Ensemble Forecasts: Can They Provide Useful Early Warnings?

François Lalaurette and Gerald van der Grijn

European Centre for Medium-range Forecasts, Reading, Berks. U.-K.

1. Medium-range ensemble forecasts: why, how - and how well?

The design of ensemble methods for medium range forecasting ten years ago has been aiming mainly at addressing the problem of limited predictability of supra-synoptic weather regimes in the 6-10 days range. From this point of view, it has been shown since then that the ensembles have delivered better forecasts compared to a purely deterministic approach, both improving single-value estimates by removing stochastic errors (ensemble mean) and providing reliable and sharp estimates of the probability distributions of large scale flow patterns. Fig. 1 gives the scores from different NWP models and shows that after the first three days, the removal of small scale, realistic but unpredictable patterns by the use of the ensemble mean filter is beneficial to the forecast skill. Fig. 2 shows that the benefit of the dynamical ensemble is more than a mere filter applied to the high resolution forecast: even if only the projection of the forecast onto large scale weather regimes is looked after, the probability of occurrence of those is more accurately predicted by the EPS than it is by the deterministic (control) forecast, or even by the ensemble of lagged deterministic forecasts valid for the same date. This allows the probabilistic forecast of weather regimes to be more skilful than a forecast of the climatological frequency of occurrence of the regime all along the currently operational 10 days forecast range, when a purely deterministic forecast would loose its skill by day 7.

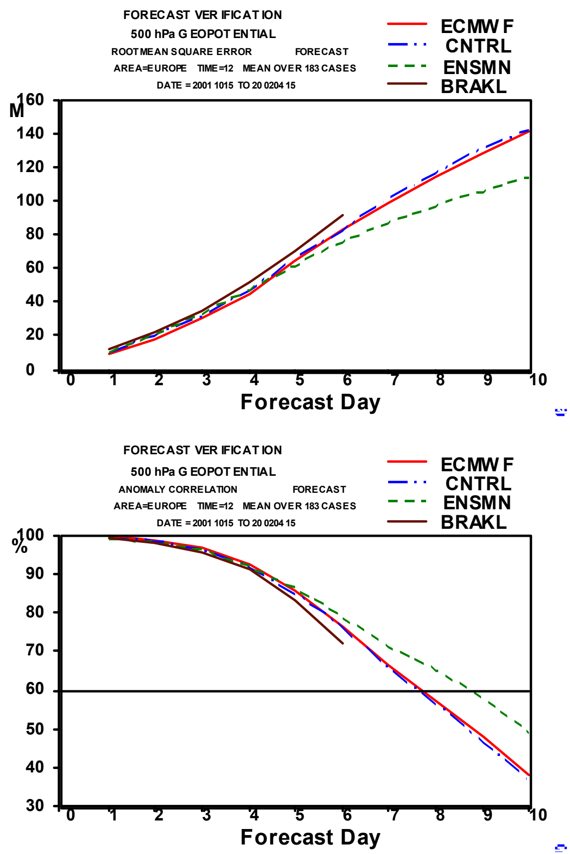

Figure 1 NWP scores (500 hPa height) over Europe (15 Apr-15 Oct 2002); upper: Root Mean Square Error; lower: Anomaly Correlation; ECMWF T511 is full red, EPS Control T255 is long dashed blue, EPS ensemble mean is dashed green and Bracknell is full brown up to D6.

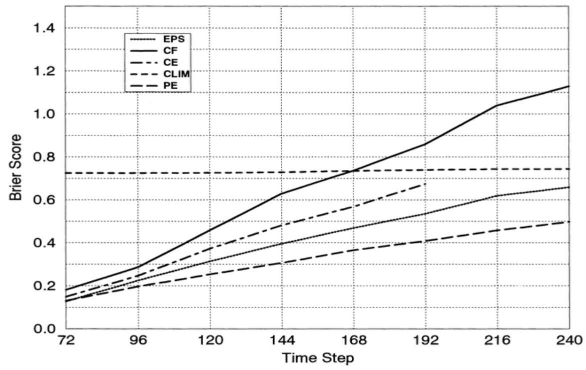

Figure 2 Brier Score (mean square probability errors) of forecasts for North-Atlantic weather regimes: from upper to lower: (deterministic) control forecast, ensemble of lagged control forecasts, ensemble of perturbed forecasts (EPS) and Perfect Ensemble (verification randomly sampled from the ensemble); the horizontal line is the score reached by the probabilistic forecast based on the climate frequency of occurrence of each regime (From Chessa and Lalaurette, 2001).

To reduce the space of variables generated by the Ensemble Prediction Systems to a projection onto a limited set of weather regime patterns, although useful to demonstrate the level of predictability achieved by the system up to ten days ahead, is however a very restrictive approach, leaving it a long way to demonstrate the potential of early warning for severe weather based on an ensemble of global forecast.

Indeed the experience in trying to use medium range forecasting systems to be alerted of the risk of severe weather more than one day in advance is very limited. In most occasions, civil security services are alerted not earlier than the day before the event, while public warnings are only issued on the same day - Tropical Cyclones being a notable exception to this common rule. The reasons usually given by forecasters as to why they do not use numerical forecasts in the early medium range are mainly twofold:

1. the global numerical models are generating nothing looking remotely like severe weather

2. if one takes signatures from the global models that are associated to severe weather, there is no consistency of the forecast from one day to the next: the rate of false alarms would be far too high to be considered.

The current status of development of global models used for medium-range forecasting is such that the first item can be debated. ECMWF is running T511 (roughly 40km) with 60 vertical levels, a resolution that was considered not so long ago as suitable for mesoscale modelling. Running a global model has also many advantages compared to a limited area approach, as it is easier to keep the forecasts under control of a global observing system without having to take care of lateral boundary problems.

The Ensemble Prediction System currently runs at half resolution (T255L40) – the truncation on the vertical compared to the high resolution version is in the stratrosphere and mesosphere, with a top level at 10hPa instead of 0.1hPa in the T511L60 version. In addition to a Control run, 50 members are run with perturbations that include:

1. adiabatic singular vectors maximising total energy growth in the extratropics over the next 2 days, as well as those (evolved) vectors that did maximise error growth over the past two days;

2. diabatic singular vectors targeted to maximise total energy growth over areas where tropical cyclones have been reported[1] (Puri et al., 2001); those are scaled with a random amplitude sampling a multidimensional Gaussian distribution;

[1] In addition, an area 0°-25°N/100°0-60°W is permanently kept as a target in the Caribbean Sea..

3. stochastic perturbations to the tendencies diagnosed by the physical parametrisations (within +/- 50%, consistently on a vertical for all variables during 6 hours and over a 10° by 10° area (Buizza et al., 1999).

As an example of the skill the EPS now achieves in forecasting weather events such like rainfall, Fig. 3 shows the evolution of Brier Skill Scores[2] of the Probability of Precipitation (PoP) EPS forecast since 1995. Since March 2001, the EPS is run experimentally from 00 in addition to 12 UTC.

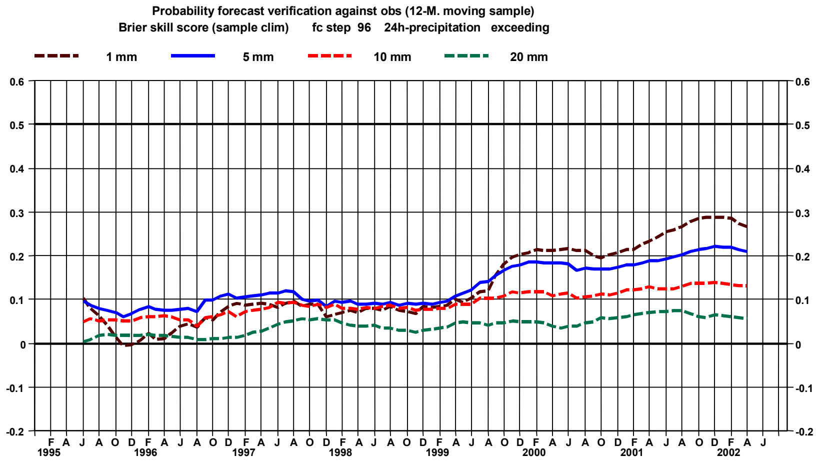

Figure 3 Evolution of the Brier Skill Score of EPS 96h forecasts verified against European SYNOP reports of daily rainfall over 12 consecutive months

The following sections will aim at providing evidence that not only do global NWP models nowadays generate severe weather signatures, but also that there is some signal to be detected up to a range of three to five days ahead using the improved signal detection capability offered by running an ensemble of forecasts - indeed that there may be hope that we can combine "Modelling and Predicting Extreme Weather Events " in the early medium range.

2. Severe weather in medium-range forecast models

The ability for the model to generate severe storms has improved in recent years, although the direct comparison from model wind speed over land with observations shows a pretty large, negative biais. Among the reasons why this occurs is that modellers have more in mind when designing boundary layer representations to have a good momentum budget than to optimise on-site validation of local effects. A step towards the post-processing of maximum wind gust values based both on explicit model winds and the subgrid scale representation of turbulent fluxes has however been taken in 2000 at ECMWF, resulting in a better adequacy between model and observations. An example of signatures now to be found for a typical, small-scale storm over Western Europe is given in Fig.4, where the good agreement between the short-range forecast and observations can be seen.

[2] Brier Skill Scores are the reduction in the rate of bad forecasts compared to climatology.

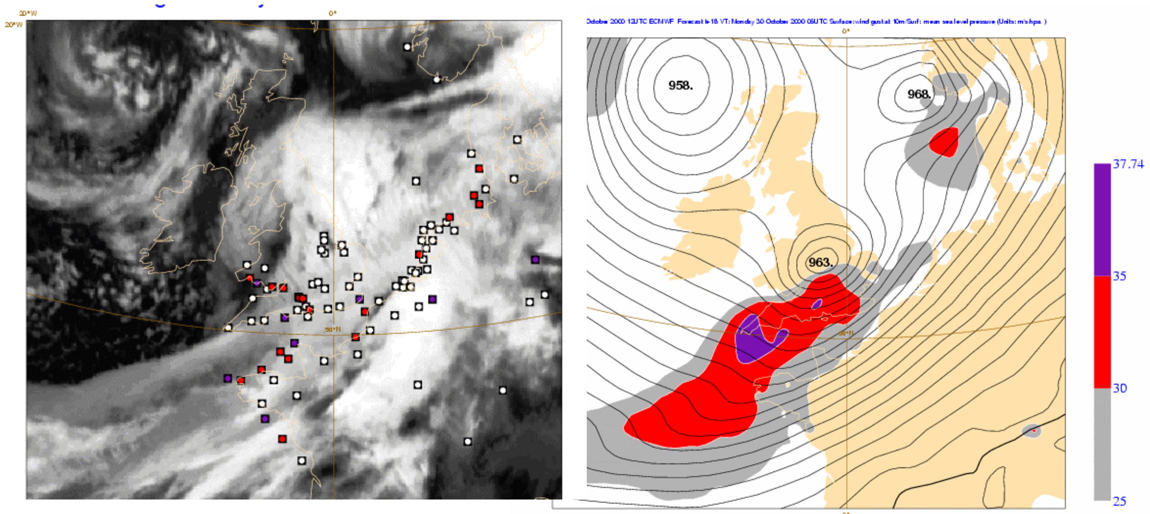

Figure 4 "Halloween" storm over the British Isles on 06UTC the 30 Oct. 2000; left: Meteosat7 Infrared picture+ SYNOP wind gust reports (grey>25m/s; red>30m/s; violet>35m/s); right: 29 Oct. 12UTC forecast for mean sea level pressure and wind gust

To identify the ability for models to generate severe weather signatures is only one side of the story of course, which tells nothing of the rate of success and failure with which the model generates them 3 or 4 days in advance. Severe weather scenarios are usually not seen by the models as the most likely scenarios at these ranges, but forecasters have expressed an interest in using even small probabilities of severe weather occurrence as a useful early warnings that will help them focus their monitoring of the situation when the severe weather eventually comes closer. The definition of thresholds on which to base such probabilities is however difficult, as it is very dependant on the local climate. In fact, it could be argued that in any populated area, indigenous populations have remarkably well adapted their activities to the local climate: what is considered a severe cold outbreak in Montreal or Cairo is not likely to be the same event, and the same could probably be said for severe storms in Reykjavik and Berlin. In order to generate maps were all locations are a priori equally likely to be hit by an unusual event, it has been proposed (Lalaurette, 2003) to map the severity of events with respect to the local climate distribution. In this way, and provided that we use the model to picture the local climate, orographic or land-sea map effects should be handled consistently in the climate and the model forecast, making the measure of departures more meaningful.

3. An Extreme Forecast Index

Although the models currently generate severe weather systems consistently throughout the medium-range, their forecast skill is quickly declining over the first few days. In the case of the Halloween storm (Fig. 4), this resulted in both an error in the location and the intensity of the storm (Fig.5) - although the model correctly predicted the large-scale rapid, perturbed flow that was to affect the area. Such errors in the smaller scales are to be expected for unstable systems where small errors in the analysis quickly amplify. It is therefore of interest to see whether the EPS is able to tackle these uncertainties in a probabilistic way.

The ensemble distribution of 90h-wind forecasts for Dover based on 12UTC the 26 Oct. can be found in Fig. 6. Although far from values that were observed (25m/s winds with gusts at 37m/s), some of the members did indicate that the situation was deviating from "normal". This is easier to see when the climate distribution based on EPS forecasts valid for this time of the year and this location are also reported, as is the case here. For example, although a value of 15m/s that is exceeded by 33 members out of 50 here would not be considered as a severe weather event in Dover, it is to be found only in slightly more than 1% of the EPS records at this time of the year and this location.

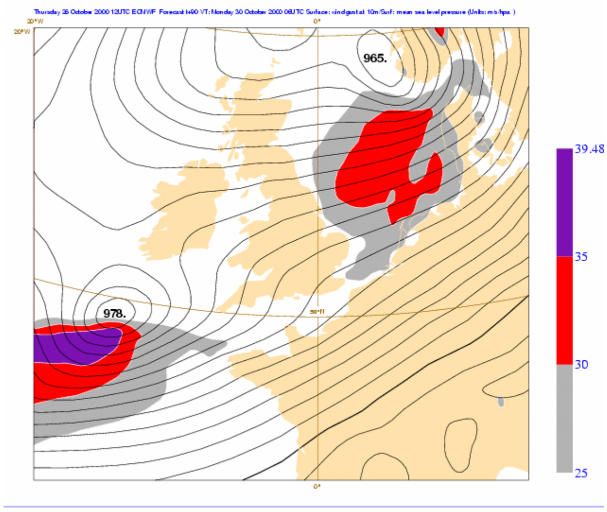

Figure 5 90h-forecast of the "Haloween" storm

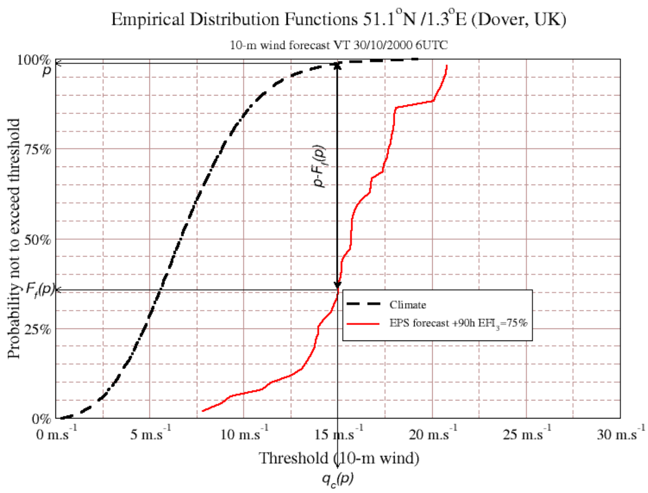

Figure 6 Climate (dashed) and EPS 90-h forecast (full) distributions of wind at Dover on 6UTC the 30 Oct. 2000 ("Halloween" storm).

The inspection of such local, empirical EPS distributions of parameters such as wind, temperature or rainfall is however not something that could easily become part of the routine work of any forecaster: the amount of information that would be needed before an assessment of the situation can be made would be far too large to be achievable in time for the forecast to be of any use. As has been the case each time a practical use of ensemble forecasts has been considered, some aggregation of the available information has to be made. Early attempts have mostly aimed at *clustering* large-scale scenarios on the basis of their similarity over subcontinental areas. Such an approach, although helpful to describe the large-scale evolution of the weather, is unlikely to provide a useful information for severe events, as scenarios that can look similar on the large scale can generate very different kinds of local events. It is rather proposed here to scale the EPS distributions with respect to the climate.

To each proportion p of the climate records is attached a value $q_c(p)$ known as a *quantile*, e.g. Wilks, 1995. How much a given EPS forecast deviates from the climate therefore can be measured by $p - F_f(p)$ (see Fig.

6) where $F_f(p)$ is the proportion of EPS members that lies below $q_c(p)$. These differences are then integrated in probability space to provide the *Extreme Forecast Index* at order *n* (Lalaurette, 2003)[3]:

$$EFI_n = (n+1)\int_0^1 \left(p - F_f(p)\right)^n dp$$

The result from such a procedure is that differences between any climate and forecast cumulative distributions functions (CDF) are scaled between -1 and +1 (-1 if the CDF in the forecast is entirely shifted below the minimum climate record, +1 if it is entirely shifted beyond the maximum). If the CDF from the forecast is the same as the climate one, then EFI=0.

An example of EFI3 field generated from 90h forecast EPS distributions based from 26 Oct. 12UTC is in Fig 7. It shows that although the deterministic forecast was misplacing the event, enough EPS members did have it at the right location for the EFI to reach high values where the event happened (see Fig. 4). Indeed the EFI3 value at Dover reached +75%, clearly indicating a well-above normal level of risk for gale force winds four days in advance.

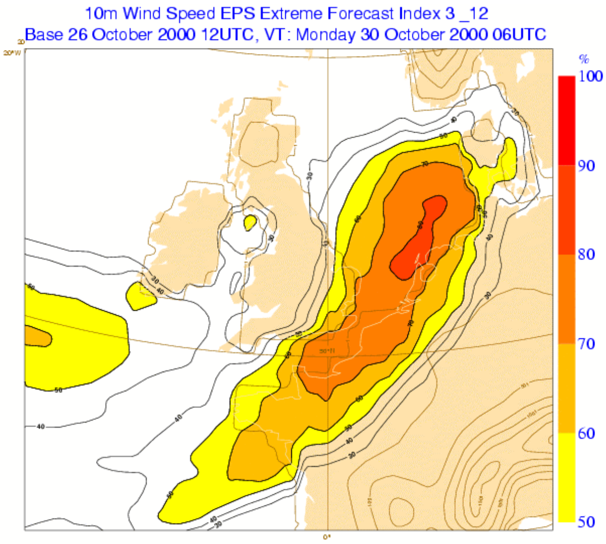

Figure 7 Extreme Forecast Index valid for the "Haloween" storm

One of the benefits of using an index that scales departures of the forecast distribution from the climate globally rather for a given threshold is that it improves the robustness of the signal. Fig. 8 for example shows the time series of probabilities and EFI values as the forecast range was becoming shorter during one of the dramatic rainfall events that have flooded large parts of Central Europe in August 2002. Although PoP >50mm and 20mm indeed were showing unusual values as far as 6 days in advance, the signal was very inconsistent with time. By comparison, the integrated values from the EFI gave a much smoother signal, although in that case the signal did not show before 2-3 days prior to the event.

Maps of the type shown in Fig. 7 have been posted on the Web server for ECMWF Member States since April 2002. Several case studies have been looked at, and forecasters have expressed a keen interest for this type of product. It is not intended of course as an automated warning system – rather a "warning light" that makes sure a potentially dangerous event does not go unnoticed by the forecaster – from this point, more

[3] If *n* is even and $\int_0^1 \left(F_f(p)\right)dp > 1/2$, then $EFI_n = -(n+1)\int_0^1 \left(p - F_f(p)\right)^n dp$

detailed investigations of the full probability distributions, either locally (Fig. 7) or by isolating “worse case” synoptic scenarios should help in detecting dynamical signatures from future observations and making well informed decisions such as issuing public warnings.

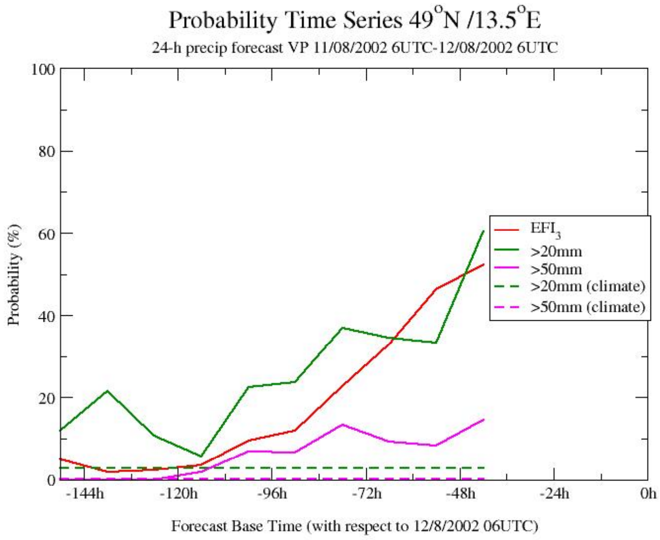

Figure 8 Time series of PoP and EFI for forecasts verifying on 11-12 August 2002 upstream of Prague before the big floods; climate values are also shown.

Another area in which the EFI can be useful however is to explore predictability of severe weather in the early medium range. Case studies are indeed well-known biased estimators of severe weather forecast performance: there is always some kind of signal that the forecaster should on retrospect have been aware of… On the other hand, conducting a verification study using forecasters expertise in real time or delayed mode is both costly in terms of human resources and biased in its own way by the forecaster’s perspective. It must also be recognised that Day3-Day5 forecasts are hardly ever looked at in the context of severe weather at this stage. Strategies to explore the potential benefit that is to be gained from using such forecasts therefore should be explored first before any consideration whether or not to extend the range of warnings by one or two days would bring more benefit than frustration…

As a first step into this direction, 6-30h model forecasts have been used as a proxy for daily precipitation analyses – a reasonable choice in the absence of a comprehensive verification network (Rubel and Rudolf, 2001). The event targeted were those with daily rainfall exceeding 95% of the model climate records, and the verification period was December 2001 to April 2002. Results in terms of Hit Rates and False Alarm Rates are shown for Europe in Fig. 9 for different values of the EFI, both for the EPS and the Control, single value forecast. Results in terms of the ROC curves look rather positive, with a large portion of the curve lying well above the zero-skill diagonal. The longer the forecast range, the larger becomes the benefit from using an ensemble – indeed at the 6-30h range, the verification used here is biased in such a way that the control forecast would be perfect, while the ensemble would still generate some false alarms and miss some events.

Although these results indicate without doubt that there is a skill in the early medium range to forecast moderately severe weather events, it should be realised that to achieve large hit rates, a significant number of false alarms will be generated as well. This fact is in a way hindered by the rarity of the event on the ROC curves. If instead of a False Alarm Rate per number of non-occurring events, the curves are shown with respect to the number of warning issued, it can be seen on Fig. 9 that to achieve a hit rate of 50%, 60 to 80% of the “warnings” would be wrong. Whether or not this is an acceptable result is certainly user dependent. It stresses in any case that using early warnings for severe whether has to be carefully balanced. There is certainly more than pure random in the skill achieved, as a “no-skill” rate of false alarms per warning issued should be 95% when you forecast an event occurring only in 5% of the cases. The rate of hits/ false alarms generated by using such an early warning system would clearly be very far from what can be obtained by waiting until a few hours before the event. But because there are protection measures that cannot wait until the last minute to be taken - releasing water from reservoirs in the case of floods, evacuating of populations for Tropical Cyclones, interrupting air and rail traffic for a major storm event for example - there may be some value attached to such early warning procedures.

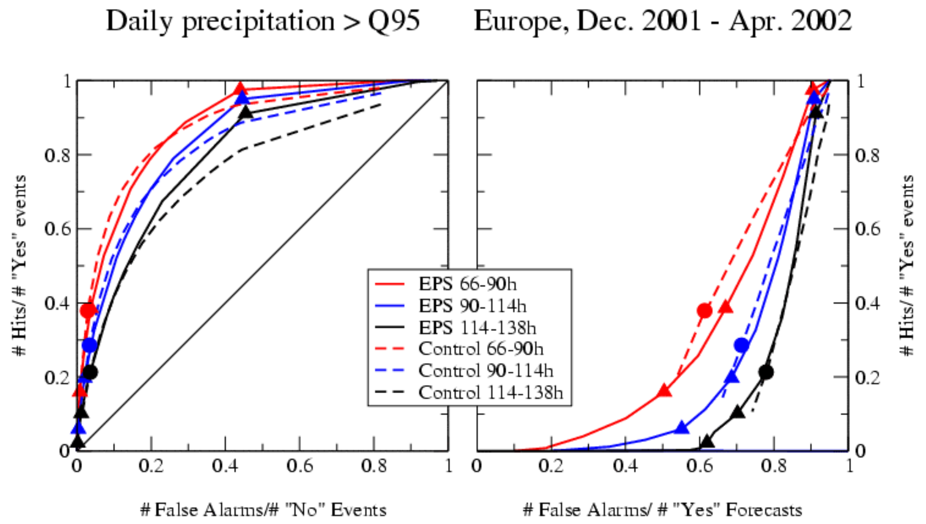

Figure 9 Verification of EPS forecasts that daily rainfall exceeds 95% of the model climate records (Q95); left: ROC curves; right: same, but false alarms are scaled with respect to total number of forecast occurrence; each point corresponds to a different EFI threshold (triangles are 50%, 30% and 0%, while the circles are for the Control forecasting exactly the Q95 value)

4. Tracking tropical cyclones in the forecast

In a recent change (January 2002) of the operational EPS, tropical perturbations were added to the initial perturbations. Barkmeijer et al. (2001) and Puri et al. (2001) found that to benefit from SV perturbations in tropical cyclone (TC) ensemble forecasting, it was necessary to define target areas in the vicinity of TC locations rather than using the entire tropical region and to ignore perturbation growth above 500 hPa in a diabatic SV computation. A more recent study regarding the impact of these so-called ‘targeted tropical singular vectors’ on ensemble TC tracks is presented in Palmer et al. (2001). In order to validate TC forecasts, a TC tracker has been developed at ECMWF in line with developments in other centres (van der Grijn, 2002). No TC genesis is handled at present - only those TCs that have been reported by the regional centres (RSMCs) with responsibility for TC are tracked. The algorithm currently uses model data on a 0.5° by 0.5° latitude-longitude grid. Starting from the analysis, TCs are tracked for 120h every 12h (EPS) and 6h (T511). The tracking algorithm is based on a weighted average of extrapolation and mid-troposphere steering to calculate the first guess position of the next tracking point. (adaptation from Sinclair, 1994). A search is then made in the vicinity of this first guess position for a local pressure minimum. This location, the minimum pressure value and the maximum 10-m wind speed are stored in the TC forecast database for the T511, the EPS T255 Control and all 50 EPS members. The tracker stops if either the TC goes extratropical (+/- 45°), is too weak (p>1010hPa or vorticity<5x10-5s-1) or losts its warm core for 2 consecutive time steps

The skill of the deterministic TC forecasts has been assessed for February to May 2002. The results are shown in Fig. 10. Scores are given for the high resolution T511 model (OPER) and lower resolution EPS control model (CTRL). The sample size decreases rapidly with increasing forecast step due to all TCs not surviving for the 5-days period considered in the verification, which is partly an observed feature (some TCs die during this period) and partly a model failure to maintain the TC activity. The forecast error in core pressure is always positive: TCs in the analysis and forecasts are on average weaker than observed. This is especially the case for the CTRL. However, this positive bias in core pressure seems to decrease, or at least to saturate, later in the forecast. Apparently the model is more capable in developing TCs with a realistic core pressure in the forecast than analysing them in the initial conditions, a feature that maybe related to the limited resolution of both structure (background) functions and incremental 4D-var (T159 inner loops).

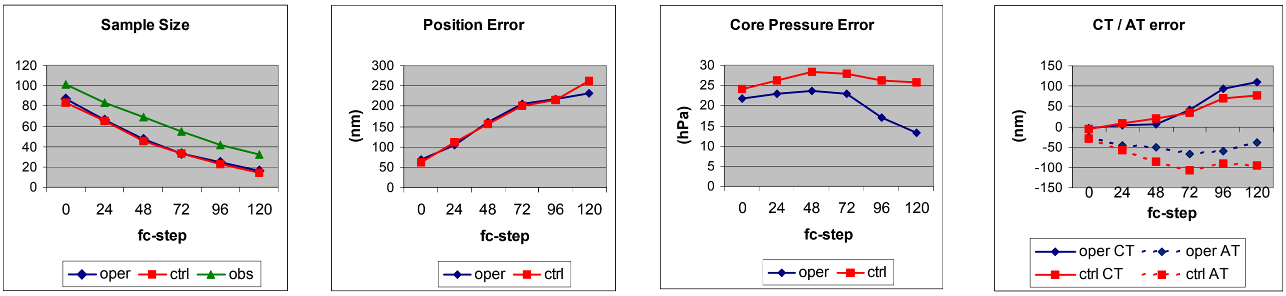

Figure 10 Sample sizes and forecast errors at different forecast steps during February, March, April and May (FMAM) for the high resolution T511 (oper) and the lower resolution T255 EPS control (ctrl)

Of particular importance for practical applications is the distance error in tracking TCs. It is on aveareage 60 nmi in the analysis and increases almost linearly to around 250 nmi at D+5. Model resolution does not seem to have much impact on this error. The-right plot in Fig. 10 shows that both resolutions forecasts have a slow and right-off track bias. From analysis until D+3 forecast, the Along Track (AT) Error for CTRL is almost twice as large as for the OPER while having a comparable Cross Track (CT) Error.

5. "Strike probability" and probabilistic verification of TC forecasts

Conventional probability maps are useful in assessing the likelihood of a certain event at a specific location. However, a drawback of such a probability map is that probability values do not always 'add up'. In theory, probability values can be very low even when all EPS members support a specific event (e.g. 10m wind speed exceeding 25m/s). This is the case when the EPS members predict the event at different locations for a given time step. To enhance the signal for severe weather one must think of a different type of probability: a forecaster is often more interested in whether a TC will affect a certain area than when that TC will hit a specific location; the exact location of the TC is of less importance - within a certain margin, the TC is likely to be equally (or even more) devastating if its centre passes not exactly following the forecasted track, or with a slight time delay.

The concept of 'strike probability' originates from this idea. The strike probability is defined as "the probability that a TC will pass within 120 km (65nmi) radius from a given location at anytime during the next 120 hours". The strike probability is based on the number of members that predict this event, each member having an equal weight. The actual value of 120 km corresponds roughly with the value (75 statute miles) that is in use at the National Hurricane Center (NHC) in Miami (USA).

One of the advantages of a strike probability map as presented in Fig.11 is the elimination of the time dimension and therefore its simplicity. It allows the forecaster to make a quick assessment of the high-risk areas regardless of the exact timing of the event. Another feature of a strike probability map is that it gives the forecaster an estimate of the skill that can be expected from the CTRL. This is because on average the CTRL track error should be equal to the ensemble track spread. In other words, the width of the probability plume is a direct measure of the spread in the EPS and would ideally be a good indicator of the expected error in the CTRL. Elsberry and Carr (2000) also found that a small spread in TC tracks from 5 different models is often indicative of a small consensus track error. In the case shown in Fig. 12 though, rather than the spread of the EPS tracks, it is their bimodal mode that is more remarkable - although the T511 track missed the recurvature of the typhoon, the EPS gave an early indication of the likelihood of such a recurvature that might have been a useful indication for a forecaster in Korea.

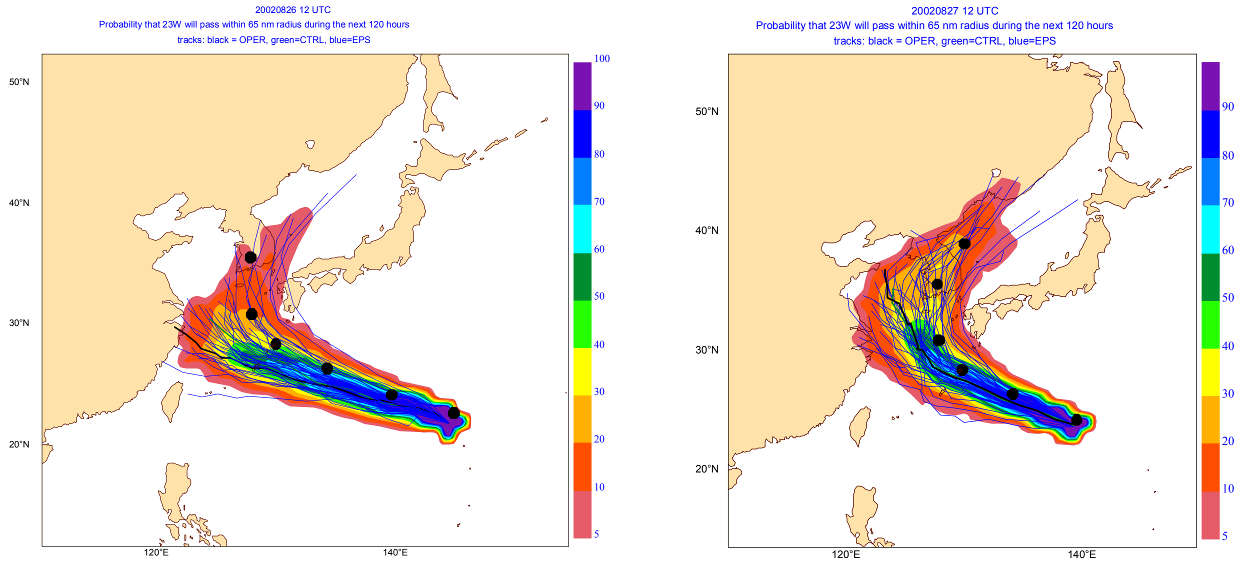

Figure 11 Strike probability of TC 23W (Rusa); probability (shaded) is based on 51 tracks from the EPS (blue tracks, green track is CTRL); the track of the high-resolution deterministic model (OPER) is shown in black; left: 0-120h forecast from 26 August 2002; right: same, but from the 27th; black dots are verifying observations.

The strike probability can be used to assess the skill of the EPS with respect to TC forecasting. Fig. 12 illustrates this both in terms of reliability and signal detection (ROC). A clear improvement in reliability can be seen from the pre-operational testing of cy24r3 that was to be introduced in January 2002. However, the system is over-confident in the high probability range. A 95% probability forecast only verifies in 60% of the cases. This might be an indication that the spread is still too low in the early forecast steps. Just like the reliability diagram (left panel) a clear improvement of 24r3 over 23r4 can be seen. It must be noted that despite the small values of the false alarm rate (FAR), the actual number of false alarms is quite large, as already commented earlier (Fig. 10). However, taken the severity of the event into account and the loss involved, it is likely that most users will favour a forecast system that is designed to reduce the number of misses at the expense of a higher amount of false alarms

6. Summary and perspectives

Preliminary work aiming at designing new products that could be used for early warnings of severe weather conditions have been described in this paper. Although a probabilistic approach is likely to be desirable in order to be able to extract the signal from the numerical forecasts in a way that can be tailored to the user's needs, the quality of the product ultimately relies on the quality of the numerical model itself - from this view point, the forecast quality has improved quite significantly in recent years, and this is what makes it possible to think of extending the range of warnings for severe weather ijto the early medium range. Developments of the EPS system is also of paramount importance if one expects to sample properly the tails of the forecast distributions - those that matter in the early medium range when severe weather is only a possible, but low probability scenario. Some results showing the impact of improving the sampling of tropical cyclone perturbations have been shown here (Fig. 12) that illustrates this point.

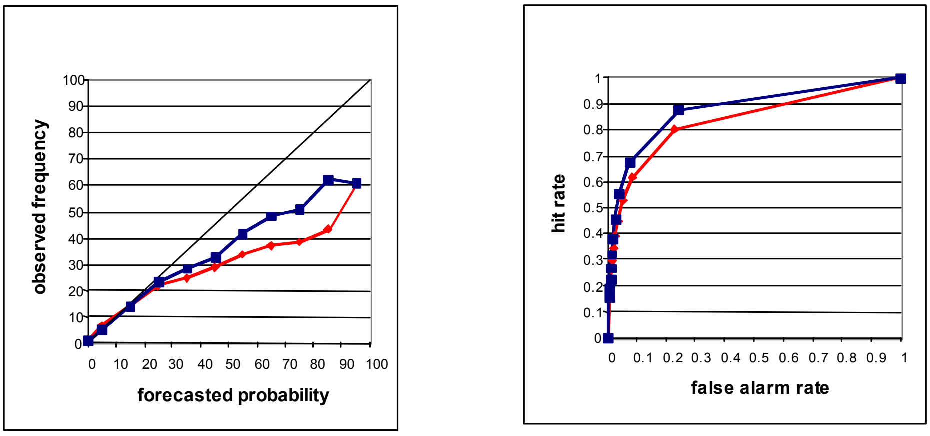

Figure 12 Reliability diagram (left) and ROC (right) for the forecast probability that a TC will pass within 65nm anytime during the first 120 hours of the forecast; red lines: cycle 23r4 (no TC targeting), blue lines: cycle 24r3 (with TC targeting)

Signal detection of severe weather events in the early medium range is likely to be difficult - although one can find some comfort in the fact that the preliminary results shown here for Tropical Cyclones and large precipitation events indicate high rates of successful forecasts, it cannot be disputed that these are achieved provided that preliminary action is taken on the basis of very small probabilities - thus generating very high false alarm rates. The skill achieved by the EPS system in this context is however not negligible: it has been shown that the false alarms are reduced significantly compare to a random forecast system, and even compared to a single-value (EPS Control) system (Fig. 10). It still remains however some way in front of us before users can be convinced that they can make valuable decisions based on such products.

Finally, the status of these products is still very experimental. Operational production is however expected that will involve both archiving in MARS and dissemination of the products in real time - including to RMSC for tropical cyclones following agreed (BUFR) WMO formats for the GTS.

7. References

Atger, F., 1999: The skill of ensemble prediction systems. *Mon. Weather Rev.*, **127**, 1941-1953

Atger, F., 2001: Verification of intense precipitation forecasts from single models and ensemble prediction systems. *Nonlinear Proc. in Geophys.*, **8**, 401-417

Barkmeijer, J., Buizza, R., Palmer, T.N., Puri, K., and J.-F. Mahfouf, 2001: Tropical singular vectors computed with linearized diabatic physics. *Quart. J. Roy. Meteor. Soc.*, **127**, 685-708.

Buizza, R., M. Miller and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System, *Q. J. R. Meteorol. Soc.,* **125**, 2887-2908.

Chessa, P. and F. Lalaurette, 2001: Verification of the ECMWF Ensemble Prediction System Forecasts: A Study of Large-Scale Patterns. *Weather and Forecasting*, **16**, 611-619.

Elsberry, R.L., and L.E. Carr III, 2000: Consensus of dynamical tropical cyclone track forecasts - Errors versus spread. *Mon. Wea. Rev.*, **128**, 4131-4138.

van der Grijn, G, 2002: Tropical Cyclones forecasting at ECMWF: New Products and Validation. *ECMWF Tech. Memo*[4] ., **386**.

Lalaurette, F., 2003: Early Detection of Abnormal Weather Using a Probabilistic Extreme Forecast Index. *Q. J. R. Meteorol. Soc*. (in press, long version also available as *ECMWF Tech. Memo*[4]. **373**).

Palmer, T.N., J. Barkmeijer, R. Buizza, C. Jakob, F.Lalaurette, T.Paccagnella and D. Richardson, 2001: Severe weather prediction. *ECMWF Tech. Memo*[4]., **352**.

Puri, K., J. Barkmeijer and T. N. Palmer , 2001: Ensemble prediction of tropical cyclones using targeted diabatic singular vectors, *Q. J. R. Meteorol. Soc*., **127**,709-732.

Richardson, D.S., 2000: Skill and relative economic value of the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc*., **126**, 649-668.

Rubel, F. and B. Rudolf, 2001: Global daily precipitation estimates proved over the European Alps. *Meteorol. Z*., 10, 403-414.

Sinclair, M.R., 1994: An objective cyclone climatology study for the southern hemisphere. Mon. Wea. Rev., 122, 2239-2256.

Wilks, D.S., 1995: Statistical methods in the atmospheric sciences. Academic Press, 59, 467pp.

[4] Recent ECMWF Technical Memoranda can be retrieved in electronic format from http://www.ecmwf.int/publications/library/ecpublications/techmemos/tm00.html