**Natural Hazards
and Earth
System Sciences**

A tool for assessing the quality of the Mediterranean cyclone forecast: a numerical index

M. A. Picornell[1]**, A. Jansà**[1]**, and A. Genovés**[2]

[1]D.T. Illes Balears, AEMET, Palma de Mallorca, Spain
[2]DDA, AEMET, Madrid, Spain

Abstract. Cyclones affecting the Mediterranean region, sometimes related to severe weather events, are often not well represented enough in numerical model predictions.

Assessing the quality of the forecast of these cyclonic structures would be a significant advance in better knowing the goodness of the weather forecast in this region, and particularly the quality of predictions of high impact phenomena.

In order to estimate the cyclone forecast uncertainty in operational models, in this work we compare two cyclone databases for the period 2006–2007: one from the operational analyses of the T799 ECMWF deterministic model; and the other from the forecasts provided by the same model in three ranges, H+12, H+24, and H+48. The skill of the model to detect cyclones and its accuracy in describing their features are assessed.

An index is presented as an indicator of the quality of the prediction, derived from the frequency distribution of errors in the prediction of four characteristics of the cyclone: position, central pressure value, geostrophic circulation, and domain. Some sub-indexes are derived to verify each of the variables separately in order to analyse the most frequent sources of error. Other sub-indexes are also defined to indicate possible biases in the numerical prediction model.

1 Introduction

An essential aspect in the weather prediction process is verification, that is, a comparison of predicted weather against observed weather or a good estimate of true outcome. In recent times much effort has been devoted to explaining that the forecast should be verified (Murphy, 1993) and to finding suitable methods for this (JWGFVR, 2008; Stanski et al., 1989). Verification is a key step to improving the forecast process, as information concerning the scale and features of forecast errors is obtained and possible sources of error can be identified, in order to monitor forecast quality and to compare the quality of different forecast systems (Brooks and Doswell, 1996; Wernli et al., 2008; JWGFVR, 2008).

This task involves some intrinsic difficulties: firstly, the observed weather has to be described properly. Many variables can be used to describe the weather and the most representative ones should be selected. Furthermore, these variables can derive from observations (e.g. remote sensing, surface observations) or from numerical model analyses; moreover, all of them entail an associated error. Another difficulty is how to handle so much information. Some suitable statistical skill scores can be selected, or created if necessary, to combine the large quantity of information obtained from this process. These indexes can also be useful to quantify the variation of the skill of numerical models in forecasting these events and to assess the improvement of the model over time (Charles et al., 2009).

Weather in the Mediterranean is sometimes related to the presence of cyclones which, from time to time, produce severe weather events. Assessing the forecast quality of these cyclonic structures would be a significant advance to better know the goodness of the weather forecast in this region and, particularly, the quality of predictions of high impact phenomena. A more or less accurate representation of these cyclonic structures in the analysis and forecasts from numerical models can be an indicator of the quality of the description of the state of the atmosphere. The skill of the numerical weather prediction model in forecasting cyclones has been the subject of previous works for the Mediterranean area (Atger, 1997) and for North America (Charles et al., 2009).

Correspondence to: M. A. Picornell
(mpicornella@aemet.es)

The aim of this paper is to present a tool for comparing the detection and description of cyclones in different outputs of the same or different models that also quantifies the similarity of the different descriptions. This tool also attempts to analyse the origin of these differences. This procedure can be applied to compare analyses from different numerical models, as well as analyses and forecasts of different scopes and predictions obtained from Ensemble Prediction Systems (EPS).

In a previous work, an objective methodology to assess and quantify cyclone forecast error was developed (Picornell et al., 2002) and applied to investigate the performance of the HIRLAM(INM)-0.5 model in predicting surface cyclones. In the present work, some aspects of this methodology have been modified in order to reduce the subjectivity in some steps. Moreover, some new numerical indexes are proposed to estimate the cyclone forecast quality, to discriminate whether the forecast has over or underestimated some magnitudes in the description of the cyclones, and even to report on the prediction of each cyclone feature. Finally, the resultant methodology has been applied to investigate the performance of the ECMWF T799 operational model in predicting surface cyclones.

2 Cyclone databases

In the present work, in order to apply the verification methodology over a period of one year, from spring 2006 to winter 2007, two cyclone databases were obtained from the T799 ECMWF deterministic model which was operational for those years. The cyclones were detected and described using a previously automated procedure described in Picornell et al. (2001). The procedure was originally designed to describe Mediteranean cyclones and their specific characteristics. It has been successfully applied for climatological purposes in different objective analysis outputs on different lat-lon regular grids (Picornell et al., 2001; Gil et al., 2002; Campins et al., 2006, 2010). Nowadays, the procedure can be considered a robust enough tool for objectively describing not only the usual extratropical Mediterranean cyclones, but also the shallow mesoscale depressions frequently observed in the area.

In the aforementioned procedure, a cyclone is defined as a relative minimum in the mean sea level pressure (m.s.l.p.) field with a mean pressure gradient around the centre greater than 0.5 hPa/100 km in at least six of eight main directions. For each cyclone, the date, location, domain, radii, geostrophic vorticity, and geostrophic circulation (among other magnitudes) are estimated. The cyclone domain (defined as the area of positive geostrophic vorticity around the cyclone centre) is obtained looking for the zero-vorticity line around the low-pressure centre in sixteen directions. Cyclone intensity is measured by the geostrophic circulation
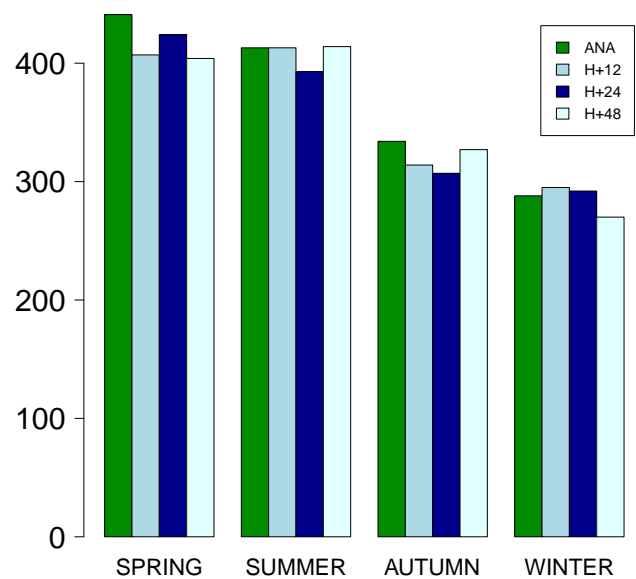

Fig. 1. Number of detected cyclonic centres in the m.s.l.p. analyses (ANA), H+12, H+24, and H+48 forecast fields from spring 2006 to winter 2007.

over the domain and it is expressed in Circulation Units, where $1\,\mathrm{CU} = 10^7\,\mathrm{m}^2\,\mathrm{s}^{-1}$ (Sinclair, 1997).

The two databases used in this work were obtained by applying the procedure on the T799 ECMWF deterministic model mslp fields twice, once on the sea level pressure analyses and then one on the mslp forecasts fields, over an area from 32° N to 56° N and from 32° W to 24° E. All fields were interpolated to a 0.25° × 0.25° lat-lon grid and smoothed using a Cressman filter with 200 km of radius in order to obtain an adequate description of the cyclones. Three different forecast ranges, +12, +24, and +48 h (two forecasts per day 00Z and 12Z), were considered to investigate temporal trend in the forecast quality.

The definition of cyclone is not very restrictive and a large number of cyclonic centres were detected, some of them small and weak. The maximum frequency of cyclones occurs in spring and summer and the lowest in winter (see Fig. 1). No big differences in the cyclone magnitude values between the three forecast ranges and the analyses were observed.

3 Verification procedure

In order to assess the model's ability to forecast cyclones, the predicted cyclone database was compared against the corresponding cyclone database from the analysis fields, which is considered a reliable reference. From this comparison an error sample was obtained for one year.

The verification procedure follows three steps (as in Picornell et al., 2002; following Atger, 1997). First of all, the numbers of forecast and analysed cyclones are compared

and, secondly, the accuracy of this forecast is assessed by comparing several cyclone magnitudes. In the third step, in order to quantify forecast quality, some indexes are defined.

3.1 Detection performance

In order to compare the analysed and forecast fields valid for the same time, a criterion to decide whether an analysed cyclone was forecast should be established. In this work an analysed cyclone is regarded as a correct forecast if in the corresponding forecast field a cyclone is located at a distance shorter than 400 km from the analysed one. If more than one forecast cyclone is located within a 400 km radius around the observed cyclone, the closest forecast cyclone is paired with the observed cyclone. The distance threshold of 400 km between analysed and forecast cyclones was selected by reaching a compromise: for a longer distance, the number of hits increases in some cases, but the number of wrong relations will also probably increase, and forecast quality can be adversely affected.

First, the location of the cyclones is compared and the distance between the positions of cyclones in the analysis in the forecast is calculated. This distance depends on the time forecast range: distance distributions for different forecast ranges reveal that, in most cases, the forecast cyclone at H+12 is closer to the analysed cyclone than the forecast cyclone at H+24 and H+48, as Fig. 2a shows. The mean distance between forecast and analysed cyclones increases from 50.18 km for H+12, to 75 km for H+24, and to 118 km for the longest range.

The frequency distribution of the distance for the whole sample (Fig. 2b) shows that less than 3 % of forecast cylones are farther than 300 km from the analysed centres. These statistical results reinforce the chosen criterion and support the hypothesis that the distance of 400 km is appropriate to relate the most analysed and forecast cyclones.

To assess the detection performance, the number of predicted and analysed cyclones are compared. The number of hits (analysed and forecast cyclones, H), false alarms (forecast but not analysed centres, FA), misses (analysed but not forecast centres, M), and correct negatives (number of charts without analysed and without forecast cyclones, CN) are counted and compiled in a contingency table for each forecast range. Hit Rate (HR) and False Alarm Rate (FAR) are obtained as well as other skill scores (see Table 1). As shown in Fig. 1 in the previous section, the number of detected cyclones is similar in the three forecast ranges, but the number of hits decreases when the forecast range increases, therefore the HR value decreases. The correct forecasts of nonocurrence dominate the contingency table and it affects some skill scores such as FAR.

Bearing this in mind, a second assessment scheme is used based on the correspondence between the distribution of forecast and analysed categories. The multi-category contingency table shows the frequency of forecasts and observa-
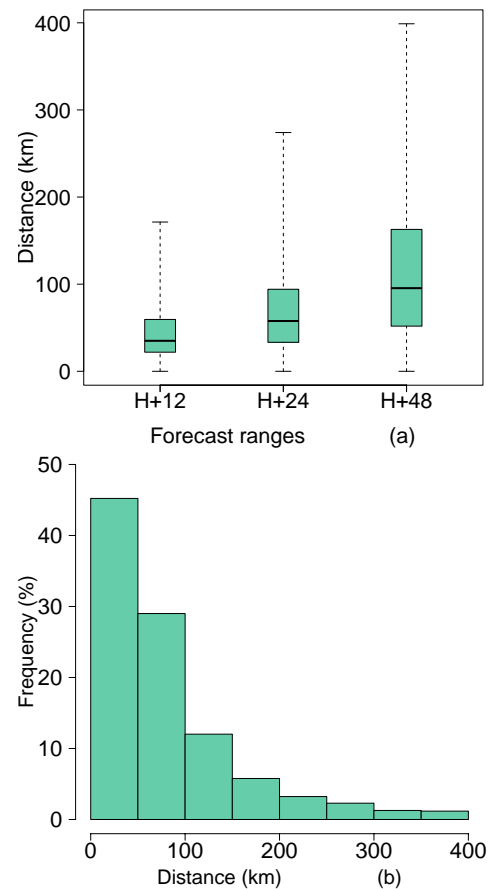

Fig. 2. Frequency distribution of distance (km) between forecast and analysed centres, **(a)** for three forecast ranges and **(b)** for all forecast ranges.

Table 1. Skill score values (HR, HSS, and HK; perfect score: 1, FAR, perfect score:0).

	HR	FAR	HSS	HK
H+12	0.8332	0.0010	0.5747	0.5734
H+24	0.7382	0.0013	0.4887	0.4881
H+48	0.6536	0.0018	0.3394	0.3397

tions in the various bins. An examination of the relationship between the elements in the multi-category contingency table gives information as to whether forecasts produce the correct distribution of cyclones when compared to the observations. From off-diagonal elements, the nature of the forecast errors can be more easily diagnosed (JWGV, 2008).

The cyclones in the sample are categorized according to their intensity, measured by geostrophic circulation. Twelve categories of 1 CU are defined and a multi-category contingency table for each forecast range is obtained. In this work, a row and a column is added to the contingency table,

corresponding to the zero-category, in order to obtain the distribution of False Alarms (FA) and Misses (M): when a cyclone is forecast but not present in the analysis (FA), it is assigned to column 0; or when the analysed cyclone is not forecast (M), it was was assigned to row 0. Cyclones whose intensity has been correctly forecast are counted in the diagonal elements, (in bold in the Table 2). In accordance with the multi-category contingency table for H+12 (Table 2), 65 % of hits were forecast in the right category. The over-diagonal elements show an underestimate of the intensity and 16 % of cyclones were forecast in an inferior category. Only about 2 % of the forecasts estimated were under-predicted with a greater difference of intensity. Under-diagonal elements show an overestimate, and 14 % were forecast in a superior category.

To summarize the performance of multi-category forecasts, two skill scores can be used, the Heidke skill score (HSS), and Hansen and Kuipers discriminant (HK), which indicate the accuracy in predicting the correct category compared with chance (JWGV, 2008). Again, the highest skill score values occur for the shortest range. For H+24 and H+48 the number of diagonal elements decreases, that is, detection performance worsens when forecast range increases, and, consequently, the HSS and HK values are lower, as shown in Table 1. Most of the analysed cyclones are correctly detected by the forecasts, but their intensity is not well forecast.

3.2 Forecast accuracy

To assess forecast accuracy, the error distribution of some of the cyclone features is examined. Only hits are considered in this step of the process. The magnitudes of some parameters characterizing the analysed and forecast cyclones are compared: central pressure value, geostrophic vorticity, geostrophic circulation, and cyclone domain. The forecast error is measured by calculating the differences in the value of these parameters. To illustrate these errors, Fig. 3a shows the distribution of differences in central pressure value, geostrophic circulation, and area. For these magnitudes, negative or positive difference values indicate that the variable is underpredicted or overpredicted, respectively. It must be taken into account that an underestimate or overestimate of the central pressure value involves an overestimate or underestimate, respectively, of the cyclone depth.

A common characteristic of all error distributions are their quasi-Gaussian shape and their dependence on the forecast range. For H+12 the difference values are more concentrated around zero, but for H+24 and especially H+48 the values vary to a much greater extent. As the difference distributions are similar for the three forecast ranges, the three error databases are jointly collected and an error sample is obtained for a reference period of one year.

Two simple measurements that convey some information concerning forecast performance are calculated: bias, which indicates whether the forecast system has a tendency to underforecast or overforecast events, and root mean square error (RMSE), which informs about the average magnitude of the errors. Figure 3b shows that for shortest ranges the bias is nearly zero and, although the bias value increases for H+48, no strong signal of bias is observed. RMSE curves become larger as the forecast range increases, so the error in the pressure value, geostrophic circulation, and area also increases with forecast range.

3.3 Index

The third step in the verification process consists in constructing an index based on the forecast errors of some characteristics (n) of the cyclone, which could quantify the accuracy of the model to forecast cyclones. In an initial attempt, an index based on the weighted sum of the errors in some characteristics of the cyclone (location, central pressure value, geostrophic circulation, and radii in 16 directions ($n = 19$) was defined (Picornell et al., 2002). The weight assigned to each difference was quite subjective and to correct this issue a new index has been defined, weighting the contribution of each error in an objective way.

In the present work, to compare the predicted and observed cyclone shape, the correlation between the sets of 16 radii is considered, rather than the actual values of the 16 radii that were used in Picornell et al. (2002). Thus, the measure of the forecast quality is based on the error of four of the cyclone's features ($n = 4$):

– distance between forecast cyclone position and observed position;

– difference between forecast and observed pressure values at the central point of the cyclone;

– difference in geostrophic circulation; and

– correlation between radii of forecast and observed cyclones in order to measure differences in their shape.

Besides, some sub-indexes are defined to verify each of the characteristics separately in order to analyse the most frequent sources of error. Other sub-indexes are also derived to indicate possible biases.

3.3.1 Contribution of each characteristic error

For each one of these features, the cumulative frequency distribution of the absolute value of the error, F_k ($k = 1, n$, $n = 4$), is obtained for the whole sample and is considered the error reference pattern. The cumulative frequency of the absolute error for one of the characteristics, for instance the geostrophic circulation (GC) $F_3(j)$, is shown in Fig. 4. From

Table 2. Multi-category contingency table for the forecast range H+12. The cyclones were categorized according to their intensity. Twelve categories of 1 CU were defined. False Alarm distribution appears at the zero observed category column (FA). Miss distribution appears in the zero forecast category row (M).

|		FA					Observed Category									TOTAL
		0	1	2	3	4	5	6	7	8	9	10	11	12	
M	0	0	32	108	39	27	9	8	2	4	2	2	0	0	233
	1	24	**16**	9	1	0	0	0	0	0	0	0	0	0	50
	2	127	19	**217**	51	2	0	0	0	0	0	0	0	0	416
	3	72	2	40	**200**	35	10	1	0	0	0	0	0	0	360
	4	30	1	4	39	**132**	30	3	0	0	0	0	0	0	239
Forecast	5	15	0	1	6	28	**93**	32	2	1	0	0	0	0	178
	6	5	0	0	1	1	16	**44**	7	1	1	0	0	0	76
Category	7	6	0	0	0	1	2	6	**24**	18	1	0	0	0	58
	8	1	0	0	0	0	0	2	5	**24**	2	2	0	0	36
	9	0	0	0	0	0	0	0	3	7	**10**	7	0	0	27
	10	0	0	0	0	0	0	0	0	1	1	**3**	2	1	8
	11	0	0	0	0	0	0	0	0	0	0	1	**2**	0	3
	12	0	0	0	0	0	0	0	0	0	0	0	0	**2**	2
	TOTAL	280	70	379	337	226	160	96	43	56	17	15	4	3	1686
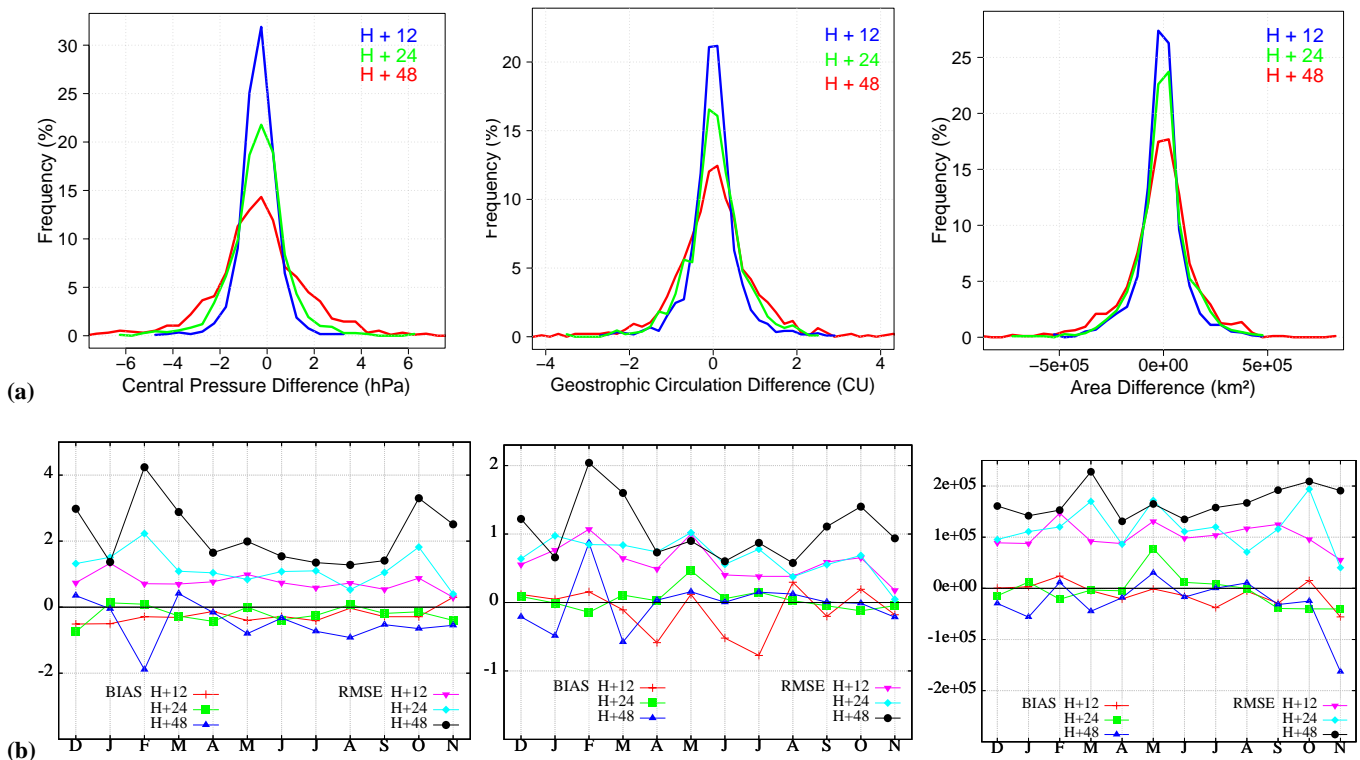
Fig. 3. (a) Frequency distribution (%) of the differences in central pressure (hPa), geostrophic circulation (CU), area (km^2), and **(b)** bias and RMSE for these differences H+12, H+24, and H+48 for each month.
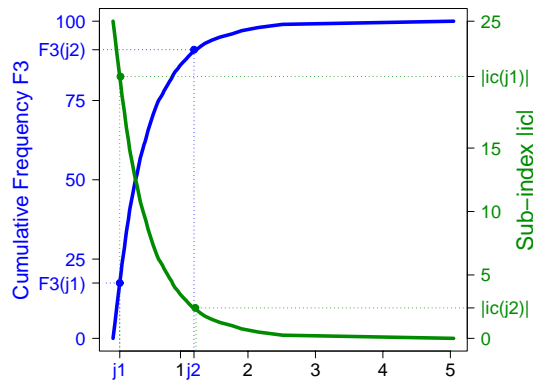
Fig. 4. Cumulative frequency (%) of geostrophic circulation error (in CU), F_3 (in blue), and partial index, $|ic|$ (in green).

Fig. 5. Cyclone distribution as a function of the index value $I(j)$.

$F_3(j)$ a partial index $ic(j)$ is defined in order to quantify the accuracy of the characteristic forecast:

$$ic(j) = (100 - F_3(j))/n \tag{1}$$

and their values for the cyclones in the sample are represented in the right vertical axis in Fig. 4. This index is constructed in such a way that maximum value ($100/n$, in this case 25) occurs when $F_3(j) = 0$ and is therefore assigned to the perfect forecast. From the cumulative frequency distribution $F_3(j)$, for an expected cyclone with a GC error $j1 = 0.1CU$ the value of cumulative frequency $F_3(j1) = 17.4$ is assigned. This value is relatively low and it indicates that a large part of cyclones of the sample have been forecast with a larger GC difference. Therefore, a high index value is assigned, $ic(j1 = 0.1) = 20.6$, and this means that the GC forecast is among the best forecasts of the reference sample, in which case it is considered as a good prediction. However, for a larger error $j2 = 1.2CU$, a high value $F_3(j2) = 91$ is assigned, that is, most of the cyclones of the sample have been forecast with a lower GC error. In this case, the very low index value, $ic(j2) = 2.4$, indicates that the forecast is considered a low quality prediction.

In a similar way, other sub-indexes are defined from the cumulative frequencies of the error of the other characteristics: distance, $id(j)$, central pressure difference, $ip(j)$, and radii correlation, $ir(j)$:

$$id(j) = (100 - F_1(j))/n \tag{2}$$
$$ip(j) = (100 - F_2(j))/n \tag{3}$$
$$ir(j) = (100 - F_4(j))/n \tag{4}$$

If the depth of cyclone or of the circulation has been underestimated, ip or ic values are multiplied by (-1), in such a way that positive values of these partial indexes indicate an overestimate and negative values indicate an underestimate. These sub-indexes can be useful to discriminate the origin of the forecast error. In an attempt to estimate the error in
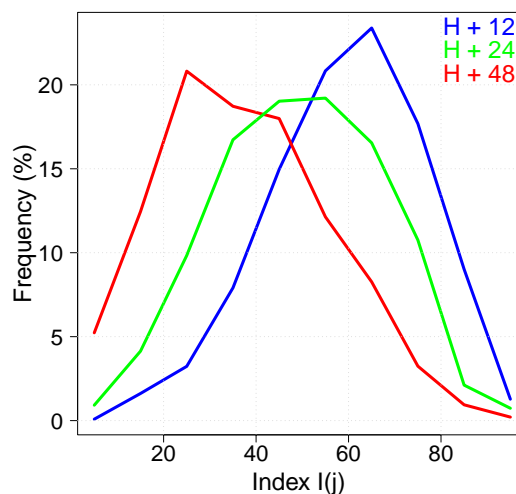
predicting the cyclone at a given time, the error in the characteristics used to describe it can be used jointly. The quality of a prediction is affected by the error in all the parameters. As these sub-indexes are non-dimensional they may be combined jointly. Thus, from these sub-indexes an index, $I(j)$, was constructed to quantify the goodness of the model at forecasting cyclones:

$$I(j) = |id(j)| + |ip(j)| + |ic(j)| + |ir(j)|$$
$$I(j) = 100 - \frac{\sum_{k=1}^{n} F_k(j)}{n}$$

$$id(j), ir(j) \in [0, 100/n] \quad ip(j), ic(j) \in [-100/n, 100/n],$$
$$I(j) \in [0, 100] \tag{5}$$

If $n = 4$

$$I(j) = 100 - \frac{F_1(j) + F_2(j) + F_3(j) + F_4(j)}{4}$$

Perfect forecast:

$$I(j) = 100, id(j) = \ldots = ir(j) = 25 \tag{6}$$

In the future, more characteristics of the cyclone can also be considered and included in the index calculation, in particular the features that describe the evolution of the cyclonic centres.

Two other sub-indexes have been derived: $I_{un}(j)$ and $I_{ov}(j)$, in order to gather information concerning whether the forecast under or overestimates the strength of the cyclone, taking into account the contribution of the differences in the pressure and GC values. Only the error of the characteristics which are underestimated contribute to the sub-index $I_{un}(j)$: if both characteristics are underestimated, $F_2(j)$ and $F_3(j)$ are involved in $I_{un}(j)$ (case 1 of Eq. 7); if only cyclone depth is underestimated, only $F_2(j)$ appears in the equation (case
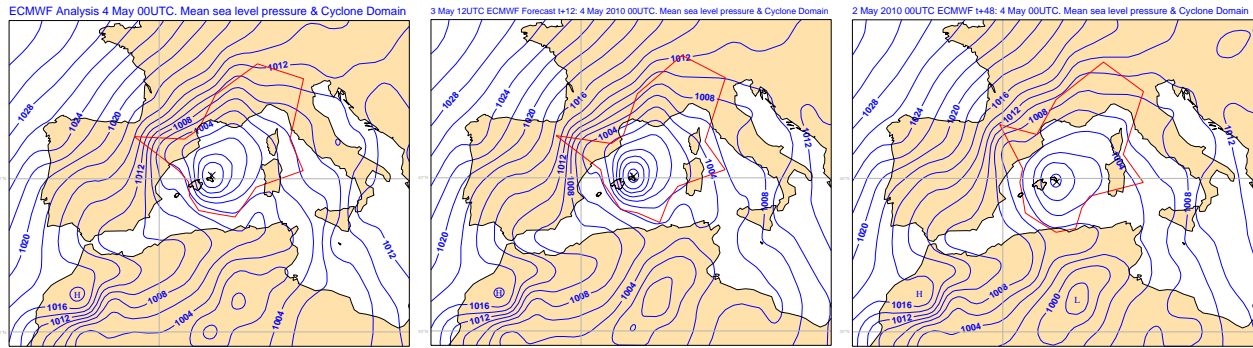
Fig. 6. Mean sea level pressure and Cyclone Domain (red), 4 May 2010 00:00 UTC. ECMWF Analysis (left), Forecasts H+12 (centre), and H+48 (right).

2 of Eq. 7); and the same for $F_3(j)$ if only GC is underpredicted (case 3 of Eq. 7). If any characteristic is underestimated, the index takes its maximum value, $I_{un}(j) = 50$ (case 4 of Eq. 7). The index $I_{ov}(j)$ is obtained similarly from the contributions of the overpredicted characteristic errors, by means Eq. (8).

$$I_{un}(j) = \begin{cases} |ip(j)| + |ic(j)| = 50 - \frac{F_2(j)+F_3(j)}{4} \\ \qquad \text{if } dP0(j) > 0, \quad dGC(j) < 0 \\ |ip(j)| + 25 = 50 - \frac{F_2(j)}{4} \\ \qquad \text{if } dP0(j) > 0, \quad dGC(j) > 0 \\ 25 + |ic(j)| = 50 - \frac{F_3(j)}{4} \\ \qquad \text{if } dP0(j) < 0, \quad dGC(j) < 0 \\ 50 \\ \qquad \text{if } dP0(j) < 0, \quad dGC(j) > 0 \end{cases} \quad (7)$$

$$I_{ov}(j) = \begin{cases} |ip(j)| + |ic(j)| = 50 - \frac{F_2(j)+F_3(j)}{4} \\ \qquad \text{if } dP0(j) < 0, dGC(j) > 0 \\ |ip(j)| + 25 = 50 - \frac{F_2(j)}{4} \\ \qquad \text{if } dP0(j) < 0, dGC(j) < 0 \\ 25 + |ic(j)| = 50 - \frac{F_3(j)}{4} \\ \qquad \text{if } dP0(j) > 0, dGC(j) > 0 \\ 50 \\ \qquad \text{if } dP0(j) > 0, dGC(j) < 0 \end{cases} \quad (8)$$

$$I_{ov}(j) \in [0, 50], I_{un}(j) \in [0, 50]$$
Perfect forecast : $I_{ov}(j) = 50, \quad I_{un}(j) = 50$

For the cylones in the reference sample, the index $I(j)$ was obtained and is shown in Fig. 5. The highest index values occur for the shortest range and lowest index values are reached for the longest range, that is, cyclone forecasts are more precise for the shorter range, which is in accordance with the above results.

3.3.2 Example of application

Finally, an example is shown to illustrate the meaning of the index values. An initial application to a cyclone that does not belong to the reference sample is presented.

Table 3. Differences and index values for two different forecasts (H+12 and H+48) of the cyclone which affected the Balearic Islands on 4 May 2010 at 00:00 UTC.

	dist (km)	difp0 (hpa)	difcir (CU)	r_0
H+12	0.0	−1.00	−0.38	0.99
H+48	86.06	2.40	−1.74	0.82
	id	ip	ic	ir
H+12	25.00	7.75	-11.48	24.54
H+48	7.75	−2.00	−1.04	17.13

	I_{ov}	I_{un}	I
H+12	32.75	36.48	68.77
H+48	50.00	3.04	27.91

On 3–4 May 2010 an intense cyclone affected the Balearic Islands. The cyclone moved from the southern Balearic towards NE, quickly intensifying, with a rate (15 hPa/24 h) slightly lower than explosive cyclogenesis (at 39° N 1 Bergeron = 17 hPa/24 h). A historical record of rain in Palma de Mallorca (471 m^{-2} in one hour), a maximum for the month of May in Mallorca mountains (2341 m^{-2} in 24 h, whith 54l m^{-2} in one hour), and exceptional values of western wind in Minorca were recorded.

The cyclone forecast at H+12 and H+48 on 4 May at 00:00 UTC are verified against the corresponding analysis.

At H+12 the location of cyclonic centre is the same as in the analysis and there is a good correlation between the shape of the forecast and analysed cyclones, as shown in Fig. 6. Therefore the partial indexes id and ir are close to 25 (see Table 3). The forecast cyclone is deeper than the analysed one, that is, there is an overestimate of the depth of the cyclone, as is indicated by the partial index value $ip = 7.75$. On the other hand, geostrophic circulation is slightly under-

estimated, with $ic = -11.48$. The under and overestimated indexes are both less than 50, because one variable is over-predicted and the other is underpredicted. In this case, the total index $I(j)$ is 68.77. This index value indicates that this can be considered a good prediction.

At H+48 the forecast cyclone is located 86 km to the east of the analysed centre (see Fig. 6) and id has a small value, $id = 7.75$ (see Table 3). The low and negative partial sub-indexes ($ip = -2.0$ and $ic = -1.04$, $I_{un} = 3.04$) indicate an underestimate of cyclone depth and geostrophic circulation. In this case no variables are overestimated and I_{ov} attains its maximum value, $I_{ov} = 50$. At H+48 the forecast is worse than at H+12 and this is reflected in a lower index value, $I(j) = 27.91$.

4　Conclusions

– A methodology to assess the quality of the Mediterranean cyclone forecast from a numerical model has been developed, based on an assessment of the differences of some cyclone magnitudes between the analysed and the corresponding forecast cyclone. An index is defined as a measure of forecast quality.

– The method has been applied to the T799 ECMWF model at three different forecast ranges. The mean distance between the locations of cyclones in the analysis and in the forecast is larger for longer ranges. The number of detected cyclones is similar in the three ranges, but the number of hits decreases when the range increases. The accuracy of the forecast also decreases as the forecast range increases.

– A sample of differences in four cyclone characteristcis for one year, without a strong trend to over or underestimation, has been obtained and is used as error reference pattern to assess the prediction of cyclones.

– The index is based on the cumulative frequency of errors and variables with different dimensions are dealt jointly. Other features can be introduced, if necessary.

– This index can be adapted to assess probabilistic forecast of cylones from Ensemble Prediction Systems, and can also be used to compare cyclone databases from different models.

Edited by: A. Bartzokas
Reviewed by: two anonymous referees

References

Atger, F.: Medium-range forecasting of Mediterranean cyclones, INM/WMO Symposium on Cyclones and Hazardous Weather in the Mediterranean, MMA/UIB, Palma de Mallorca, 729–737, 1997.

Brooks, H. E. and Doswell III, C. A.: A comparison of measures-oriented and distributions-oriented approaches to forecast verification, Weather Forecast., 11, 288–303, 1996.

Campins, J., Jansà, A. and Genovés, A.: Three diemnsional structure of Western Mediterranean cyclones, Int. J. Climatol, 26A, 323–343, 2006.

Campins, J., Genovés, A., Picornell, M. A., and Jansà, A.: Climatology of Mediterranean cyclones using the ERA-40 dataset, Int. J. Climatol, article first published online: 14 June 2010, doi:10.1002/joc.2183, Early View, 2010.

Charles, M. E. and Colle, B. A.: Verification of extratropical cyclones within the NCEP operational models, Part I: Analysis errors and short-term NAM and GFS forecasts, Weather Forecast., 24(5), 1173–1190, 2009.

Gil, V., Genovés, A., Picornell, M. A., and Jansà A.: Automated database of cyclones from the ECMWF Model: Preliminary comparison between West and East Mediterranean Basisns, Mediterranean Storms, 4th Plinius Conference 2002, Mallorca, Spain, 2–4 October 2002, Edited by: Jansà, A. and Romero, R., CD-ROM, ISBN: 84-7632-792-7, 2002.

Joint Working Group on Forecast Verification Research: Forecast Verification – Issues, Methods and FAQ, JWGFVR/WWRP/WCRP WWRP/WGNE, available at: http://www.cawcr.gov.au/projects/verification, last access: 6 October 2008, 2008.

Murphy, A. H.: What is a Good Forecast? An Essay on the Nature of Goddness in Weather Forecasting, Weather Forecast., 8, 281–293, 1993.

Picornell, M. A., Jansa, A., Genoves, A., and Campins, J.: Automated database of mesocyclones from the HIRLAM(INM) 0.5 analyses in the Western Mediterranean, Int. J. Climatol., 21, 335–354, 2001.

Picornell, M. A., Carrassi, N. A., and Jansà, A.: A study on the forecasts's quality of the Mediterranean Cyclones, Mediterranean Storms, 4th Plinius Conference 2002, Mallorca, Spain, 2–4 October 2002, Edited by: Jansà, A. and Romero, R., CD-ROM, ISBN: 84-7632-792-7, 2002.

Stanski, H. R., Wilson, L. J., and Burrows, W. R.: Survey of common verification methods in Meteorology, WMP, World Weather Watch, WMO/TD 358, 114 pp., 1989.

Sinclair, M.: Objective Identification of Cyclones and Their Circulation Intensity, and Climatology, Weather Forecast., 12(3), 595–612, 1997.

Wernli, H., Paulat, M., Hagen, M., and Frei, C.: SAL A Novel Quality Measure of the Verification of Quantitative Precipitation Forecasts, Mon. Weather Rev., 136, 4470–4487, 2008.