

Probabilistic prediction of raw and BMA calibrated AEMET-SREPS: the 24 of January 2009 extreme wind event in Catalunya

P. Escribà¹, A. Callado¹, D. Santos², C. Santos², J. A. García-Moya², and J. Simarro³

¹Delegation of AEMET in Catalunya, Barcelona, Spain

²AEMET Central Headquarters, Madrid, Spain

³Delegation of AEMET in Valencia, Valencia, Spain

Received: 26 February 2010 – Revised: 5 November 2010 – Accepted: 11 November 2010 – Published: 22 November 2010

Abstract. At 00:00 UTC of 24 January 2009 (24Jan09) an explosive cyclogenesis placed at the Gulf of Vizcaya, reached its maximum intensity with observed surface pressures below 970 hPa on its center. During its path through the south of France there were strong westerly and north-westerly winds over Iberian Peninsula (above 150 km/h). These extreme winds leaved 8 casualties in Catalunya, the north-east region of Spain.

The aim of this work is validating the skill of the Spanish Meteorological Agency (AEMET) Short Range Ensemble Prediction System (SREPS) in forecasting this event. Two probabilistic forecasts of wind are compared, a non-calibrated (or raw) and a calibrated one using the Bayesian Model Averaging (BMA).

AEMET runs a daily experimental SREPS twice a day (00:00 and 12:00 UTC). This system consists on 25 members that are constructed by integrating five different Limited Area Models (LAMs) at 0.25 degrees of horizontal resolution. Each model uses five different initial and boundary conditions from five Global Models (GMs). Thus it is obtained a probabilistic forecast that takes into account initial, contour and model uncertainties.

BMA is a statistical tool for combining predictive Probability Distribution Functions (PDFs) from different sources. BMA predictive PDF is a weighted average of PDFs centered on the individual bias-corrected forecasts. Each weight is a measure of the corresponding forecast skill. Here BMA is applied to calibrate probabilistic forecasts of wind speed.

In this work two time forecast ranges (H+60 and H+36) of 10-m wind speed over Catalonia are verified subjectively at 12:00 UTC of 24Jan09 valid time. We focus on the location

and intensity of 10-m wind speed maximum values. Observations at 29 automatic ground stations of AEMET are used for the verification.

On one hand results indicate that raw AEMET-SREPS is able to forecast 60 h ahead mean winds higher than 36 and 54 km/h and that it correctly locates them in three different areas. On the other hand, predicted probability loses its skill after BMA calibration of the ensemble. This is due to the fact that BMA bias correction underestimates the intensity of wind.

1 Introduction

Mediterranean region has a meteorological behaviour dominated by the interaction of synoptic flow with orographic small scale features and Mediterranean sea. Such interaction produces mesoscale structures that are difficult to model using GMs or even LAMs.

In order to improve the forecast of these structures AEMET proposed some years ago to create a SREPS. The main goal of ensemble prediction respect to deterministic one is that it takes into account forecast uncertainty by means of a probabilistic prediction. Using this technique, another valuable aspect is that PDFs are constructed dynamically, not statistically.

There are several ways of making a SREPS. Our ensemble is based on the multi-model multi-analysis and multi-boundary conditions technique. On the other hand, as numerical models have systematic errors, it is desirable to calibrate the ensemble. In this work we validate the skill of AEMET-SREPS when forecasting the 24Jan09 wind event. We compare probabilistic prediction of raw ensemble with BMA calibrated prediction.



Correspondence to: P. Escribà
(pescrbaa@aemet.es)

Table 1. LAM (files), GM (columns) and the 25 members of AEMET-SREPS.

	ECMWF	GME	AVN	CMC	UKMO
HIRLAM	IEC	IGM	IAV	ICM	IUK
UM	UEC	UGM	UAV	UCM	UUK
LM	LEC	LGM	LAV	LCM	LUK
HRM	HEC	HGM	HAV	HCM	HUK
MM5	MEC	MGM	MAV	MCM	MUK

2 AEMET-SREPS, BMA and the probabilistic forecast of wind

2.1 AEMET-SREPS

AEMET-SREPS (García-Moya et al., 2009) is a multi-model, multi-analysis and multi-boundary conditions system. Uncertainties of model, analysis and boundary conditions are sampled by construction. The ensemble has 25 members that are generated combining five LAMs with five GMs (Table 1). Each one of these numerical models are daily integrated and verified at different National Weather Services all over the world, so they are *state-of-the-art* Numerical Weather Prediction (NWP) models.

Hou et al. (2001) in SAMEX and Palmer et al. (2003) in DEMETER experiments demonstrated that this ensemble technique gives more skill in probabilistic forecasting than any other one. This fact was supported by one of the conclusions of the last workshop on SREPS hold at Exeter on June 2009: “Multi-model ensembles give the best results in all temporal and space scales” (4th Workshop, 2009).

AEMET-SREPS runs twice a day (00:00 and 12:00 UTC) at AEMET Central Headquarters. GM boundary conditions are received operationally and then five LAMs are integrated locally at our CrayX1e vectorial supercomputer. Horizontal resolution of members is 0.25 degrees of longitude and latitude and they have 40 vertical levels. Figure 1 shows the integration domains of each LAM, and in blue the common one. Forecast range is 3 days (H+72).

2.2 Bayesian Model Averaging (BMA)

BMA is a bayesian statistical technique for ensemble calibration (Raftery, 2003). We mean calibration as the correction of predictive PDF in order to adjust it to the actual (and unknown) forecast uncertainty.

$$\text{PDF}(y) = \sum_{K=1}^{K=25} w_K \text{PDF}_K(y | \bar{f}_K, \sigma_K) \quad (1)$$

The basic point of BMA is making a supposition of the shape of the actual PDF of forecast (1). Having the shape, parameters w_K and σ_K are estimated by maximum likelihood comparing predictions with observational data in a training period previous to the forecast initial time. PDF_K is the PDF

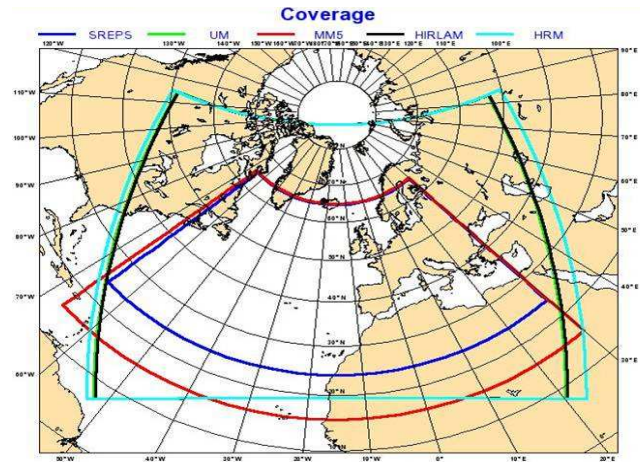


Fig. 1. Integration domains of AEMET-SREPS LAMs (LM has the same domain as HRM).

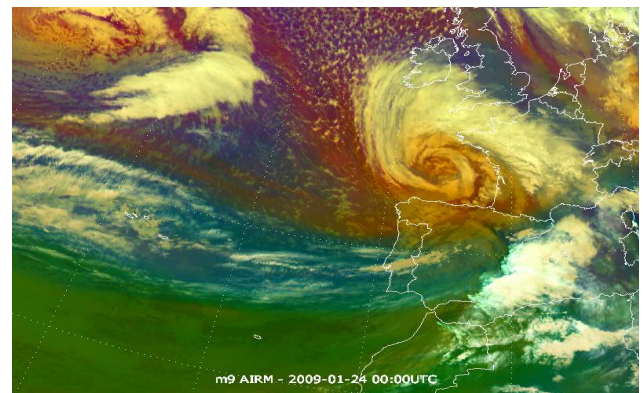


Fig. 2. RGB composition of Meteosat-9 at 00:00 UTC of 24 January 2009. Here RGB is the composition of channels (μm) WV6.2-WV7.3 (Red), IR9.7-IR10.8 (Green) and WV6.2 (Blue).

of ensemble member K and it is centred in the corresponding bias corrected deterministic forecast \bar{f}_K . It represents the uncertainty of the member and its shape depends on the meteorological parameter (Gaussian for temperature, Gamma for precipitation, etc.). w_K is the weight of each member and a measure of its skill.

2.3 Wind speed calibration with BMA

Wind speed is a definite positive meteorological variable. It means that its PDF maximum is not at zero. According to literature Weibull PDF is the one that approximates it better. In this work we use a Gaussian PDF for wind speed because this is the available one we have in the BMA computation software. Here we point out that this approximation can introduce errors in the calibration.

Previous experiments have been done with variable temperature in order to find the proper BMA training period. A

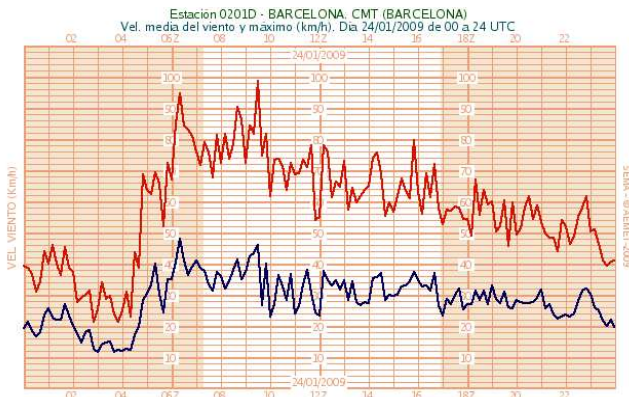


Fig. 3. 10-minutal wind mean speed (blue) and wind gust (red) registered at the automatic ground station of the Delegation of AEMET in Catalunya, Barcelona. Day 24 of January 2009.

short training benefits a rapid adaptation to weather changes whereas a long training ensures a better estimation of parameters. In this case we have proved that a 3-days training gives similar results to a 20-days one. Here we take the work for temperature as a reference for wind speed calibration, because both variables are generated with the same NWP system. In order to save computational time we choose the 3 previous days to the event as the training period.

3 Case of study: the extreme wind of 24Jan09 in Catalunya

3.1 Description of the event

Figure 2 is the Airmass RGB image of the Klaus extratropical cyclone placed at Gulf of Vizcaya at 00:00 UTC of 24Jan09. The red colour of the image which encloses the cyclone indicates descending stratospheric air related with a jet stream. The white and whitish colours indicate high and mid-level clouds rotating around the centre of the low pressure. In the south part of Spain the green colour is related with warmer air. This system was a deep low pressure that evolved rapidly from Atlantic Ocean to central Europe. As it passed through the south of France it caused very strong westerly and north-westerly wind gusts of more than 150 km/h at some points in Catalunya. In addition to multiple damages, 8 people dead.

In Fig. 3 is represented the 10-min average wind speed of 24Jan09 registered at the Delegation of AEMET in Catalunya, Barcelona. According to the graphic, episode started at 06:00 UTC reaching 95 km/h of wind gust (maximum wind) and 50 km/h of mean wind. At 12:00 UTC wind gust was 80 km/h and mean wind 40 km/h. It decreased steadily down to 40 km/h of wind gust and 20 km/h of mean wind at 00:00 UTC of 25 January.

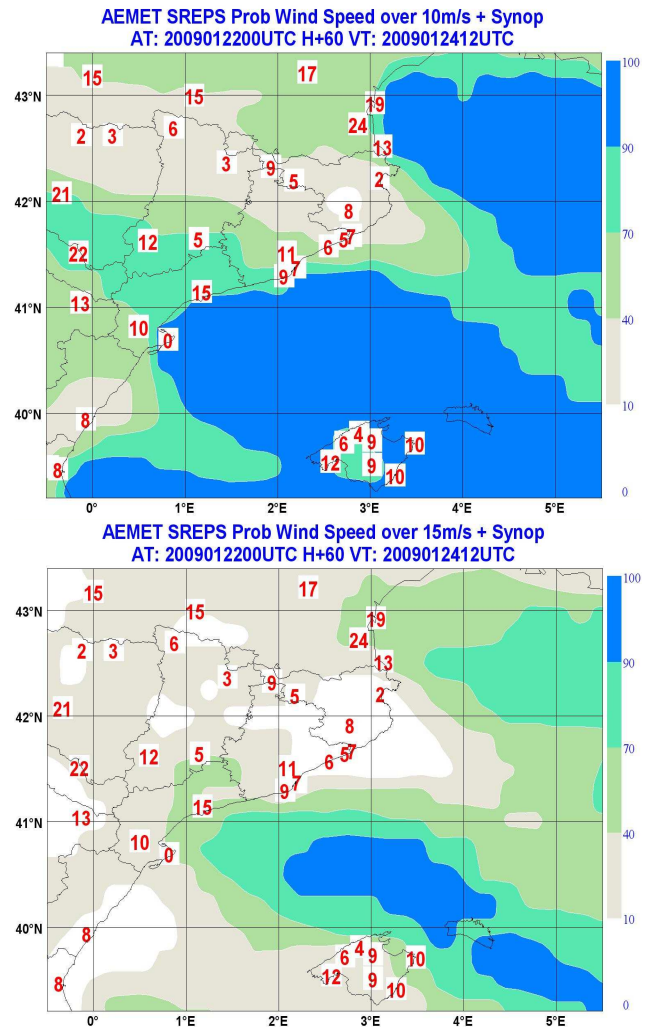


Fig. 4. H+60 raw AEMET-SREPS. Probabilistic prediction of mean wind speed greater than 10 m/s (top) and 15 m/s (bottom). Observations are plotted in red (m/s).

3.2 Questions to answer

With this work we try to answer two questions: (a) Did AEMET-SREPS hit prediction of the 24Jan09 extreme wind event? And (b) did BMA improve probabilistic forecast?

3.3 Methodology

The study consists on carrying out a subjective verification of probabilistic forecasting of 10-m wind speed greater than four thresholds, 10, 15, 20 and 25 m/s (36, 54, 72 and 90 km/h respectively). We follow the rule that a predicted event is true if its probability is greater than 50%. We look at skill of SREPS to forecast spatial location and intensity of wind. The instant of study is 12:00 UTC of 24Jan09 and two forecast ranges are taken into account, H+36 and H+60. As

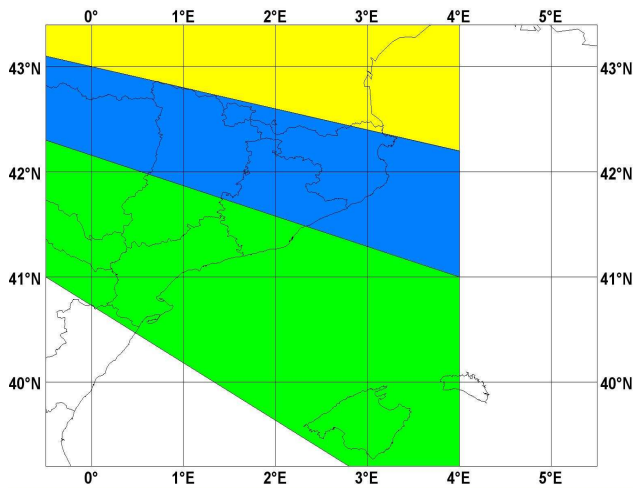


Fig. 5. Areas of different affection of wind. By convention we name them Tarragona (green), Pirineus (blue) and Perpinyà (yellow).

observational data we take mean 10-m wind speed at 29 automatic ground stations in Catalunya. This is firstly studied for raw ensemble and secondly for the BMA calibrated one.

Although verification is done for mean wind speed (this is the variable explicitly resolved by NWP models) results can be extrapolated to wind gust as far as both variables may have a quasi linear relationship (see Fig. 3).

4 Results

4.1 Did AEMET-SREPS hit prediction of the 24Jan09 extreme wind event?

Yes. Figure 4 shows H+60 predicted probability of wind speed greater than 10 m/s (left) and 15 m/s (right), for raw ensemble. Observations are plotted in red. Looking at observation wind a general behaviour can be inferred considering the map in three affected areas. These three areas are plotted in Fig. 5. We name them Tarragona, Pirineus and Perpinyà. These three areas show strong, moderate and again strong wind speeds, respectively. This geographical distribution of observed winds is due to the existence of the Pirineus mountains (with altitudes higher than 3000 m). This chain of mountains act as a wedge that splits the north-westerly winds in two parts, one to the north and the other to the south of the mountains.

Here SREPS defines very well these three zones and predicts quite well the intensity of wind for the 10 m/s threshold. For the 15 m/s threshold SREPS hits the right intensity at almost all stations when following the rule of 50%. It has to be noticed that the skill of the system in forecasting very strong winds in the coasts of Tarragona and Perpinyà, 60 h before the event, is quite remarkable.

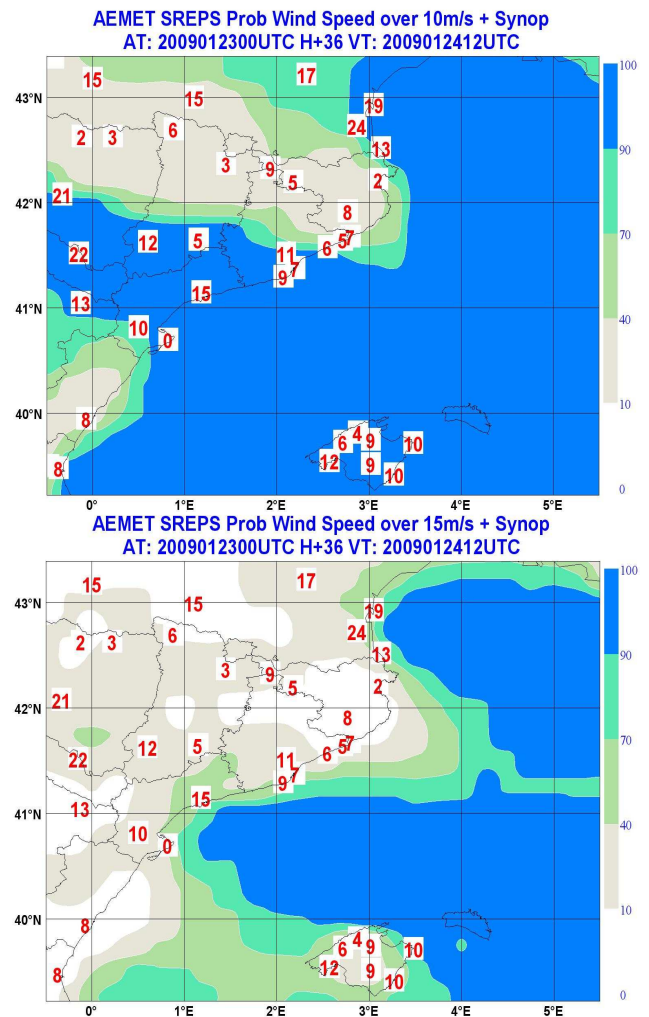


Fig. 6. H+36 raw AEMET-SREPS. Probabilistic prediction of mean wind speed greater than 10 m/s (top) and 15 m/s (bottom). Observations are plotted in red (m/s).

In Fig. 6 is represented the same as in Fig. 4 but for H+36 forecast range. As expected, forecast improves with respect to H+60 in the whole domain. Now the three areas (Fig. 5) are completely defined, in particular probability of mean wind speed greater than 10 m/s at Tarragona is greater than 90%.

20 m/s and 25 m/s thresholds for H+36 predictions are plotted in Fig. 7. For 20 m/s threshold the skill is lost at the three stations that have observed values greater than 20 m/s. This result is not rare and it mainly responds to the fact that NWP models underpredict wind speed when the observed wind is strong or very strong. This affirmation will be analyzed in the next section. Verifying 25 m/s threshold is not possible because there is not any station measuring values greater than this.

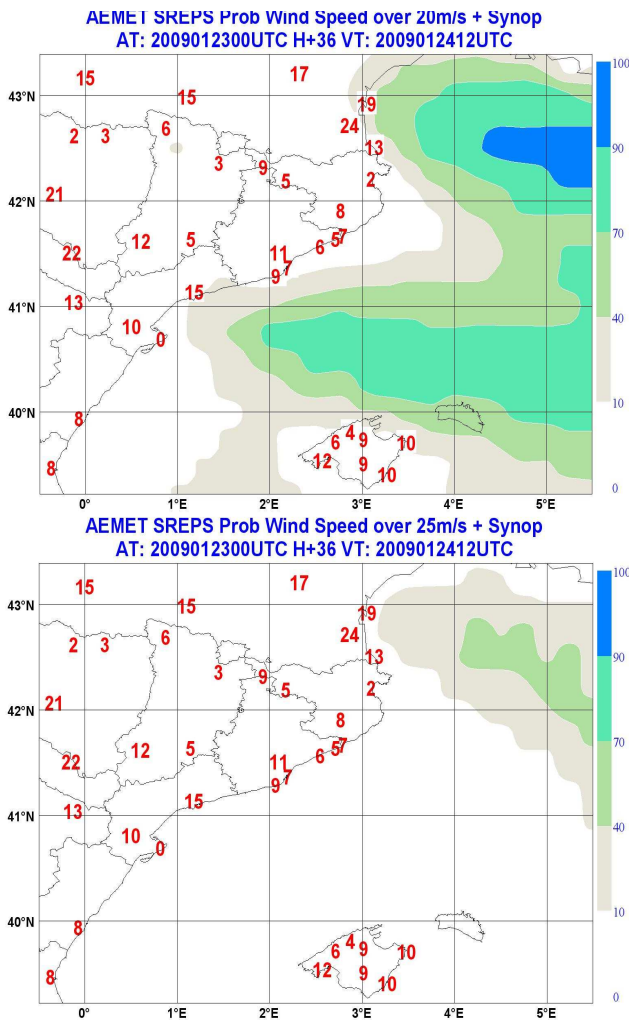


Fig. 7. H+36 raw AEMET-SREPS. Probabilistic prediction of mean wind speed greater than 20 m/s (top) and 25 m/s (bottom). Observations are plotted in red (m/s).

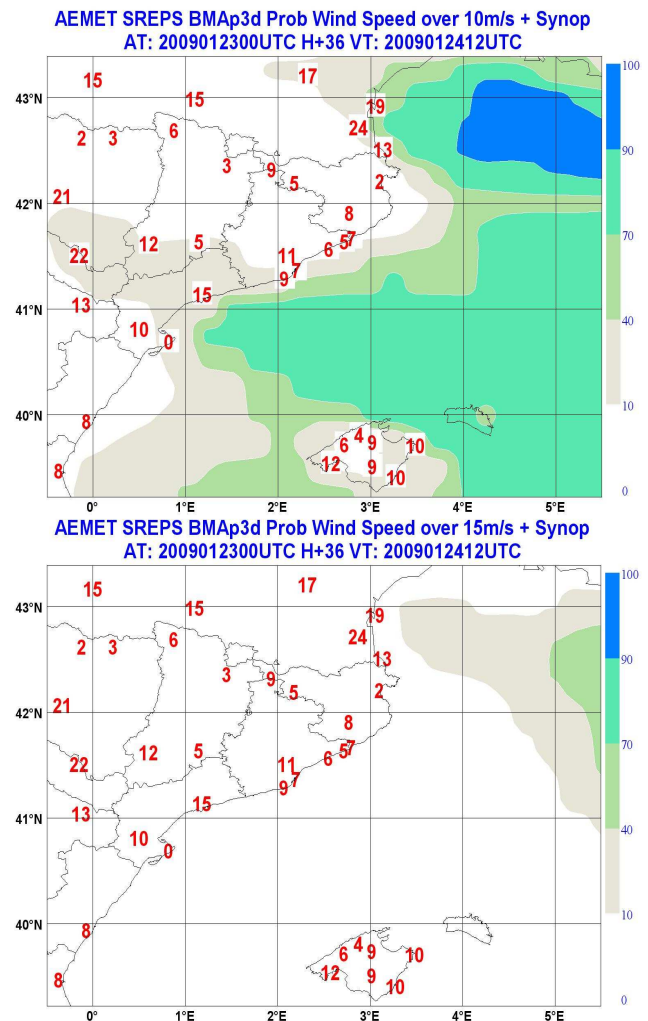


Fig. 8. H+36 AEMET-SREPS calibrated with BMA. Probabilistic prediction of mean wind speed greater than 10 m/s (top) and 15 m/s (bottom). Observations are plotted in red (m/s).

4.2 Did BMA improve probabilistic forecast?

No, in fact here BMA makes the forecast worse. As it is shown in Fig. 8 probabilistic forecast is significantly worse for H+36 with respect to raw ensemble (Fig. 6). There is a general tendency of underestimating wind speed.

In order to find the reason of this result we take a look at Eq. (1). \bar{f}_K is the bias-corrected forecast of the ensemble member K . In our case we apply a simple linear correction of bias, that is,

$$\bar{f}_K = a + bf_K \quad (2)$$

Here a and b are bias coefficients and f_K the raw forecast.

Figure 9 shows the magnitude of this correction when computing bias using the 3 previous days as training period. As it can be seen, for high values of predicted wind (above 4 m/s), bias correction reduces wind values. This correction

does not correspond to the actual behaviour of NWP models on predicting wind. In Fig. 10 are represented 568 pairs of observed and predicted values of wind speed using IEC member of SREPS (Table 1), corresponding to the training period. Looking at the figure this member tends to overpredict wind speed for low values whereas it underpredicts wind when it is strong. It can be demonstrated that this behaviour is common for NWP models. So that BMA bias correction of 24Jan09 extreme wind event calculated from the previous 3 days eventually reduces wind intensity, giving an erroneous forecast.

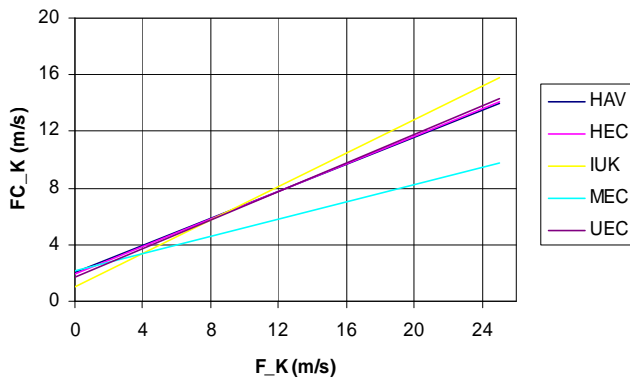


Fig. 9. Linear correction of bias per member. FC_K is the bias-corrected forecast, \hat{f}_k in (1), and F_K is the raw forecast, f_k in (2). Only the members that have a weight significantly different from 0 are represented.

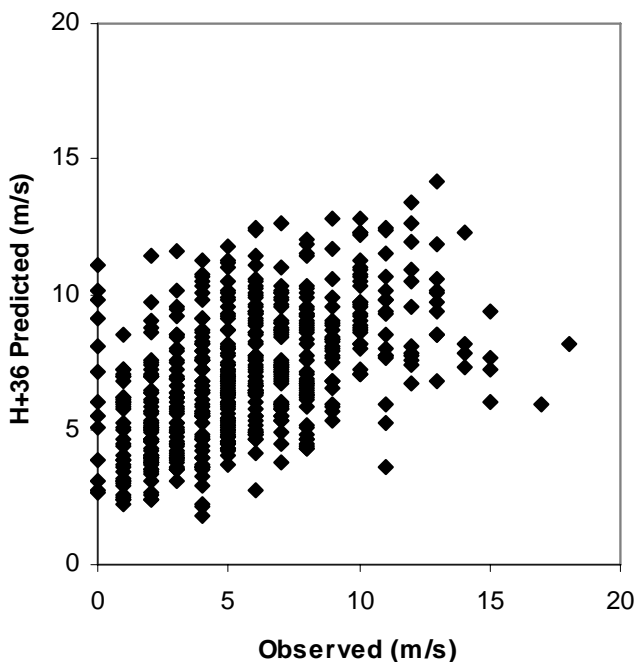


Fig. 10. Scatter plot of IEC (Table 1) H+36 predictions of wind speed against observations. 568 realizations during the 21–23 January 2009 training period.

5 Conclusions

In this work we have shown the skill of AEMET-SREPS predicting the 24Jan09 extreme wind event in the local area of Catalunya. 60 hours before the event, raw AEMET-SREPS is able to define very well the three main affected areas, say Tarragona, Pirineus and Perpinyà (Fig. 5). It is of particular interest the skill of the system in forecasting very strong winds in the coasts of Tarragona and Perpinyà. A probability of wind speed greater than 10 m/s at Tarragona is larger than

90% 36 hours before the event.

When we apply the standard formulation of BMA to this high intensity wind speed event, probabilistic forecast is bad calibrated. This problem is mainly due to the fact that strong winds are rare events and they have a high variability in space and time. Then training periods which don't have strong winds may cause that computed bias cannot be applied to calibrate probabilistic forecasts of these events. Thus we conclude that in the case of strong winds no calibration of predictive PDF is desirable.

This work supports the decision taken by AEMET to launch a SREPS based on the multi-model multi-analysis and multi-boundary conditions technique. As the ensemble system approximates the daily dynamics of the real atmosphere (It is constructed combining several NWP models) it can be inferred that these results can also be found in other similar situations. Despite its high cost in human and computational resources AEMET-SREPS seems to be a powerful tool to make skilful probabilistic forecasts in the mesoscale and the short range.

Acknowledgements. The authors would like to especially thank the National Weather Services of Germany (DWD), United Kingdom (UKMO), Canada (SMC) and United States (NCEP) and the ECMWF for daily providing us the data of their global operational forecasting models. We also thank Adrian Raftery and his group for developing and sharing BMA software for ensemble calibration. Finally we thank Raquel Sanclaudio for helping us to plot graphics with Metview.

Edited by: A. Mugnai

Reviewed by: one anonymous reviewer

References

- 4th Workshop on Short Range Ensemble Prediction Systems, Exeter, UK, online available at: http://www.metoffice.gov.uk/conference/srnwp_workshop, June 2009.
- García-Moya, J. A., Callado, A., Santos, C., Santos-Muñoz D., and Simarro, J.: Predictability of Short-Range Forecasting: A multi-model approach", AEMET Technical Memorandum, NT SPPE-1, 30 pp., 2009.
- Hou, D. and Kalnay, E.: Objective verification of the SAMEX98 ensemble forecasts, *Mon. Weather Rev.*, 129, 73–91, 2001.
- Palmer, T. N., Alessandri, A., Andersen, U., Cantelaube, P., Davey, M., Délecluse, P., Déqué, M., Díez, E., Doblas-Reyes, F. J., Feddersen, H., Graham, R., Gualdi, S., Guérémy, J.-F., Hagedorn, R., Hoshen, M., Keenlyside, N., Latif, M., Lazar, A., Maisonave, E., Marletto, V., Morse, A. P., Orfila, B., Rogel, P., Terres, J.-M., and Thomson, M. C.: Development of a European multi-model ensemble system for seasonal to inter-annual prediction (DEMETER), *B. Am. Meteorol. Soc.*, 85, 853–872, 004.
- Raftery, E.: Using a Bayesian Model Averaging to Calibrate Forecasts Ensembles. Technical Report no. 40, Department of Statistics, University of Washington, USA, 2003.