ECMWF forecasts: a tale about signal, noise, error and value

François Lalaurette, ECMWF

The quest for perfect forecasts

Improving the forecasts through improved data assimilation, numerics and physics is the aim of all Numerical Weather Prediction (NWP) centres – ECMWF is no exception! Impressive progress has been achieved in recent years as can be seen from the classical measures of performance, such as the anomaly correlation and root mean square errors of forecasts of the topography of isobaric surfaces [1]. Although all aspects of model development did contribute to this remarkable improvement, new data assimilation techniques (4D-Var) making it possible to extract useful information from a wide range of satellite instruments have certainly played a major role. Thanks to a better control of errors in relatively small scales, it has become possible to get a clear forecast benefit out of increased resolution - something that could not be taken for granted in the early 1990s [2]

The impact of model time and space resolution on the forecast of rapid cyclogenesis has recently been shown to be larger than expected (Figure 1). Such cases illustrate the intimate relation between the physics and the dynamics that lead to a successful forecast. These features that have to be kept in the model version used in the Ensemble Prediction System (EPS) even though the time and space resolution have to be degraded for the system to remain affordable.

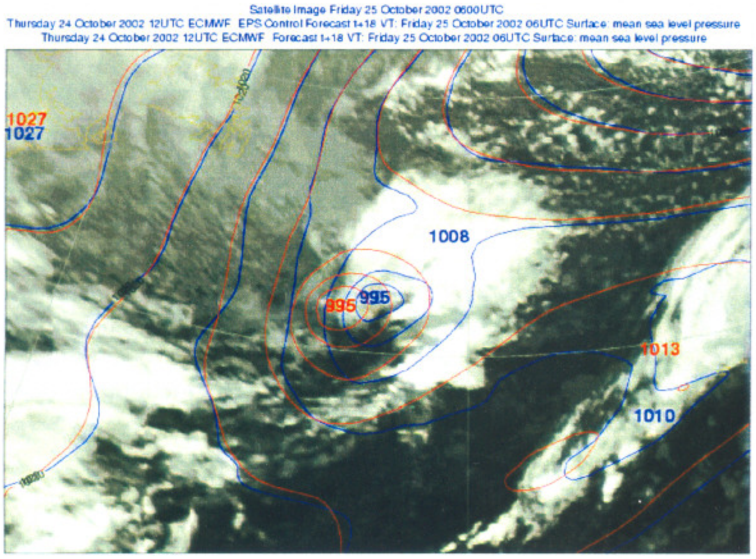

Fig. 1 Early stage of development of the storm that hit England, Netherlands and Germany on 27–28 October 2002. The 18h forecast for sea-level pressure from the T511 (blue) and T255 (red) versions of the model are superimposed on Meteosat-7 infrared imagery from 25/10/2003 06 UTC

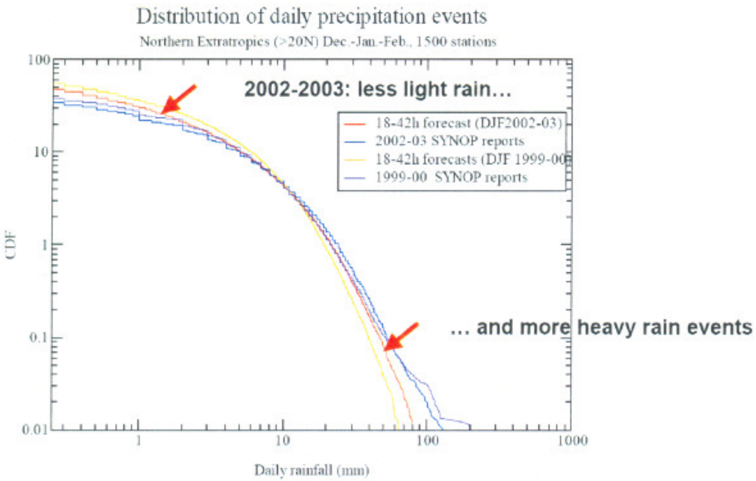

Fig. 2 Distribution of winter daily rainfall events in the model (orange: 1999-2000; red: 2002–2003) compared to SYNOP reports (blue and magenta)

Figure 2 demonstrates that the current model physics are increasingly realistic in their representation of processes directly related to sometimes-severe weather conditions such as heavy rainfall. Although the limited resolution does not allow the representation of the most extreme events in excess of 100 mm/day, both the changes in the physics and increases in resolution have brought the model distribution significantly closer to the observed ones over the last five years.

Error and value

Although objective measures of forecast errors clearly show an improvement in our forecasts, including those for high-impact weather events, this improvement has to be brought to the user in the form of a *useful* forecast. This means that for any given application, the ultimate measure of how good a forecast was is not how close it was to observations, but whether or not it can be used for making good decisions. Models providing an evaluation of the impact on decision making of NWP forecasts have been proposed [3]. They show a clear benefit from using the forecast in a probabilistic format, as it allows tailoring the decision-making for each application provided that a good statistical knowledge of model errors is provided.

Signal extraction for meteorological forecasts has little to do with forecast validation. Statistical techniques can usefully extract value from a range of parameters that may only have an indirect relation with the meteorological event to which a given application is sensitive. Since ensemble forecasts, even when formulated in a probabilistic way, are prone to errors, it may be better for any given application to be fed with ensemble values and error statistics rather than to try and calibrate the probability distribution separately. Therefore the evaluation of EPS forecasts broadly follows two main streams:

- the first aims at checking that the ensemble products provide unbiased probability distributions (reliability)
- the second aims at providing the users with a view whether or not the ensemble helps in making useful decisions (resolution)

The latter is only tentative of course, and should be seen only as a first guess to be further refined by the users for each application. In the former category are measures of the spread/skill relationship. Because EPSgrams [4] have become a popular way to represent the EPS forecast for any location, it may be of interest to check that the spread as represented by the size of the boxes (namely the interquartile distance) gives an a-priori estimate of whether or not the forecast of the day is likely to be more or less accurate. No direct correlation between the EPS spread and the individual forecast errors should of course be expected - rather the spread is providing an estimate of the average level of error. If we define the spread as half the interquartile distance, the spread should match the median of the error distribution - and this should apply for any predefined range of spread values, provided that the EPS provides a large enough sample of cases to compute a robust estimate of the error distribution median. This verifies rather well and can be seen in Figure 3: the spread matches pretty well the different estimates of the median of wintertime, day 6 2m-temperatures errors, even though it is slightly underestimated. A more surprising result found recently when performing this type of verification. The spread of summer rain forecasts is usually an overestimate of the local error, when it is expected that limited resolution of the model should lead to an underestimate of errors [5]. New formulations of the stochastic physics formulation are expected to provide more realistic estimates of the spread for summer time precipitation.

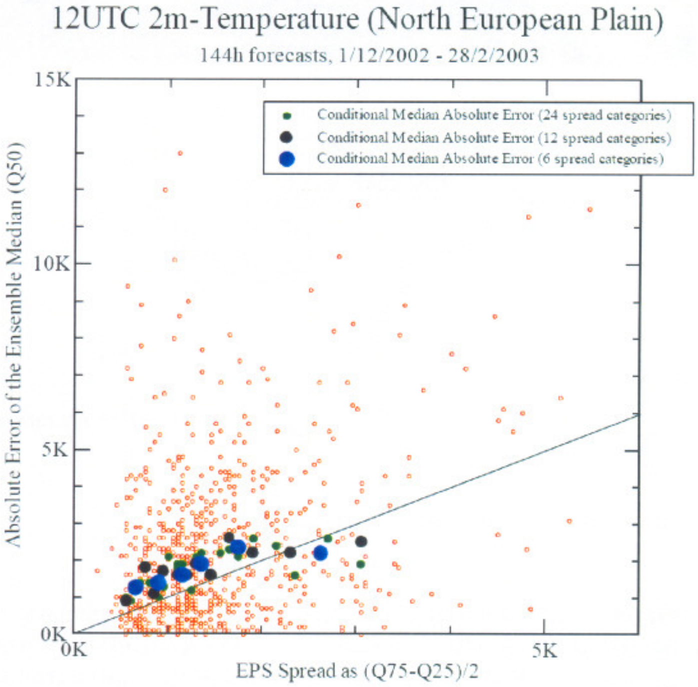

Fig. 3 EPSgram verification

Extreme weather

Although the statistical calibration of deterministic and ensemble forecasts is enhancing their value, statistical post-processing usually fails to improve the forecast near the extremes of the climatological distribution. The reason for this is simply because it would take too long to 'learn' of model errors in such cases, meaning that significant model changes occur on a time scale that is short compared to this learning period.

Rather than trying to address the problem of calibrating the model forecasts for extreme events, the approach developed at ECMWF over recent years has been to detect extreme events in the model climate rather than the real one. The Extreme Forecast Index [6] for example locates areas where the EPS forecast deviates most from the model climate, on a scale going from -1 (all members being low-value outliers of the climate) to +1 (all members being high value outliers). The normalisation is using climate distributions that are both space and time dependent. EFI maps therefore provide a convenient 'first guess' for early warnings on a continental scale. This allows to correct for some of the known model deficiencies such as the underestimation of wind gusts over land or the effect of limited resolution on strong orographically forced convective events, while using the probabilistic formulation of the forecast to allow detection of the event even if only a small proportion of members are really extreme forecasts (Figure 4)

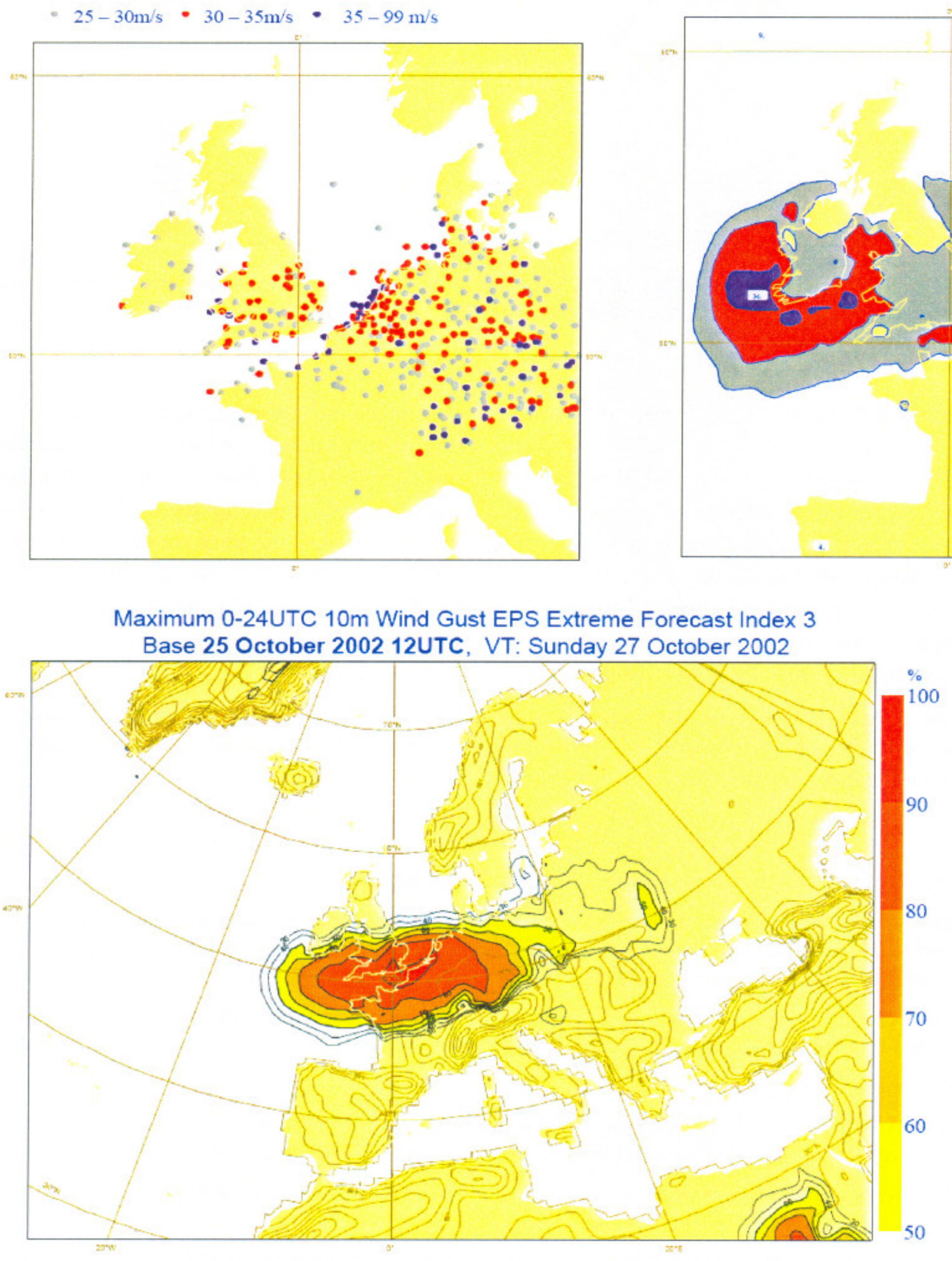

Fig. 4 27 October 2002 storm (same event as in Figure 1): Clockwise from top left 1) Maximum wind gusts reported on that day; 2) T511 forecast from 26/10 12 UTC; 3) EFI from 25/10 12 UTC.

Another example of probabilistic product helping at the early forecasting of extreme events is the Tropical Cyclone Strike Probability map (Figure 5). Although the deterministic forecast of tropical cyclones tracks has improved in recent years [7], such probabilistic products are likely to improve the planning and early decision making for actions that might have to be taken at a later stage such as moving populations or restoring public services.

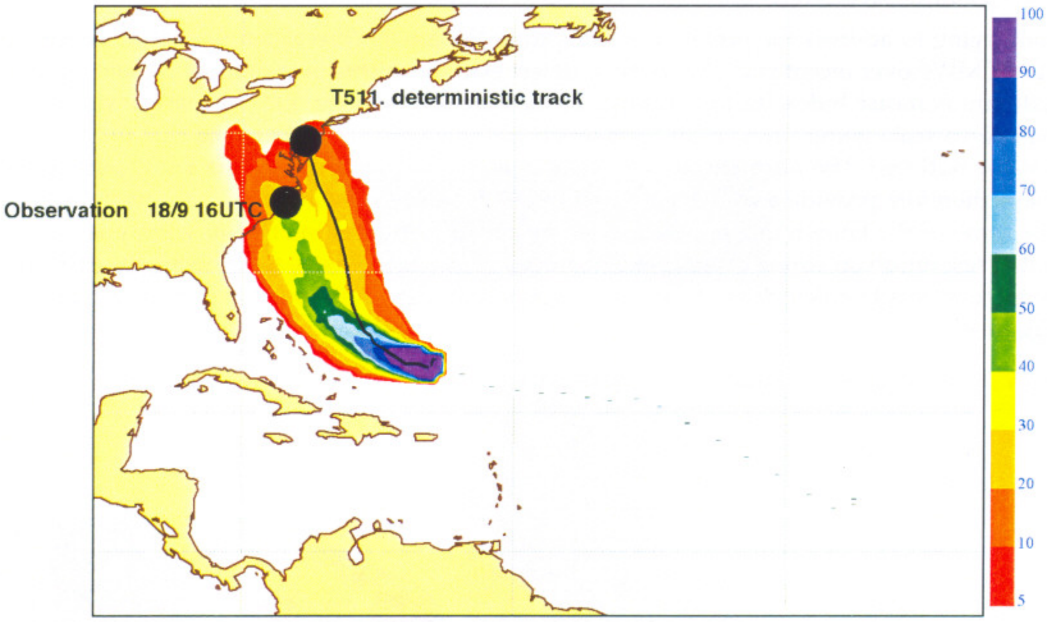

Fig. 5 TC Isabel: both the T511, deterministic forecast track and the EPS probabilities that the TC track will come closer than 120 km within the next five days are shown (Forecast base time 14/9/2003 12 UTC).

Conclusion

The improvement in the quality of NWP forecasts in recent years has led to more and more users taking some of their critical decisions with the help and on the on the basis of meteorological forecasts. It is however a difficult task to convince these users that forecast errors are not all related to human failures, but may also be due to the intrinsic limitations of atmospheric predictability. Also, even if the trend towards pushing the limits of predictability further into the future has proved very successful over the past 5 to 10 years, the pace remains too slow for users to wait for much improved forecasts to be available before they start using them. Therefore moving from the 'perfect forecast' dream to a more pragmatic approach where forecast errors are fully accounted for is likely to increase the value of the forecast. Also the ensemble technique, by bringing some information on the flow dependent characteristics of these stochastic errors, helps further increase the forecast value.

The next challenge for global models is to extend the useful forecast range of extreme weather conditions. Some first steps have been demonstrated here (EFI, Tropical Cyclones maps) but a lot remains to be done to assess the value of such forecasts in operational environments.

References

[1] www.ecmwf.int/products/forecasts/d/charts/verification/timeseries/monthly_mean

[2] **Simmons, A.J.**, 2003: Observations, assimilation and the improvement of global weather prediction - some results from operational forecasting and ERA-40, ECMWF Seminar on Recent developments in data assimilation for atmosphere and ocean. www.ecmwf.int/publications/library/ecpublications/_pdf/sem2003_simmons.pdf

[3] **Richardson, D.S.**, 2002: Predictability and economic value. ECMWF Seminar on Predictability of weather and climate. www.ecmwf.int/publications/library/ecpublications/_pdf/sem02_richardson.pdf

[4] www.ecmwf.int/products/forecasts/epsgrams/

[5] **Lalaurette, F., L. Ferranti, A. Ghelli,** and **G. van der Grijn**, 2003: Verification statistics and evaluations of ECMWF forecasts in 2002-2003, *ECMWF Tech. Memo*, **432**. www.ecmwf.int/publications/library/ecpublications/_pdf/tm432.pdf

[6] **Lalaurette, F.**, 2003: Early detection of abnormal weather conditions using a probabilistic extreme forecast index, *Q.J.R. Meteorol. Soc.*, 3037-3057

[7] **van der Grijn, G.**, 2002: Tropical cyclone forecasting at ECMWF: new products and validation. *ECMWF Tech. Memo*, **386**. www.ecmwf.int/publications/library/ecpublications/_pdf/tm386.pdf