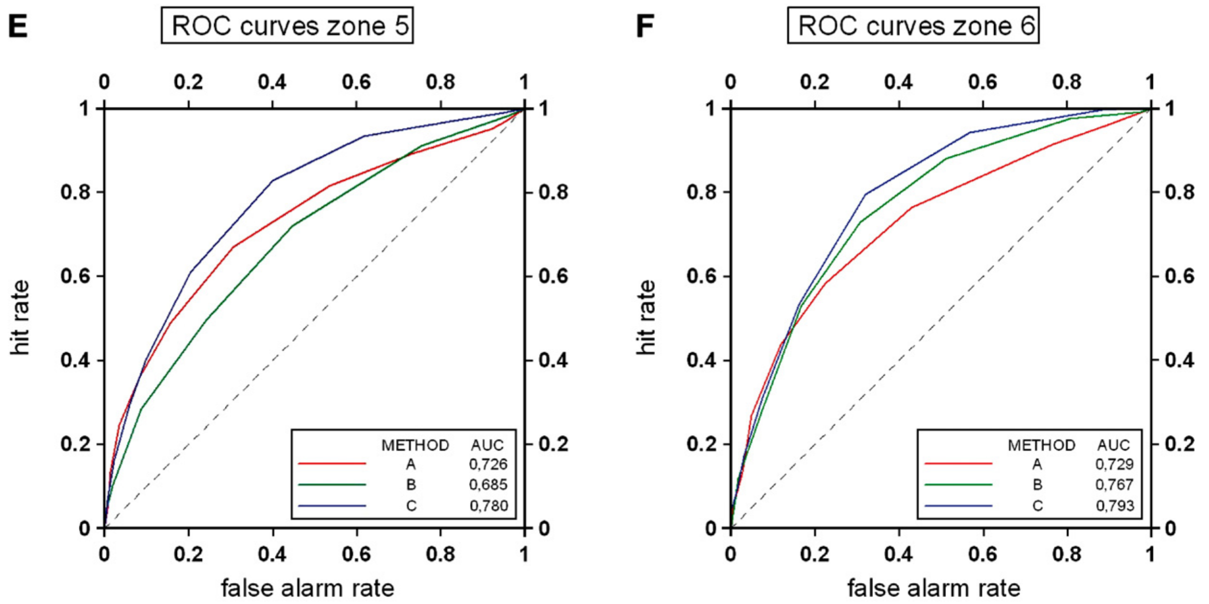

Fig. 8 (*continued*).

4.2. Forecast in different intervals of precipitation

Let N_a be the number of analogues selected for a particular day. If among them there is a number N_i that lies within a rainfall amount interval i, then, the probability of rain in that interval will be given by the percentage of analogous days selected which are in the interval i, that is:

$$p_i = \frac{N_i}{N_a} \tag{9}$$

It is therefore necessary, firstly, to establish the number of rainfall intervals and their limits. On the basis of a climatic study of precipitation in the region (Martin Vide, 1995), the following precipitation intervals were established:

– Class 1: $P<1$ mm No rain or negligible rainfall recorded
– Class 2: $1 \text{ mm} \leq P < 10$ mm Light rain
– Class 3: $10 \text{ mm} \leq P < 25$ mm Moderate rain
– Class 4: $25 \text{ mm} \leq P < 50$ mm Heavy rain
– Class 5: $50 \text{ mm} \leq P < 100$ mm Very heavy rain
– Class 6: $100 \text{ mm} \leq P$ Extreme rain

where P is the daily rainfall amount (mm/24 h).

The verification scores most commonly used to evaluate categorical rain forecasts are:

4.2.1. Brier Probability Score (Brier, 1950)

$$\mathrm{PS} = \frac{1}{N}\sum_{n=1}^{N}\sum_{g=1}^{G}(p_{gn} - \delta_{gn})^2 \tag{10}$$

where:

N is the number of forecasts made using a particular method.
G is the number of rainfall intervals or classes.
$p_n = (p_{1n}, p_{2n}, \ldots, p_{Gn})$ is the vector of foreseen probabilities, in which p_{gn} is the probability of the n-th forecast being in class g.
$\delta_n = (\delta_{1n}, \delta_{2n}, \ldots, \delta_{Gn})$ is the vector of probabilities obtained from observed data, in which $\delta_{gn}=1$ if class g is observed and $\delta_{gn}=0$ otherwise; for the n-th forecast.

We have $0 \leq \mathrm{PS} \leq 2$, so $\mathrm{PS}=0$ if each forecast is correct and $\mathrm{PS}=2$ if every forecast is incorrect. The lowest value of PS is for the best forecast. This index measures the accuracy of the forecast. Nevertheless it has got the disadvantage of penalizing the incorrect forecast equally, whether they are in neighboring class or not.

4.2.2. Ranked Probability Score (Epstein, 1969)

The expression used in this paper is the simplest one (Dequé et al., 1988), which is the quadratic error between the vectors of probability observed and forecast:

$$\mathrm{RPS} = \frac{1}{N}\sum_{n=1}^{N}\sum_{g=1}^{G}\left[\sum_{k=1}^{g}(p_{kn}-\delta_{kn})\right]^2 \tag{11}$$

We now have $0 \leq \mathrm{RPS} \leq G-1$, so the lower this value reaches the better the forecast is. In addition to accuracy, this index also measures the reliability of the forecast, since it penalizes the incorrect forecast to a lesser extent when it is close to the observed value.

5. Results

5.1. Results obtained in the predictors selection

Once the quality control was applied and variables indicated in Section 2.2 were calculated for all the radiosoundings series (00 UTC and 12 UTC) of Palma de Mallorca, discriminant analysis and correlation analysis with daily precipitation were carried out (Gibergans-Báguena, 2001). A similar discriminant process was applied by Tudurí and Ramis (1997) to relate severe weather in Mallorca and radiosounding data. The result of the discriminant analysis allowed the selection of the following variables, both for the 00 UTC and for the 12 UTC radiosoundings:

- precipitable water mass in 700–500 hPa stratum
- relative temperature at 850 hPa
- instability indices LI and KI
- isozero pressure
- CAPE
- potential temperature gradient in 850–700 hPa stratum
- equivalent potential temperature gradient of the surface — 950 hPa stratum.

These seven variables have been applied to classify a 90% of the observations in "rain/no rain" days, in order to obtain the discriminant functions. 10% of the observations are withheld during the discriminant analysis in order to validate the results and to compare with the corresponding climatologic percentage. Results obtained show that at 00 UTC 72.6% observations were correctly classified, in comparison to 70.6% that climatology indicated. At 12 UTC 76.5% observations were properly sorted, what means more than the 70.0% value obtained from the climatology. These results reveal the suitability of the criteria followed for the variables selection, as well as the good selection of the 12 UTC radiosounding.

Table 4
Results obtained for POD and RPS by different Analogues Methods applied to the prediction of daily precipitation in Catalonia and the reference prediction (climatology or persistence)

	Climatology	Persistence	Method A	Method B	Method C	Zone
POD	75.01	75.53	75.24	76.27	**77.37**	1
RPS	26.06	33.82	24.66	26.49	**23.23**	
POD	76.41	76.48	76.26	76.91	**79.33**	2
RPS	26.97	37.07	**25.17**	28.91	25.60	
POD	76.12	77.44	76.33	76.81	**78.63**	3
RPS	26.03	34.94	24.23	26.02	**24.13**	
POD	65.57	70.62	71.72	66.77	**75.56**	4
RPS	31.68	44.40	28.40	32.45	**24.76**	
POD	70.40	72.63	72.97	72.38	**76.26**	5
RPS	27.12	34.11	25.40	26.82	**22.42**	
POD	80.81	80.73	80.95	79.40	**81.98**	6
RPS	23.20	31.87	22.49	24.65	**21.36**	

The best forecasting appears in bold letters.

Table 5
Major relative reduction, in POD and RPS, obtained by the corresponding Analogue Method in relation to percentage not explained by the best reference prediction (climatology or persistence)

Zone	Relative reduction (%) POD		Relative reduction (%) RPS	
1	7.5	Method C	10.9	Method C
2	12.1	Method C	6.7	Method A
3	6.1	Method C	7.3	Method C
4	16.8	Method C	21.6	Method C
5	13.8	Method C	17.3	Method C
6	6.1	Method C	7.9	Method C

5.2. Results obtained by the different analogues methods

The methods presented above were applied to forecast daily precipitation in Catalonia for the fall season. The period studied was 1975–1983, corresponding to the intersection of the databases involved: daily precipitation in Catalonia, principal components of the geopotential fields at 00 UTC and Palma de Mallorca radiosoundings at 12 UTC.

Firstly, the verification scores corresponding to the forecasts by climatology and by persistence were calculated, so they can be used initially as standard references for evaluating the skill of these standard forecasting schemes (Hall et al., 1999; Kuligowski and Barros, 1998). They were calculated for all the zones, and in Table 4 are shown and compared with the results obtained by the various analogues methods tested. In this table, the values of the scores indicating best forecasting are emphasized in boldface type. It is important to notice that the column corresponding to Method C presents better values in all

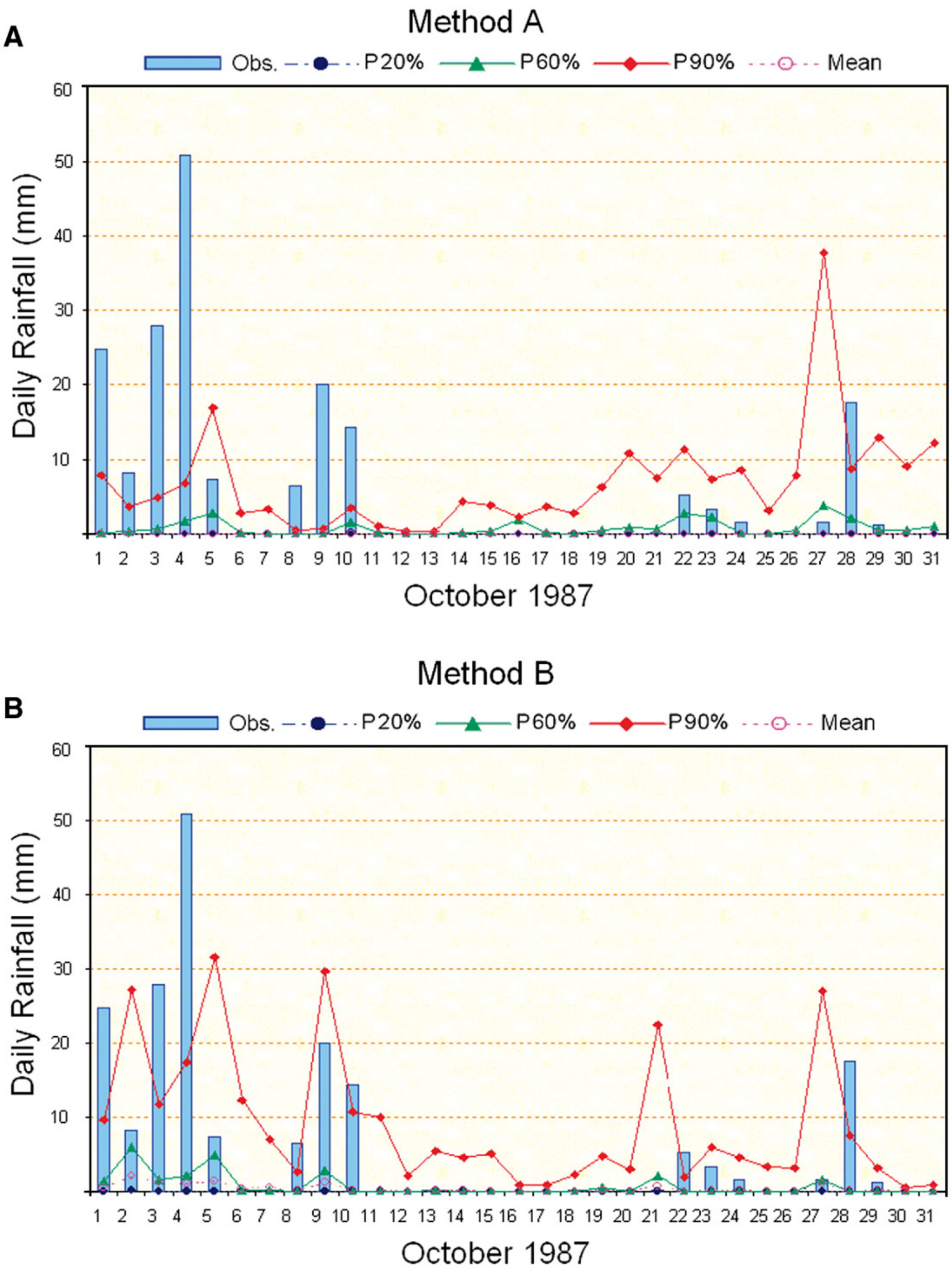

Fig. 9. Results obtained for Zone 1 by Analogues Methods (A, B and C) for October 1987. These figures contain for each day in October: the precipitation observed value and the forecasted quantiles: 20, 60 and 90% and the mean precipitation.

the verification scores for nearly all the pluviometric studied zones.

Table 5 summarizes the relative reduction by the best forecast by analogues of the part of the forecast not explained by the best reference forecast. For example, for Zone 1, the best prediction of reference of rain/no rain would be the persistence with a POD value of 75.53%, while the best prediction for analogues

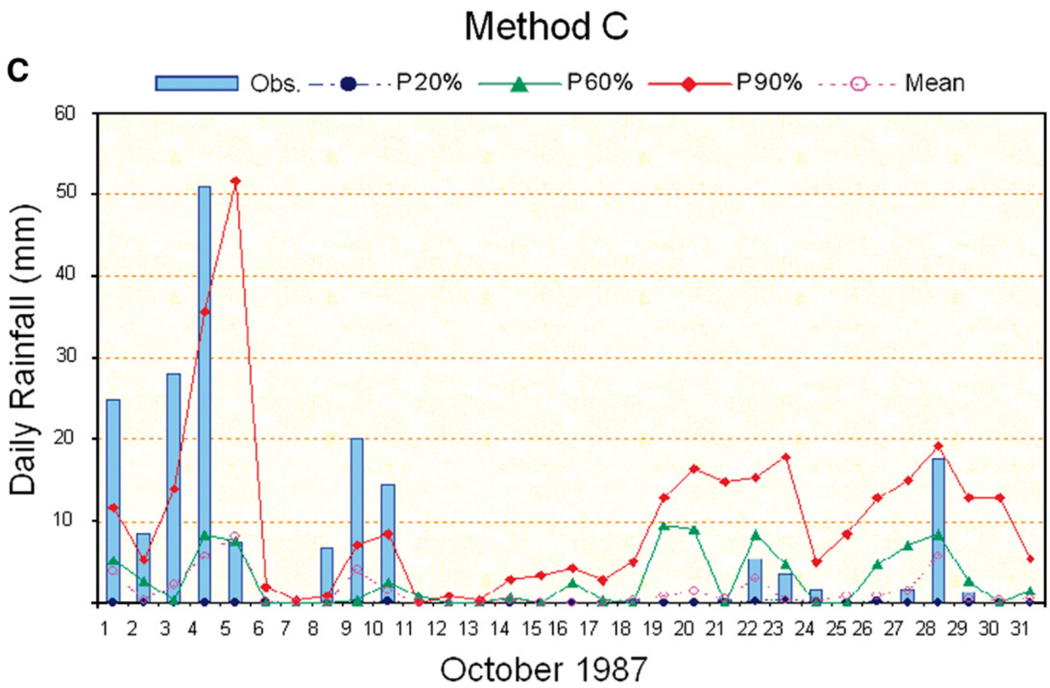

Fig. 9 (*continued*).

corresponds to the Method C with 77.37%. It would imply an improvement of 1.84% (the difference between both percentages) when Method C is applied. Therefore, with the Method C this part not explained would be decreased by 7.5% (calculated as the rate between the improvement and the part not explained by persistence). In the same Table 5 it can be observed that for the rain/no rain forecasting, the Method C provides the best improvement with values between 6% and 17%. Actually, Method A appears only once in this table and Method B does not even appear. The Zone 4 is the one that shows the greatest improvement by means of Method C.

Fig. 8 shows the ROC curve for the forecasting of rain/no rain. These curves indicate that the predictions generated by the three analogue forecasting methods possess real discriminant power. Method C is in all the cases the best discriminant method, with areas ranging from about 0.74 to 0.80 in Zone 4, Zone 5 and Zone 6. The curves corresponding to Method B and Method C corresponding to Zone 6 present a slight asymmetry, indicative of an overestimation of false alarms. The same behaviour is observed for Method C in Zone 5. On the contrary, the asymmetry observed in Zone 4 shows an underestimation trend of rainfall forecasting. Thus, to forecast rain/no rain Method C presented in this paper, which combines the information at synoptic scale with the information from the Palma de Mallorca radiosounding, is the best of all the methods tested.

Method C is also the one that shows the best results when it is applied over daily rainfall classes (mm/24 h): $P<1$, $1\leq P<10$, $10\leq P<25$, $25\leq P<50$, $50\leq P<100$, $100\leq P$. Table 3 shows that the best reference forecast for rainfall classes is given by climatology, with an RPS value of 31.68. On the other hand, when analogues technique is applied, Method C gives the best results, with a value of 24.76 for that score. That means a relative improvement of 21.6% in forecasting of precipitation by classes (Table 4).

With regards to Method B, which takes account only of the information provided by the selected and calculated variables from the Palma de Mallorca radiosounding for 12 UTC, we find that in Zone 1, Zone 2 and Zone 3 it improves upon the forecast offered by Method A for forecasting rain/no rain, while the quality of forecasting by rainfall classes remains the same or slightly lower, though in any case higher than that given by climatology and persistence. With this same method, for Zones 4 and 5, we find that the results obtained are quite significantly worse than those provided by Method A, both as relates to forecasting of rain/no rain and rainfall classes. It is to a certain extent logical that this would happen, for in this case the forecasting was carried out only with information from the Palma de Mallorca radiosounding, and these zones are very far from those affected by invasions from the East.

Finally, with regards to Method A, the POD index values obtained for Zone 1, Zone 2 and Zone 3 are lower than those obtained by persistence, which suggests that persistence is a better method for discriminating between rain/no rain. For the other three zones, Zone 4, Zone 5 and Zone 6, Method A is slightly better. When forecasting in rainfall classes the RPS indices show that, for all the zones, the forecasting by Method A is better than the reference forecasts. It is also important to observe that Zone 5 is one of what has got most rainy days in fall as well as the lowest percentages of isolated days of rain, which makes it hardly surprising that the persistence index gives a better result in this zone than in the others. By means of climatology, the index of probability of detection of rain/no rain, POD, shows values lower than or equal to those of Method A for all the pluviometric zones considered. The same happens with forecasting in rainfall classes; the values of RPS are in all cases very much lower with Method A, which indicates better forecasting in rainfall classes with this method. It is also interesting to observe the results obtained in Zone 4, corresponding to the stations situated more to NW Catalonia and the French Pyrenees zones. This is the zone with the greatest difference in the percentage of successes for the POD index, with 71.72% for Method A compared with the 65.57% given by climatology. Another good result is obtained in Zone 5, also pertaining to the NW Catalonia. It shows a value of 72.97% by Method A, compared with a climatological value of 70.40%. Both zones pertain to the most westerly sector, and so can be more greatly affected by organized weather systems entering from the West and reaching these sectors in fall, as a consequence of the withdrawal towards the South of the high subtropical pressures. These types of precipitations are well detected by Method A, which seeks analogous situations at the synoptic scale.

Table 6
Pluviometric stations used for the Catalonia regionalization: the number corresponds to its location in the map of Fig. 3

1	Pobla de Masaluca
2	Amposta
3	Perello
4	Montblanch
5	Castellvi de la Marca
6	Sant Sadurni d'Anoia
7	Artesa de Segre
8	Terradets-Embalse
9	Gavet-Central
10	Organyà
11	Seu d'Urgell
12	Port de Suert
13	Esterri d'Aneu
14	Tavescan-Presa
15	Viella
16	Arties
17	Berga
18	La Pobla de Lillet
19	Campdevanol
20	Camprodon
21	Calaf
22	Manresa
23	Balsareny
24	Moià
25	Prats de Lluçanets
26	Bongonyà
27	Les Planes d'Hostoles
28	Castellfollit de la Roca
29	Martorelles
30	Caldes de Montbui
31	Cànoves Can Garriga
32	Dosrius
33	Argentona
34	Girona Bell Lloch
35	Calella de Paqlafrugell
36	Jafre
37	Castelló d'Empuries
38	Cadaqués
39	Darnius
40	Le Tec
41	Mont-Louis
42	Hospitalet
43	Usson
44	Salou
45	Artigues

5.3. Results obtained to the heavy rainfall event of October 1987

On the evening of 28 September 1987, heavy rainfall brought the dry Mediterranean summer to an abrupt end in the Catalonian region. Heavy rainfall continued intermittently in Catalonia and southern France for the eight-day period ending 5 October 1987, but most of the precipitation fell during the last four days. Much of the coastal section of Catalonia was affected by flooding, resulting in road closures and disruption of the railway system. By the end of the flood event, ten people had been killed and damage in North-East Spain was estimated at more than 10^9 euros.

The analysis of Palma de Mallorca radiosounding data indicates that the atmospheric structure over the Western Mediterranean was favourable to the development of deep moist convection (Llasat and Ramis, 1993; Ramis et al., 1994).

This section presents the mean quantitative daily rainfall estimated in Zone 1 by the analogue forecasting method, for the entire month of October 1987. Zone 1 has been selected because it was the most affected by the heavy rainfall on 3rd October and 4th October. Although the three methods present good results (Fig. 9), Method C is the best one. When attention is paid to the 90% quantile, Method A presents very low values in comparison with rainfall recorded, but it overestimates the amount of precipitation when low values are recorded. Although Method B is better than Method A for the 90% quantile, it underestimated accumulated precipitation of 3rd October and 4th October. The best estimates are for the precipitation of days 8, 9 and 10 of October.

Finally, Method C shows a good agreement in the monthly evolution as well as on the accumulated precipitation during the heavy rainfall event when 90% quantile is considered. The displacement observed on day 5 might be related with the time interval of 24 h for which accumulated rainfall was recorded. The lap between the maximum cumulated rainfall observed on day 4 and the maximum forecasted rainfall on day 5 could be related to the integration interval of the 24 h cumulated rainfall (between 0700UTC day 4 and 0700UTC day 5). Indeed, a great part of the rainfall was recorded on the night between days 4 and 5.

6. Conclusions

This paper has shown three different analogues methods to obtain quantitative daily rainfall. Those methods were applied over NE of the Iberian Peninsula. In order to have a better spatial coverage, average rainfall values were calculated for homogeneous zones. Method A takes into account only information from the atmosphere at synoptic scale, based on geopotential fields at 700 and 1000 hPa. The other two methods include the use of thermodynamic data to improve the discriminant method. With this aim, a previous quality control and selected thermodynamic variables had been introduced. Method B involves using only information from the Palma de Mallorca radiosounding at 12 UTC. The second improvement consists of an integration of the two preceding methods, that is, taking as a basis an initial selection of analogous days from information at synoptic scale at 00 UTC using the base method, a second selection was then made of analogous days in relation to the Palma de Mallorca radiosounding at 12 UTC. This method has been designated as Method C.

All three methods have been applied to daily data over the period 1975–1989. A heavy rainfall event recorded in October 1987 has also been analyzed. The results from all these methods have been evaluated by calculating three verification scores of successful forecasts: the Probability of Detection (POD) and ROC curves for evaluating the forecasting rain/no rain and the Ranked Probability Score (RPS) to evaluate the forecasting into rainfall classes. Here, six daily rainfall classes (mm/24 h) have been used: $P<1$, $1\leq P<10$, $10\leq P<25$, $25\leq P<50$, $50\leq P<100$, $100\leq P$. The results obtained have been compared with those obtained by taking as a point of reference the methods of forecasting by persistence and by climatology. The values obtained from them led us to draw the following conclusions:

- Method C shows itself to be the best of the three, as well as to forecast rain/no rain and rainfall classes.
- Method B shows good results in the discrimination of days of rain, especially when they originate from Mediterranean situations which are reflected very well by the Palma de Mallorca radiosounding. In such cases it offers better results than Method A.
- Finally, we might note that Method A discriminates days of rain well. On the other hand, it has a tendency to underestimate days with quite significant rainfall and to overestimate days with lighter precipitation.

Although verification results performed here demonstrate that Method C is superior to the other methods, an independent verification analysis would need to be performed before this method was used operationally. This is part of the future work. Besides, future research considers the specific improvement of the quantitative precipitation related with high rainfall events, by using the analogues method.

In order to complete the current work, based on theory of averages values (such as the central limit theorem) future research will consider the specific treatment that extreme values require Coles (2001a,b), Katz et al. (2002), Gilleland and Katz (2005), Beirlant et al. (2004).

Acknowledgements

Our acknowledgement to Électricité de France, to the Instituto Nacional de Meteorología and to the Agencia Catalana de l'Aigua for providing the Palma de Mallorca radiosounding database and the precipitation data. Thanks to Professor Charles Obled of Laboratoire d'Étude des Tranferts en Hydrologie et Environnement (LTHE), of Grenoble, for providing the Nîmes radiosounding database and for his collaboration. Thanks to Dr. Javier Yves Rodriguez, of Électricité de France, and to Dr. Sophie Guilbaud, for all their comments and suggestions. This work has been partially developed into the AMPHORE (Interreg IIIB Medocc 2003-03-4.3-I-079) European project. Finally, our most sincerely acknowledge to Laura Bota (Group of Analysis of Adverse Meteorological Situations, GAMA) for her contribution to the improvement of the reviewed paper.

References

Alpert, P., Ben-Gai, T., Baharad, A., Benjamini, Y., Yekutieli, D., Colacino, M., Diodato, L., Ramis, C., Homar, V., Romero, R., Michaelides, S., Manes, A., 2002. The paradoxical increase of Mediterranean extreme daily rainfall in spite of decrease in total values. Geophys. Res. Lett. 29 (11), 135–154.

Atger, F., 2001. Verifiaction of intense precipitation forecasts from single models and ensemble prediction systems. Nonlinear Process. Geophys. 8, 401–417.

Barret, E.C., Cheng, M.C., 1996. The identification and evaluation of moderate to heavy precipitations areas using IR and SSM/I satellite imagery over the Mediterranean region for the STORM Project. Remote Sens. Rev. 14, 119–149.

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., 2004. Statistics of Extremes. Wiley, Chichester, England.

Bolton, D., 1980. The computation of equivalent potential temperature. Mon. Weather Rev. 108, 1046–1053.

Bonell, M., Summer, G., 1992. Autumn and winter daily precipitation areas in Wales, 1982–1983 to 1986–1987. Int. J. Climatol. 12, 77–102.

Boni, G., Conti, M., Dietrich, S., Lanza, L., Marzano, F.S., Mugnai, A., Panegrossi, G., Siccardi, F., 1996. Multisensor observations during the event of November 4–6, 1994 over Northern Italy. Remote Sens. Rev. 14, 91–117.

Bontron, G., 2004. Prévision quantitative des précipitations: Adaptation probabiliste par recherche d'analogues. Utilisation des Réanalyses NCEP/NCAR et application aux précipitations du Sud-Est de la France. (Quantitative forecast of precipitations: Probabilistic adaptation by search for analogues. Use of Réanalyses NCEP/NCAR and application to precipitations of the South-east of France). Institut National Polytechnique de Grenoble. 276 pp.

Brier, G.W., 1950. Verification of forecast expressed in terms of probability. Mon. Weather Rev. 78, 1–3.

Coles, S., 2001a. S-plus Functions for Extreme Value Modeling: An Accompaniment to the Book An Introduction to Statistical Modeling of Extreme Values. Available online at http://www.stats.bris.ac.uk/~masgc/ismev/uses.ps.

Coles, S., 2001b. An Introduction to Statistical Modeling of Extreme Values. Springer-Verlag, London.

Daan, H., 1984. Scoring rules in forecast verification. PSMP Publication Series, vol. 4. World Meteorological Organization. 62 pp.

Daan, H., 1985. A standardised verification scheme for local weather forecasts. Technical Memorandum, vol. 116. ECMWF, Shinfield Park, Reading. December.

Dequé, M., Royer, J.F., Veysserire, J.M., 1988. Estimation de la qualité des prévisions probabilistes par classes ordonnées. (Estimation of probabilistic forecast quality by ordered classes). Note de travail n° 297 de l'Etablissement d'Etudes et Recherche Météorologiques.

Diomede, T., Nerozzi, F., Paccagnella, T., Todini, E., 2006. The use of meteorological analogues to account for LAM QPF uncertainty. Hydrol. Earth Syst. Sci. Discuss. 3, 3061–3097 (www.hydrol-earth-syst-sci-discuss.net/3/3061/2006/).

Dobryshman, E.M., 1972. Review of forecast verification techniques. WMO Technical Note, vol. 120.

Doswell, C.A., Davies-Jones, R., Keller, D.L., 1990. On summary measures of skill in rare event forecasting based on contingency tables. Weather Forecast. 5, 576–585.

Doswell, C.A., Brooks, H.E., Maddox, R.A., 1996. Flash flood forecasting: an ingredients-based methodology. Weather Forecast. 11, 560–581.

Doswell, C.A., Ramis, C., Romero, R., Alonso, S., 1998. A diagnostic study of three heavy precipitation episodes in the western Mediterranean region. Weather Forecast. 3, 102–124.

Duband, D., 1970. Reconnaissance dynamique de la forme des situations météorologiques. Application à la prévision quantitative des précipitations. (Dynamic recognition of the weather situations form. Application to rainfall quantitative forecast). Thèse de 3ème cycle de la Faculté des Sciences de Paris.

Duband, D., 1974. Reconnaissance dynamique de la forme des situations météorologiques. Application à la prévision quantitative des précipitations (Dynamic recognition of the weather situations form. Application to rainfall quantitative forecast). Congrès de la Société Hydrotechnique de France, XIIIèmes journées de l'Hydraulique, Paris, France.

Duband, D., 1980. Dynamic selection of analogue flow patterns to produce quantitative precipitation forecasts. WMO symposium on Probabilistic and statistical methods in weather forecasting, Nice-France, Septembre, pp. 487–492.

Duband, D., 1981. Prévision spatiale des hauteurs de précipitations journalières. (Spatial forecast heights of daily precipitations). La Houille Blanche,° 7/8, 497–512.

Epstein, E.S., 1969. A scoring system for probability forecasts of ranked categories. J. Appl. Meteorol. 8, 985–987.

Fujita, T.T., Bradbury, D.L., Van Thullenar, C.F., 1970. Palm Sunday tornadoes of April 11, 1965. Mon. Weather Rev. 98, 26–29.

Galway, J.G., 1956. The lifted index as a predictor of latent instability. Bull. Am. Meteorol. Soc. 37, 528–529.

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

Gibergans-Báguena, J., 1993. Aproximación a una tipología de lluvias extremas: Análisis de las precipitaciones superiores a 100 mm en 24 h. (Approach to an extreme rainfall typology. Analysis of precipitations above to 100 mm in 24 h). Ph.D. dissertation, Dept. Astronomy and Meteorology, University of Barcelona. 92 pp. [Available from University of Barcelona, Avda. Diagonal, 647, Spain].

Gibergans-Báguena, J., 2001: Métodos hidrometeorológicos aplicados a la predicción de precipitación diaria en Cataluña. (Applied hydrometeorological methods for the daily rainfall forecast in Catalonia). Thesis. Dept. Astronomy and Meteorology. University of Barcelona. 350 pp. [Available from University of Barcelona, Avda. Diagonal, 647, Spain].

Gilleland, E., Katz, R.W., 2005. A Tutorial for Extremes Toolkit: Weather and Climate Applications of Extreme Value Statistics. Available online at http://www.assessment.ucar.edu/toolkit.

Guilbaud, S., 1997. Prevision quantitative des precipitations jounalières par une méthode tatistico-dynamique de recherche d'analogues (Prediction of quantitative precipitation by a statistical-dynamical method by finding analogs), Ph.D Thesis, L'Institute National Polytechnique de Grenoble, France.

Gutzler, D.S., Shukla, J., 1984. Analogs in the wintertime 500 mb height field. J. Atmos. Sci. 41, 177–189.

Hall, T., Brooks, H.E., Doswell III, C.A., 1999. Precipitation forecasting using a neural network. Weather Forecast. 14, 338–345.

Jansà, A., Alpert, P., Buzzi, A., Arbogast, P., 2001a. MEDEX, Cyclones that Produce High Impact Weather in the Mediterranean. Available at http://medex.inm.uib.es.

Jansà, A., Genovés, A., Picornell, M.A., Campins, J., Riosalido, R., Carretero, O., 2001b. Western Mediterranean cyclones and heavy rain. Part 2: statistical approach. Meteorol. Appl. 8, 43–56.

Kaiser, H.F., 1958. The Varimax criterion for analytical rotation in factor analysis. Psychometrika 23, 187–200.

Katz, R.W., Parlange, M.B., Philippe, N., 2002. Statistics of extremes in hydrology. Adv. Water Resour. 25, 1287–1304.

Kuligowski, R.J., Barros, A.P., 1998. Experiments in short-term precipitation forecasting using artificial neural networks. Mon. Weather Rev. 126, 470–482.

Lana, X., Fernandez, G., Burgueño, A., 1995. Daily precipitation maxima in Catalonia (north-east Spain): expected values and their spatial distribution. Int. J. Climatol. 14, 341–354.

Llasat, M.C., 1987. Episodios de lluvias copiosas en Cataluña: génesis, evolución y factores coadyuvantes. (Episodes with abundant rainfall in Catalonia: helping genesis, evolution and factors). Thesis. Dept. Astronomy and Meteorology. University of Barcelona, n° 40. 543 pp. [Available from University of Barcelona, Avda. Diagonal, 647, Spain]. University of Barcelona, no. 40. 543 pp.

Llasat, M.C., 2004. Les crues mediterraneennes récentes et historiques (Espagne, France, Italie), consequences-enseignements-projects (The recent and historical Mediterranean risings (Spain, France, Italy), consequence-teachings-projects). Colloque SHF Etiages et crues extrêmes régionaux en Europe. Perspectives historiques. Lyon. Publications S.H.F.

Llasat, M.C., Puigcerver, M., 1992. Pluies extremes en Catalogne: Influence orographique et caractéristiques synoptiques (Extreme rainfall in Catalonia: orographical influence and synoptic characteristics). Hydrol. Cont. 2, 99–115.

Llasat, M.C., Puigcerver, M., 1994. Meteorological factors associated with floods in the North-Eastern part of the Iberian Peninsula. Nat. Hazards 9, 81–83.

Llasat, M.C., Ramis, C., 1993. Anàlisi de la situació sinòptica que desenvolupà les inundacions de l'octubre de 1987 a Catalunya. Llibre Homenatge al Dr. Jardí, pp. 93–111.

Llasat, M.C., Ramis, C., Barrantes, J., 1996. The meteorology of high-intensity rainfall events over the west Mediterranean region. Remote Sens. Rev. 14, 51–90.

Llasat, M.C., Gibergans-Báguena, J., Guilbaud, S., Rodriguez, J.Y., 1997. The use of the vertical thermodynamic data in the daily rain classification. Application to the improvement of the analog technique. INM/WMO International Symposium on Cyclones and Hazardous Weather in the Mediterranean Regions, pp. 395–402.

Llasat, M.C., Ramis, C., Lanza, L., 1999. Storm tracking and monitoring using objective synoptic diagnosis and cluster identification from infrared meteosat imagery: a case study. Meteorol. Atmos. Phys. 71, 139–155.

Llasat, M.C., Rigo, T., Barriendos, M., 2003. The 'Montserrat-2000' flash-flood event: a comparison with the floods in the Northeastern Iberian Peninsula since the 14th century. Int. J. Climatol. 23, 453–469.

Lorenz, E.N., 1969. Atmospheric predictability revealed by naturally occurring analogues. J. Atmos. Sci. 26, 636–646.

Mariani, S., Casaioli, M., Accadia, C., Llasat, M.C., Pasi, F., Davolio, S., Elementi, M., Ficca, G., Romero, R., 2005. A Limited Area Model intercomparison on the "Montserrat-2000" flash-flood event using statistical and deterministic methods. NHESS 5, 565–581 (HYDROPTIMET).

Martin Vide, J., 1995. La pluviometría en Catalunya. Butll. Soc. Catalana Ciènc. XV (1).

Mason, I., 1982. A model for assessment of weather forecasts. Aust. Meteorol. Mag. 30, 291–303.

Mason, S.J., Graham, N.E., 1999. Conditional probabilities, relative operating characteristics and relative operating levels. Weather Forecast. 14, 713–725.

Miller, R.C., Bidner, A., Maddox, R.A., 1971. The use of the computer products in severe weather forecasting (the SWEAT index). Preprints Seventh Conf. Severe Local Storms. Kansas City. Am. Meteorol. Soc., pp. 1–6.

Miller, R.C., Bidner, A., Maddox, R.A., 1972. Notes on analysis and severe storm forecasting procedures of the air force global weather control. AFGWC Tech. Rep 200, Air Weather Service. US Air Force. 102 pp.

Mood, A.M., Grybill, F.A., Boes, D.C., 1974. Introduction to the Theory of Statistics. McGraw-Hill. 480 pp.

Muirhead, R.J., 1982. Aspects of Multivariate Statistical Theory. Wiley-Interscience. 673 pp.

Murphy, A., 1985. Proposed standard procedures for verification of local weather forecasts. WMO PSMP, vol. 15.

Murphy, A., 1990. Forecasts verification: its complexity and dimensionality. MWR, vol. 119.

Murphy, A., Winkler, R.L., 1987. A general framework for forecast verification. Mon. Weather Rev. 115, 1330–1338.

Pellerin, G., Lefaivre, L., Houtekamer, P., Girard, C., 2003. Increasing the horizontal resolution of ensemble forecasts at CMC. Nonlinear Process. Geophys. 10, 463–468.

Obled, Ch., Bontron, G., Garçon, R., 2002. Quantitative precipitation forecasts: a statistical adaptation of model outputs though an analogues sorting approach. Atmos. Res. 63, 303–324.

Pickup, N.M., 1982. Considerations of the effect of 500 hPa cyclonity on the success of some thunderstorm forecasting techniques. Meteorol. Mag. 111, 87–97.

Radinovic, D., 1975. An analogue method for weather forecasting using the 500/1000 mb relative topography. Mon. Weather Rev. 103, 639–649.

Ramis, C., Llasat, M.C., Genovés, A., Jansà, A., 1994. The October-87 floods in Catalonia. Synoptic and mesoscale mechanisms. Meteorol. Appl. 1, 337–350.

Ramis, C., Llasat, M.C., Genovés, A., Jansà, A., 1995. Las observaciones de la atmósfera libre en Mallorca: Una breve historia y algunos resultados (The free atmosphere observations in Majorca: A brief history and some results). Rev. Ciènc. (IEB) 17, 41–58.

Rigo, T., Llasat, M.C., 2007. Analysis of Mesoscale Convective Systems in Catalonia (NE of Spain) using radar for the period 1996–2000. Atmos. Res. 83, 458–472.

Riosalido, R., Ferraz, J., Alvarez, E., Cansado, A., Martín, F., Elízaga, F., Camacho, J.L., Martín, A., 1997. A flash flood event in the Spanish Pyrenees: the Biescas case. INM/WMO International Symposium on cyclones and hazardous weather in the Mediterranean, Palma de Mallorca, April 1997, pp. 151–158.

Roebber, P.J., Schultz, D.M., Colle, B.A., Stensrud, D.J., 2004. The risks and rewards of high resolution and ensemble numerical weather prediction. Weather Forecast. 19, 936–949.

Romero, R., Ramis, C., Alonso, S., 1997. Numerical simulation of an extreme rainfall event in Catalonia: role of orography and evaporation from the sea. Q. J. R. Meteorol. Soc. 123, 537–559.

Schultz, D.M., Schumacher, P.N., Doswell, C.A., 2000. The intricacies of instabilities. Mon. Weather Rev. 128 (12), 4143–4148.

Showalter, A.K., 1953. A stability index for thunderstorm forecasting. Bull. Am. Meteorol. Soc. 34, 250–252.

Stanski, R., 1989. Survey of common verification methods in meteorology. WMO WWW TR, vol. 8.

Swets, J.A., 1973. The relative operating characteristic in psychology. Science 182, 990–1000.

Timn, N., 2002. Applied multivariate analysis. Springer Texts in Statistics. 693 pp.

Toth, Z., 1989. Long-range weather forecasting using an analog approach. J. Climate 2, 594–607.
Tudurí, E., Ramis, C., 1997. The environments of significant convective events in the Western Mediterranean. Am. Meteorol. Soc. 12, 294–306.
Wackernagel, H., 2003. Multivariate Geostatistics: An Introduction With Applications. Springer-Verlag. 291 pp.
Weisman, M.L., Klemp, J.B., 1986. Characteristics of Isolated convective storms. In: Ray, P.S. (Ed.), Mesoscale Meteorology and Forecasting. Am. Meteorol. Soc., pp. 331–358.
Wilks, D.S., 1995. Statistical Methods in the Atmospheric Sciences. Academic Press. 467 pp.