CLASSIFICATION OF ENSEMBLE FORECASTS WITH THE HELP OF AN ARTIFICIAL NEURAL NETWORK

P. Eckert, D. Cattani, J. Ambühl
Swiss Meteorological Institute

Summary: The problem of clustering the members of an ensemble forecast is addressed. 500 hPa fields are classified by a self-organizing Kohonen artificial neural network. The members of the ensemble forecast are classified according to the learned situations. This classification technique also provides a measure for comparing forecasts (operational and ensemble) with their verifying analysis.

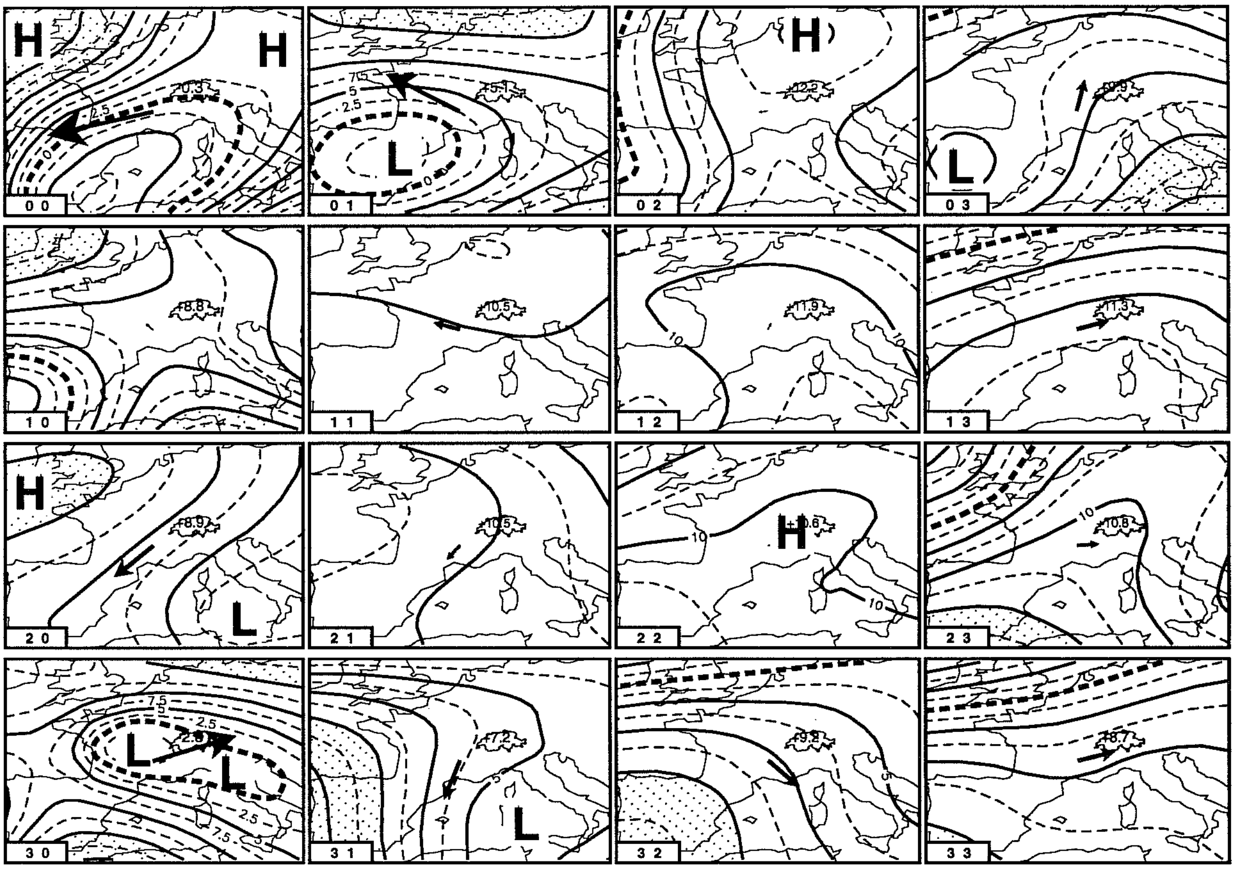
1. INTRODUCTION

With the availability ensemble products, the forecaster is faced to the interpretation of the huge amount of numerical material. Various techniques related to the treatment of the information can be used to group the forecasts. However, at the very end of the process, the graphical results presented to the forecasters have to be of meteorological relevance.

Following method has been used here: a neural network is designed to learn a certain amount of different meteorological situations presented to the system. This can be looked as a system for classifying the situations. This learning step is performed once for all, the classified situations remain fixed afterwards. In operations then, each member of the ensemble forecast is attributed to its closest analogous situation. The consideration of the number of realised situation is supposed to reveal facts about the predictability, the multimodality or alternative scenarios of the forecast.

2. THE LEARNING PROCESS

The neural network we use is a so called Kohonen chart. It is composed of 8x8 neurons set on a square. Each neuron is supplied with 10x10 synapses connected to the 10x10 grid points where 500 hPa geopotential values are provided (in our case, but any other field could be used). The covered area goes from 9W to 18E in longitude and from 36N to 54N in latitude. Three years of 00z and 12z analyses have been fed into the system for learning. The patterns of synaptic weights (= learned meteorological situations) for the 16 first neurons is presented in the following figure:



In each map, it is possible to recognize synoptic features like highs, lows, ridges and troughs in addition to the height of the field and the shape of the circulation over the alpine region.

In the view of understanding the verification scores given later, we give now explicitly the formulas for computing the distances. If H_{ij} is the height, then

$$|H|^2 = \frac{1}{N^2 M^2} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(H_{ij} - \hat{H}\right)^2$$

where $\hat{H}$ is an overall monthly mean, M and N are the dimensions x and y of the physical fields.

The normalized field is given by $h_{ij} = \frac{H_{ij} - \hat{H}}{|H|}$

After the learning process, each neuron holds a map of patterns in terms of synaptic weights w_{ij}. It is now possible to compute the distance between neurons or between neurons and normalized fields:

$$d = \langle h, v \rangle = \sum_{i=1}^{M} \sum_{j=1}^{N} \left(h_{ij} - w_{ij}\right)^2$$

3. CLASSIFICATION OF ENSEMBLE FORECASTS

The classification of ensemble members into the fixed situations presented by the neural network is done by computing the distances to each of the 8x8 neuron. The elected neuron minimizes the distance. When all members are classified, it is possible to look at the distribution of elected neurons on the neural map. On the graphical representation shown on next page, one can rapidly assess the spread, bimodality,... of the ensemble. In order to quantify the spread we introduced the entropy

$$s = \sum_{k,l}^{neurons} p_{kl} \log\left(\frac{1}{p_{kl}}\right) \qquad p_{kl} = \frac{n_{kl}}{E}$$

n_{kl} is the number of hits in neuron kl

and E is the number of members of the ensemble (34 here).
s ranges from 0 (1 neuron hit) to log(34)=3.53 (34 neurons hit).

On the next page we present an example of an operational output displayed each morning to the forecasters. The distribution maps for the timesteps T+96 to T+240 are plotted on one single sheet and the last figure shows the evolution of entropy.

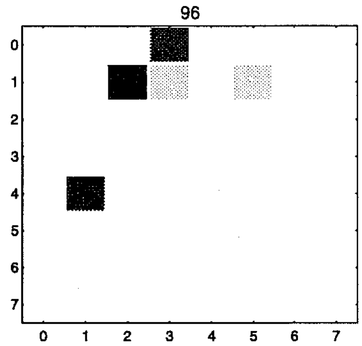
At T+96, most members of the ensemble and the operational forecast agree on the region of the 1,2 neuron (high field, ridge over central Europe), but 7 members point to a slightly different option (4,1 neuron, ridge over the Atlantic, north-westerly winds over the Alps).

The spread of the T+144 forecast is still small and the confidence in a 4,3 forecast (linear westerly situation, height slightly above average) can be considered as good. At further timesteps, the EPS members cover a good portion of the chessboard.

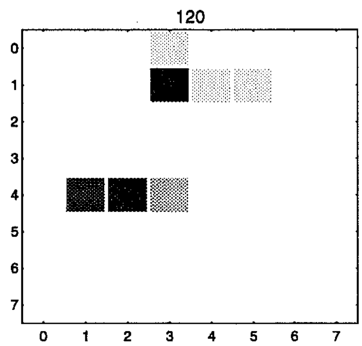
The study of a few cases led us to the following (provisional) conclusions:

- When the spread is small, it gives good confidence in the forecast. As will be seen later, the forecast to take in this case is the operational one.
- When the spread is large, the EPS gives alternative scenarios, but the way to choose them is rather subjective.
- On occasions, in rapidly evolving situations, there are timing problems in the forecast (as well EPS as operational). To avoid this, we will try to make our system learn evolutions rather than snapshot situations.

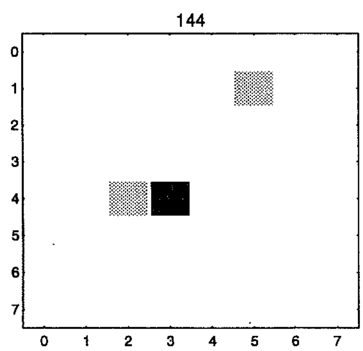
RUN : 26 6 95 12



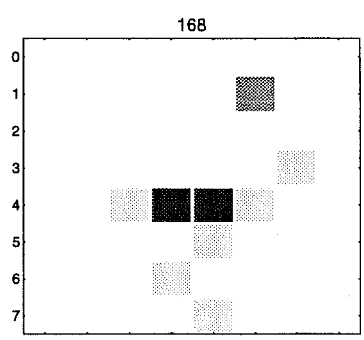
N (1, 2) : 16 members + opr + nop
N (0, 3) : 7 members
N (4, 1) : 7 members
N (1, 3) : 1 member
N (1, 5) : 1 member



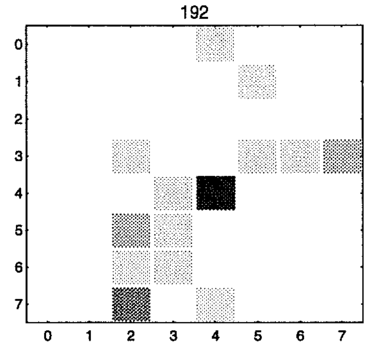
N (1, 3) : 9 members + opr
N (4, 2) : 8 members + nop
N (4, 1) : 7 members
N (4, 3) : 4 members
N (1, 4) : 2 members
N (0, 3) : 1 member
N (1, 5) : 1 member



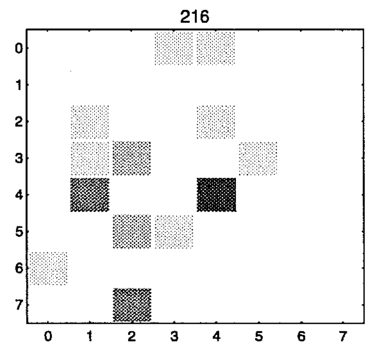
N (4, 3) : 24 members + opr + nop
N (1, 5) : 4 members
N (4, 2) : 4 members



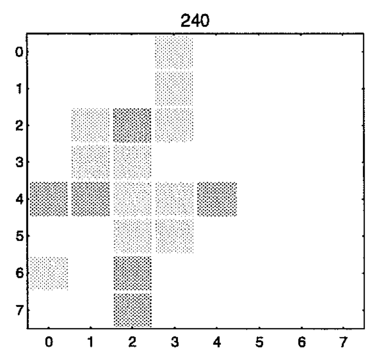
N (4, 3) : 11 members
N (4, 4) : 9 members + nop
N (1, 5) : 5 members + opr
N (3, 6) : 2 members
N (4, 2) : 1 member
N (4, 5) : 1 member
N (5, 4) : 1 member
N (6, 3) : 1 member



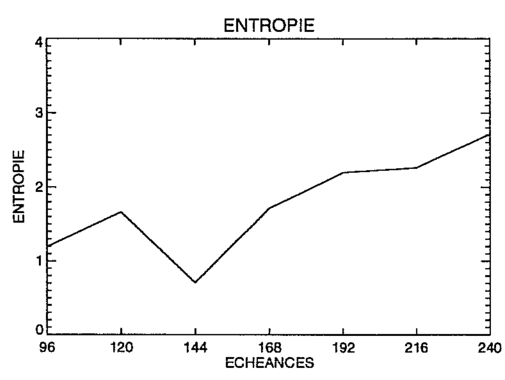
N (4, 4) : 10 members + opr
N (7, 2) : 5 members + nop
N (3, 7) : 3 members
N (5, 2) : 3 members
N (3, 6) : 2 members
N (0, 4) : 1 member
N (1, 5) : 1 member
N (3, 2) : 1 member



N (4, 4) : 8 members
N (4, 1) : 5 members
N (7, 2) : 4 members + opr
N (3, 2) : 4 members
N (5, 2) : 3 members
N (3, 1) : 1 member + nop
N (0, 3) : 1 member
N (0, 4) : 1 member

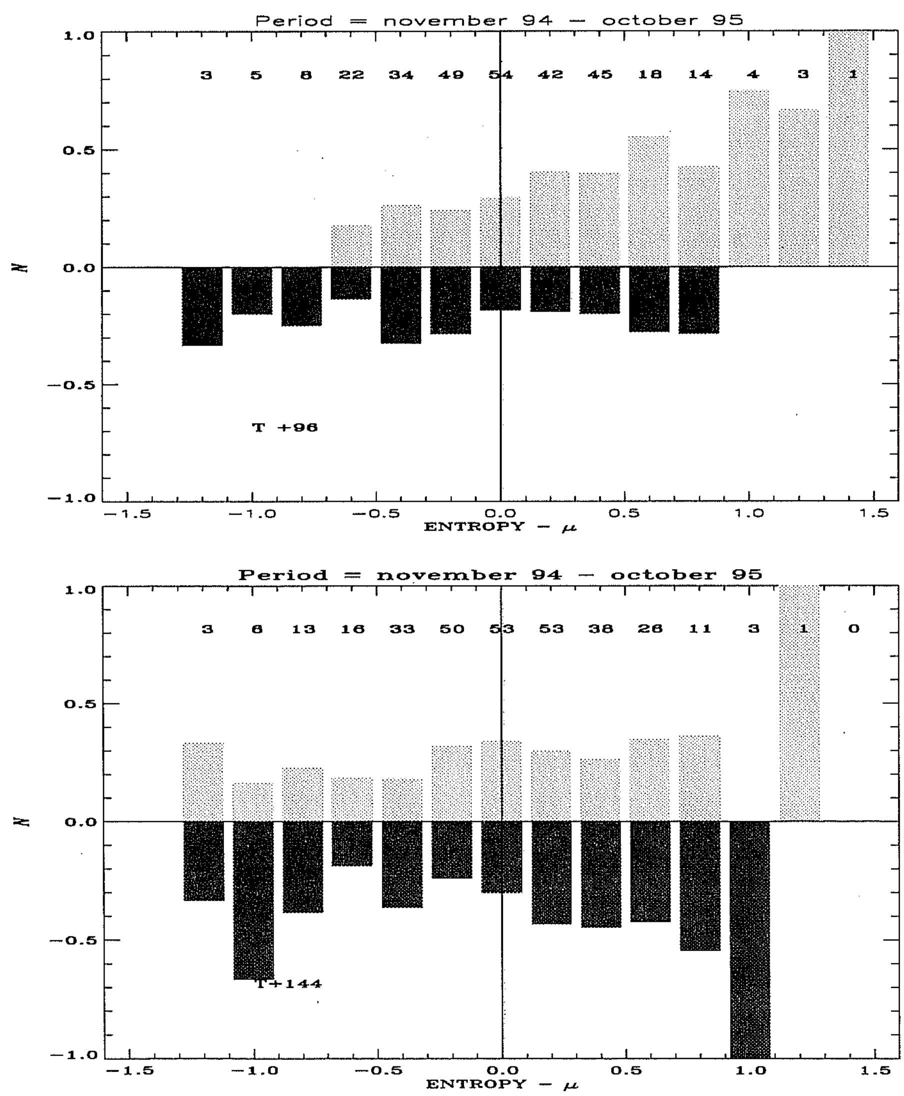


N (2, 2) : 4 members
N (4, 0) : 2 members + nop
N (4, 1) : 3 members
N (4, 4) : 3 members
N (6, 2) : 3 members
N (7, 2) : 2 members + opr
N (3, 2) : 2 members
N (4, 2) : 2 members



3. VERIFICATION

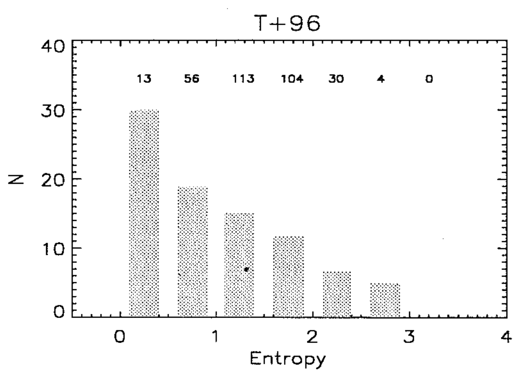
The classification of situations by an artificial neural network has been verified for the period from November 1994 to October 1995. In the following figure, the neuron associated to the operational T213 forecast is compared with the neuron containing most ensemble members (called the peak). The quality of the forecasts is plotted as a function of the deviation of the entropy from its mean. The vertical axis has positive values when the neuron associated with the operational forecast is closer to the analysis than the neuron corresponding to the peak of the EPS (according to the *RMS* distance) and negative when the peak is better. The deviation is divided by the total number of members (operational better + peak better + operational equal peak) yielding a number between -1 and +1.



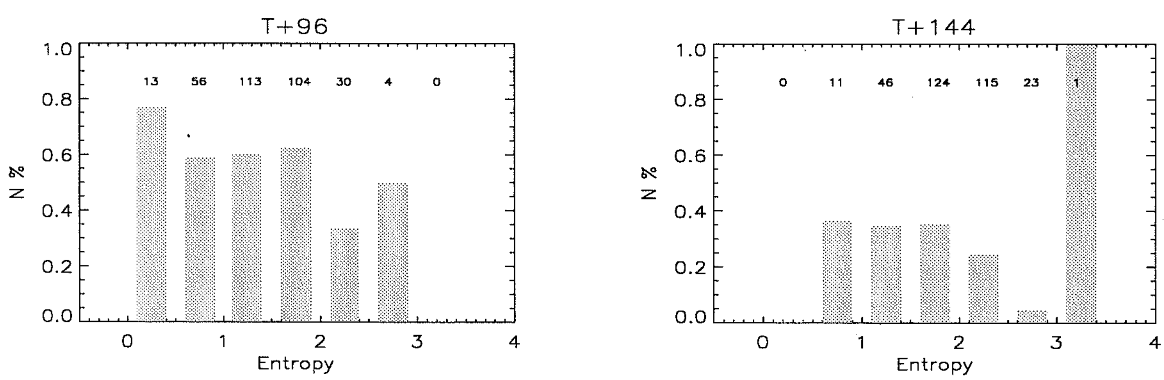
Relative number of cases with respect to the deviation of mean entropy, positive if operational is better than the peak, negative if the reverse holds. The numbers on top show the total amount of cases for each entropy class.

At T+96 and for a small spread, the peak often coincides with the T213 run. When this is not the case the peak is often better but the operational run is then quite close to the peak. This feature can be related to the well known nervousness of the T213 model. A further analysis shows that if the spread is small then the quality of the forecasts (operational and ensemble) are above average. On the other hand, if the entropy is large then the quality tends to be poor and the T213 forecast rarely coincides with the peak. There is a clear tendency for the operational forecast to give a better guidance in this case. Altogether, it can be said that until day 5 a deterministic forecast based on the operational model is possible and of good quality. The ensemble can be used to some extent to assess the confidence level of the forecast.

In order to explore the behaviour of the longer forecast ranges, let us introduce two more type of figures. The next one shows the amount of ensemble members which verify close to the analysis.



Following figures show the relative amount of good operational forecasts



At T+96 and big spread, it clearly can be seen that the amount of good ensemble members falls below 10, while the amount of good operational forecasts remains around 50%. At T+144 and high spread, the situation is mainly unpredictable. At lower spread instead, more than 10 ensemble members are still of good quality, whereas the number of good operational forecasts drops below 40%. We think that at day 6 and 7, it is still possible to give a deterministic forecast, but this time by taking a choice between the operational model and one of the ensemble peaks. Obviously, it is also possible to issue probabilistic forecasts in this time range. Beyond day 7, only probabilistic forecasts should be considered, but then it should be questioned if the ensemble probabilities are better than the climatology.