Combined use of ECMWF Ensemble Prediction System and high resolution model guidance in operational medium range forecasting

Frederic ATGER, ECMWF

1. Introduction

An evaluation of the ECMWF Ensemble Prediction System (EPS) is presented in this paper. This evaluation has been achieved by the medium-range forecasters of the national forecasting service of Meteo-France, in Toulouse, since the start of the daily production in May 1994. The method which has been used, based on a combined interpretation of EPS clustering products and ECMWF high resolution model (T213) traditional products, is described in section 2. The results of the experiment are presented in section 3. A possible application of the results to operational forecasting is proposed in section 4.

As a forecasting tool, the EPS is supposed to be designed to help forecasters in their operational tasks, but the usefulness of the system depends on the methodology which has been chosen. Basically, two approaches are possible in order to extract some relevant information from the EPS products: a statistical approach and an interpretative one.

There is a common statistical use of EPS based on the following assumption: "the probability of a meteorological event is the proportion of ensemble members forecasting this event", in which an "event" is typically the fact that a meteorological parameter (pressure, wind, rainfall...) crosses a threshold over a given point (or an area), at a given time (or for a period). This approach leads to a wide range of products as probability maps or graphs, and to direct probabilistic forecasts able to satisfy any user, whatever his point of interest. But the quality of these end forecasts directly depends on the model skill and cannot be improved by means of forecasters interpretation. Moreover, possible inconsistencies with deterministic forecasts based on high resolution models may appear.

Alternatively, an interpretative use of EPS could be based on a similar assumption: "the probability of a meteorological pattern is the proportion of ensemble members forecasting this pattern", in which a "pattern" is an ensemble of various atmospheric elements as jets, fronts, troughs, etc. An appropriate clustering of ensemble members yields to such different patterns a forecaster is supposed to be able to subjectively recognise and associate with typical weather characteristics. Following this approach, the quality of end forecasts does not depend only on the model skill, but on the forecaster's ability too. Furthermore, a combined use of EPS and high resolution model guidance is possible, since the method is not limited at the ensemble members but can be applied at any model.

Such an interpretative methodology has been followed by french medium-range forecasters in the described experiment.

2. Description of the experiment

2.1 Multi-scenarios forecast

The first step consists in a clustering of ensemble members, in the day of reception of data, in order to highlight the different meteorological patterns forecast in the ensemble. From the forecaster's point of view, the perfect clustering should be purely subjective, since the "meteorological distance" between two patterns does not yet exist. Such a clustering is not easily practicable, just because of the huge quantity of data to be checked. In this experiment, the automatic ECMWF clusters are used as an input: "super-clusters" are built by grouping clusters according to their large scale pattern similarities during the D5/D7

period (corresponding to a day+4/day+6 period from the day of forecasting). Synoptic details are naturally disregarded since they have a low predictability at these ranges (Atger and Mornet, 95). As a result, the so called *ensemble scenarios* group all those ensemble members supposed to lead to a similar evolution of the weather over France, such an evolution exclusively depending on the large scale pattern.

In a similar way, a *T213 scenario* is based on the high resolution model pattern.

The result is a *multi-scenarios weather forecast* over France.

2.2 Verification

In a second step, 8 days later, each scenario is subjectively compared to the real evolution and is marked as *good* or *bad* according to the differences between the forecast and the observed patterns.

In the following, the proportion of *good* scenarios during a period or for a specified category of scenarios will be regarded as a skill.

3. Results

In this paper we focused on the 6 months period from 1st September 1994 to 27 February 1995. Each forecast was composed with 1 to 6 ensemble scenarios, generally 2 or 3, and a T213 scenario which was generally (but not systematically) supported by one of the ensemble scenarios.

3.1 Reality is not always verified by an ensemble scenario

The *best ensemble scenario*, i.e. the one which is the closest to the real evolution, is not systematically good (fig.1). The proportion of cases in which the EPS is not able to capture the real evolution is about 20%, confirming the Talagrand's diagram results (Strauss and Lanzinger, 95).

3.2 T213 scenarios are more skilful than ensemble scenarios

In average, the T213 scenario is more often in accordance with reality than ensemble scenarios (fig.1). Even the *largest* ensemble scenario, i.e. the more populated, is not more skilful than the T213 scenario. This result confirms the importance of high resolution in numerical forecasting, even in the longest ranges when the common skill tools as RMS seem to show the contrary (Tracton and Kalnay, 93).

Furthermore, the relatively high proportion of good D5/D7 forecasts based on the T213 large scale scenarios (59%) is to be compared to the proportion of good D6 forecasts (about 20%) based on a traditional synoptic interpretation of the same model (Mornet and Lefort, 93). This comparison clearly shows that the choice of a methodology based on the large scale evolution was judicious.

3.3 The T213 scenario is more skilful when supported by an ensemble scenario

The fact that the T213 scenario is supported by one of the ensemble scenarios appears as a strong criterion of predictability (fig.2). In these cases:

- the T213 scenario is more skilful;
- ensemble scenarios (even those which do not support the T213 scenario) are more skilful;
- the ability of the ensemble to capture reality is increased.

When the T213 scenario is supported by an ensemble scenario, it is generally by the largest one (in 83% of cases) which could explain these last results. But even when the T213 scenario is supported by small sce-

narios it is not less skilful than when it is not supported at all (compare fig.2 and fig.6).

3.4 The skill is higher when the ensemble spread is small

In case of small spread, the subjective clustering yields to a single ensemble scenario, almost systematically supporting the T213 scenario. In these cases, the skill is maximum (fig.3). On the other hand, in case of large spread, when there are several scenarios, the skill does not decrease when the number of ensemble scenarios increases. This result confirms the generally noticed link between spread and skill: small spread indicates high skill, but large spread does not indicate anything.

3.5 There is a correspondence between the skill of an ensemble scenario and its number of members

The reliability diagram (fig.4) shows a quite good verification of theoretical probabilities by observed frequencies of good scenarios. The underestimate of low probabilities and the overestimate of high probabilities are usual on this kind of diagram (Strauss and Lanzinger, 95).

As one could expect knowing the second result of 3.3, the correlation is much less verified when the T213 scenario is not supported by one of the ensemble scenarios, with a systematic overestimate of forecast probabilities (fig.5).

3.6 The T213 scenario is the most likely available scenario

The ensemble scenarios which support the T213 scenarios are more skilful than others, whatever their population (which is generally large): in average, even the smallest scenarios supporting the T213 scenario are as skilful as the largest "not supporting" scenarios (fig.6).

As a consequence, when the largest ensemble scenario is not in agreement with the T213 scenario and could be an alternative scenario, it is not more likely than the T213 scenario itself (fig.7).

4. Conclusion: application to operational forecasting

4.1 Probabilistic forecasting

A probabilistic forecast could be based on a combined interpretation of EPS and T213 guidance. Such a forecast should be composed with:

- one (when a single ensemble scenario supports the T213 scenario) or several evolutions of the weather based on the different available scenarios;
- the probabilities of each of these evolutions to be marked as good in a subjective evaluation.

The probability of an evolution based on an ensemble scenario not supporting the T213 scenario should be inferred from its number of members, through a correspondence rule taking into account the characteristics of the observed frequencies distribution, according that the T213 scenario is supported or not by an ensemble scenario.

The probability of a T213 scenario which is supported by an ensemble scenario should not be the corresponding ensemble scenario probability, but a probability inferred from its number of members through a different correspondence rule, since the observed frequencies distribution is not the same for the supporting and the not supporting ensemble scenarios.

The probability of a T213 scenario which is not supported by an ensemble scenario should be the mean observed frequency of good scenarios, according that the T213 scenario is supported or not by an ensem-

ble scenario.

4.2 Deterministic forecasting

The deterministic forecast inferred from the previous probabilistic one is the most likely weather evolution, so the one based on the T213 scenario. Alternatively, the forecast could be sometimes based on an ensemble scenario which would not support the T213 scenario and would be populated enough to be considered as likely as (but not more than) the T213 scenario itself. Excepted in that case, the usefulness of an EPS interpretation is none for a "pure" deterministic forecast.

However, a deterministic forecast becomes more informative when completed with an estimate of its quality. Such an estimate could be a level of confidence according to the type of situation:

- higher confidence when there is only one ensemble scenario supporting the T213 scenario;
- normal confidence when the T213 scenario is supported by one of the several ensemble scenarios;
- lower confidence when the T213 scenario is not supported by one of the ensemble scenarios.

5. References

Atger, F. and B. Mornet, 1995: Operational medium range weather forecasting, 2nd European Conference on Applications in Meteorology, ECAM, Toulouse

Mornet, B. and T. Lefort, 1993: Synthese des controles subjectifs du modele T213 du CEPMMT, note SCEM/PREVI, METEO-FRANCE

Strauss, B. and A. Lanzinger, 1995: Validation of the ECMWF EPS, Proceedings of the 1995 Seminar on Predictability, ECMWF

Tracton, S. and E. Kalnay, 1993: Operational ensemble prediction at the NMC, practical aspects, Weather and Forecasting vol.8, pp 379-398

Figure 1 : T213 and ensemble scenarios

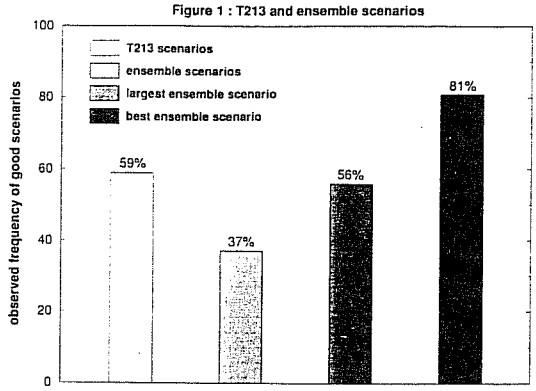

Figure 2 : T213 and ensemble scenarios

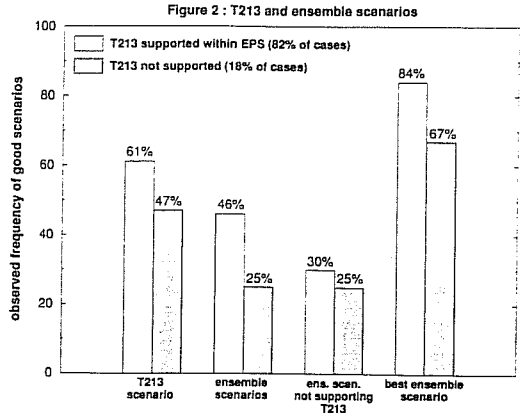

Figure 3 : T213 scenarios

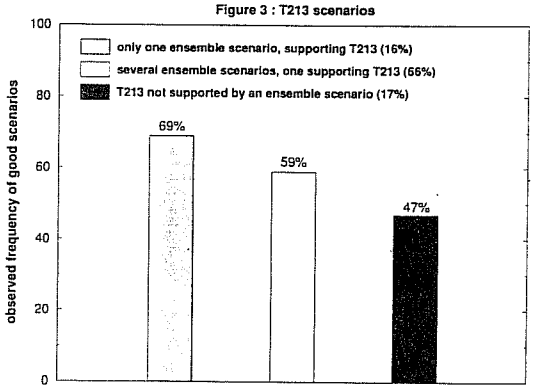

Figure 4 : ensemble scenarios

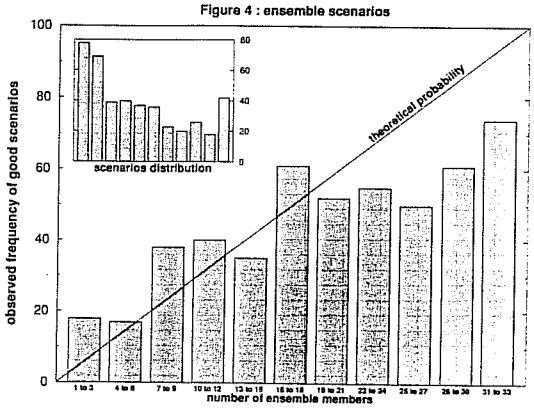

Figure 5 : ensemble scenarios

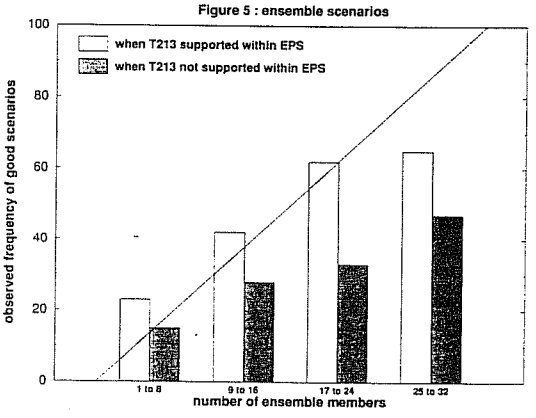

Figure 6 : ensemble scenarios

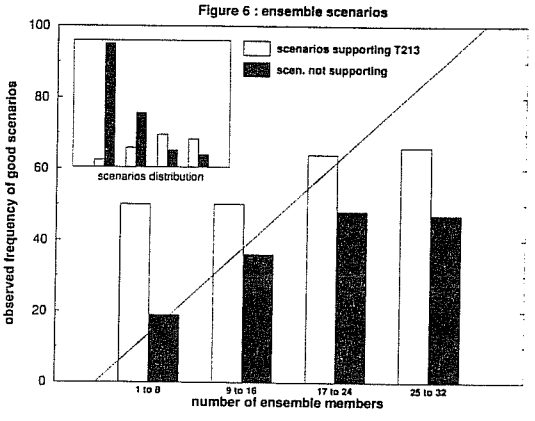

Figure 7 : largest ensemble scenario vs T213