TECHNICAL DOKUMENTATION:

The probabilistic multi-trend filter: theoretical formulation and practical application

June 2006

Autor: Corinna Möhrlen

Co-autors: J.U. Jørgenseni(2), M. Pahlow (3), D. Entekhabi (4)

Addresses of Authors:

(1)WEPROG GmbH, Altdorf, Germany Email: com@weprog.com

(2)WEPROG ApS, Ebberup, Denmark Email: juj@weprog.com

(3)Ruhr-University Bochum, Germany Email: m.pahlow@rub.de

(4)Massachusetts Institute of Technology, Boston, USA Email: darae@mit.edu

Abstract

Determining the most probable forecast from an ensemble of forecasts requires suitable statistical tools. They must enable an forecaster to interpret the model output, to condense the information and to provide the desired product. For this purpose, a probabilistic multi-trend filter (pmt-filter) for statistical post processing of ensemble forecasts is introduced. It provides an alternative to ensemble classification methods that are in use today. In essence, the algorithm is a forward-backward clustering method that strips off those ensemble members that do not follow a group or are only temporarily the most probable forecast. Here the underlying theory is developed and a practical application of the pmt-filter to data from a multi-scheme ensemble prediction system are shown.

1 Introduction

The use of ensembles is intended to provide a set of forecasts which cover the range of possible uncertainty, recognising that it is impossible to obtain a single deterministic forecast which is always correct (Legg et al., 2002). Ensemble forecasts are widely used. A well known application are weather ensemble predictions (e.g. Molteni et al., 1996; Tracton and Kalney, 1993; Pellerin et al., 2003). These weather ensemble predictions in turn are employed in electricity demand forecasting (Taylor and Buizza, 2003; Lang et al., 2006), as input for hydrological models (e.g. Gupta et al., 2002; Gouweleeuw et al., 2005), or ocean models (e.g. Farina, 2002; Vialard et al., 2005). Ensembles also find widespread application in earthquake studies (e.g. Liu et al., 2004; Rundle et al., 2006) and ensemble stream flow forecasting (e.g. Georgakakos et al., 2004; McIntyre et al., 2005; Moradkhani et al., 2005). Palmer (2002) lists further applications of ensemble forecasts, ranging from ship rerouting, pollution modelling, weather- and climate-risk finance, disease prediction and crop-yield modelling. Realising the full potential of an ensemble forecast requires statistical post processing of the model output (Gneiting and Raferty, 2005). A major challenge of ensemble prediction is to condense the large amounts of information provided by ensembles into a user-friendly format that can be easily interpreted and used by forecasters (Tracton and Kalnay, 1993). Various approaches are in use to aid in determining the most probable forecast from the ensemble. Cluster analysis is a well established multivariate technique which is commonly employed for this purpose (see Gong and Richman (1995) for an extensive overview; further applications can be found in e.g. Alhamed et al., 2002; Yussouf et al., 2004; Nakaegawa and Kanamitsu, 2006). An alternative method to classify ensemble forecasts is tubing (Atger,1999).

Although these methods are useful tools at single points in time, none of these methods consider past and future states. When dealing with time series of data, it is however desirable to create a smooth function over time. To avoid the algorithm to jump from one likely outcome to another, a time filtering method is required. For this reason the probabilistic multi-trend filter (pmt-filter), a time dependent clustering method for time series analysis of ensembles, is introduced. Its practicality and potential are shown by presenting a sample application to wind power ensembles.

2 A new ensemble classification method

The classification of the ensemble members into groups of probable outcomes of the meteorological future is an important and challenging operation. The detail of interpretation of derived probabilities from ensemble predictions depends strongly on the specific requirements of the end user. A computer cluster, running an Ensemble Prediction System (EPS), will produce large amounts of data in a rather short period of time. However, generally only a fraction of the ensemble contains relevant information for the end user. Further analysis and presentation of the data is therefore equally important to the generation of the ensemble itself. An efficient way to reduce the amount of information is to use the ensemble mean. However, the mean is only a suitable choice, if the predictions of all ensemble members are equally accurate. At present two methods are commonly used to preselect ensemble members. These are clustering and tubing. The clustering procedure yields an unbiased selection and groups ensemble members around hypothetical centroids. It is therefore a basic selection procedure towards similarities in the data. Tubing on the other hand groups ensemble members according to the criterion of similar distance from the ensemble mean. It groups members along axes starting from the ensemble mean and reaching towards the extremes of the distribution. These axes represent the variation of the ensemble members deviating from the mean (Atger, 1999).

Here a new method is proposed as an alternative to these classification methods: *the probabilistic multi-trend filter* (pmt-filter). The pmt-filter is based on the classical clustering method, yet it selects groups of ensemble members by taking the past and future of the ensembles into account. It is a forward-backward clustering method that strips off those members that do not follow a group or are only temporarily the most probable outcome. The algorithm is designed as a method to determine a conservative guess of the most probable outcome of e.g. the output from an ensemble forecast. It should help to build up confidence for interpreting the probability distribution and estimate the risks for certain actions due to the uncertainty that is inherent to forecasting.

It was found that the classical clustering method produced unacceptably abrupt changes in the computations of the most likely meteorological future. In fact, it was observed that when computing the most likely outcome, computed as the group with the highest probability, the classical clustering algorithm "hopped" from one possible future to the next within one time step. This caused the algorithm to become very unstable whenever there were two or more larger groups of members that had similar probabilities.

Therefore, a method was developed that utilises the past and future probability distribution as weighting function. This means that those members that had highest probability in the previous time step start in the current time step with a higher weight than the other members. In that way, the individual members follow certain groups with characteristic structure, once the selection has passed the forking point.

Figure 1 is a graphical demonstration of the pmt-filter. In this example, we focus on six ensemble members (A through F). The ensemble members could also be groups of ensemble members. These members are ranked with a traditional clustering technique. The difference between the pmt-filter algorithm and a traditional cluster analysis is that the pmt-filter forms groups that persist over time. In this graphical example replicates B,C,D and F are grouped closely during the last 3 time steps (n-3 to n). In the next time step replicate F leaves the group and forms a new group with replicate E. Over the three time windows backward (n-1..n-3), the present state n and forward states (n+1..n+3), the replicates B,C and D represent a clustered group. This group is a robust and consistent forecast.

The method introduced here to employ the past and future probability distribution will be referred to as forward-backward stepping. Long term statistics are used in the algorithm for the initial guess, but the selection becomes gradually more dynamic as the iterations progress. The algorithm can also be applied in two-dimensional space, i.e. for two-dimensional fields.

2.1 Mathematical formulation of the pmt-filter

The pmt-filter developed here is a mathematical filter to compute probabilities for ensemble members. However, it can certainly be utilised in other applications. It must be noted, that the probabilities are entirely computed from the density of the ensemble, while one parameter, the *best guess* is computed implicitly. This *best guess forecast* should reflect the most likely outcome of the "weather" in contrast to the ensemble mean.

In the first step the vector $fc_p(i,j)$, which are the forecasts of all ensemble members (i) over the forecast length j and weight coefficients $c_w(i,j,k)$ for all ensemble members (i) over the forecast length (j) are coupled in time-space (k) to take the past and future development into account. The coupled terms are solved in an implicit iteration process and then decoupled again. The iteration algorithm uses forward-backward stepping. A time-averaging filter is applied that iterates three times to smooth the time series. The width of the time window spans from $-k$ to $+k$. The "best guess" forecast $fc_{bestg}(i,j)$ with the highest probability over the increased time window can then be defined as:

$$fc_{bestg}(i,j) = \widetilde{fc_p}(i,j) + fc_{mean}(i,j) \tag{1}$$

with

$$fc_{mean}(i,j) = \frac{1}{n_{eps}} \sum_{i=1}^{n_{eps}} fc_p(i,j) \tag{2}$$

where $i = 1, 2, ...n_{eps}$ is the number of ensemble members, $j = 1, 2, ...fc_{len}$ is the time step variable over the forecast length (fc_{len}) or data dimension of each ensemble member. The $\widetilde{fc}(i,j)$ is the forecast that is updated after each iteration process of the pmt-filter algorithm. The mean of the ensemble fc_{mean} is subtracted from each forecast to let the function vary around zero and added again after each iteration step.

As mentioned before, the weight function c_{wp} is coupled in time, and the ensemble mean normalised forecasts are integrated over the time range $-k$ to $+k$. The coupling in time of the weight functions and the forecasts is performed by

$$\widetilde{fc_p}(i,j) = fc_{pmax}(i,j+l) - fc_{mean}(j+l) \tag{3}$$

where $i = 1, 2, ...n_{eps}$, $j = 1, ...fc_{len}$ and l is the increased time interval over past and future time steps $-k, ..k$ and $A(l)$ is a weight function in this interval. Note, that the mean fc_{mean} is computed over the time interval l. The forecast fc_{pmax} with the highest probability over the time interval -k to k is computed from the probability distribution of the ensemble and the weight function c_{wp}:

$$fc_{pmax}(i,j+l) = p_{eps}(i,j) \cdot c_{wp}(i,j) \tag{4}$$

$$c_{wp}(i,j) = \sum_{l=-k}^{k} \left[c_{wp}(i,j+l) \cdot A(l)\right] \tag{5}$$

The indices are the same as for $\widetilde{fc_p}$. The two parameters c_{wp} and fc_{pmax} are passed to the implicit algorithm to compute the probability distribution and the forecast with maximum probability (fc_{pmax}) in the "time-coupled" system (-k..k).

The weighting factor c_{wp} is in the first step estimated from a long-term statistical coefficient. If this coefficient is unknown *a priori*, it can initially be set to 1. The function A(l) could also contain inherent weight coefficients, if it would be known that the quality of the individual ensemble member is not the same. In that case the matrix A(l) can be used to simulate a ensemble with equally skillful ensemble members.

After the first iteration process c_{wp} is updated with $\widetilde{c_w}$ over the full forecast length and decoupled again by inverting the weight function A(l) in the following way

$$c_{wp}(i, j, k+1) = \sum_{l=0}^{fc_{len}} (\widetilde{c_{wp}}(i, j+l, k) \cdot A^{-1}(l)). \tag{6}$$

Note, that the functions are decoupled after completion of the integration process to the actual time step.

Inside the pmt-filter algorithm, the ensemble forecasts are first evaluated according to their probability density with an Euclidean distance measure. The individual forecasts are now integrated to a sum, minima, maxima and the mean of the forecasts $\widetilde{fc_{mean}}$. A matrix of parameter bins and time is built to select, which members are contained in the bins. A probability density function can then be calculated.

$$sum_{cw} = \sum_{i=1}^{n_{eps}} c_{wp}(i). \tag{7}$$

The pmt-filter function δ_{eps} is used for the selection procedure of the forecast with the maximum probability to occur (fc_{pmax}), also referred to as "best guess forecast" earlier.

The intervals for the integration of the probability density function p_{eps} are defined by δ_{eps}. For fc_{eps} within the interval z_{min1} and z_{max1}, j=1, δ_{eps} is the sum of all weights in this interval. If fc_{eps} lies within the interval z_{min2},z_{max2}, j=2 and δ_{eps} is the sum of all weights within this interval

$$\delta_{eps}(j) = \sum_{i=1}^{n_y} c_{wp}(n) \begin{cases} j = 1 \quad for \quad z_{min1} < fc_{eps}(i) > z_{max1} \\ \\ j = 2 \quad for \quad z_{min2} < fc_{eps}(i) > z_{max2} \end{cases} \tag{8}$$

where $n = 1, 2, ...n_y$.

The intervals are updated in each iteration step and parts of the intervals are cut off, until the function converges to the maximum probability, "best guess", value of fc_{pmax}.

The minima and maxima z_{min1}, z_{max1}, z_{min2},z_{max2} define two intervals of the probability function and thereby reduce the required number of iterations. The pmt-filter function has therefore a second implicit level. The definition of the minima and maxima are:

$$z_{min1} = w_{min}$$

$$z_{max1} = 0.25w_{min} + 0.75w_{max}$$

$$z_{min2} = 0.75w_{min} + 0.25w_{max} \tag{9}$$

$$z_{max2} = w_{max}$$

The pmt-filter function is now used to define the boundaries of the intervals. Thorough test revealed that for $\delta_{eps}(1) = \delta_{eps}(2)$, the coefficients $a_1 = 0.875$ and $b_1 = 0.125$ proved to be most suitable. Hence, for

$$\delta_{eps}(1) = \delta_{eps}(2) \begin{cases} z^*_{min_1} = a_1 z_{min_1} + b_1 z_{max_1} \\ \\ z^*_{max_2} = a_1 z_{max_2} + b_1 z_{min_2} \end{cases} \tag{10}$$

and for

$$\delta_{eps}(1) > \delta_{eps}(2) \begin{cases} z^*_{max_2} = a_2 z_{max_2} + b_2 z_{max_1} \\ z^*_{max_1} = a_2 z_{min_1} + b_2 z_{max_2} \\ z^*_{min_2} = a_2 z_{min_1} + b_2 z_{max_1} \end{cases} \tag{11}$$

and

$$\delta_{eps}(1) < \delta_{eps}(2) \begin{cases} z^*_{max_1} = a_2 z_{min_2} + b_2 z_{min_2} \\ z^*_{min_2} = a_2 z_{min_1} + b_2 z_{max_2} \\ z^*_{max_1} = a_2 z_{min_2} + b_2 z_{max_1} \end{cases} \tag{12}$$

For the second and third set of boundaries, thorough tests have shown that suitable approximations for these coefficients are $a_2 = 0.75$ and $b_2 = 0.25$.

Then the maximum probability fc_{pmax} is updated by updating the intervals/bins of the probability distribution with the new z_{min} and z_{max}, hence

$$fc_{pmax} = max[p_{eps}(n)] \tag{13}$$

fc_{pmax} is computed for the intervals $1, ..n$ and $n, ..ny$, where n is:

$$n = (n_y \cdot \frac{z_{mean} - w_{min}}{w_{max} - w_{min}} + 1)(\epsilon W) \tag{14}$$

where $z_{mean} = \frac{1}{2}(z_{min_1} + z_{max_2})$

The "best guess" forecast with the maximum probability over the time step fc_{pmax} is then:

$$fc_{pmax} = f\widetilde{c_{mean}} \begin{cases} \sum_{i=1}^{n_{eps}} c_{wp}(i) = 0 \\ p_{eps}(n) = 100 \end{cases} \tag{15}$$

or for

$$fc_{pmax} = \frac{fc_{sum}(n)}{sum(n)} \begin{cases} \sum_{i=1}^{n_{eps}} c_{wp}(i) \neq 0 \\ P_{eps}(n) = 0 \end{cases} \tag{16}$$

The distribution of the remaining probabilities is split into an upper and a lower part. The upper part is defined as

$$fc_{pmax} < n_{y,u} \leq 100 \tag{17}$$

and the lower part is defined as

$$0 < n_{y,l} \leq fc_{pmax} \tag{18}$$

The probability distribution is computed for both the upper and the lower part by integrating over the probability bins n_y

$$sum_{eps} = \begin{cases} \frac{100}{\sum_{i=1}^{n-1} p_{eps}(i)} for \quad 1 < n_y \leq n-1 \\ \frac{100}{\sum_{i=n+1}^{n_y} p_{eps}(i)} for \quad n+1 < n_y \leq 100 \end{cases} \tag{19}$$

The index $n = 0..n_y$ is used to reduce the boundaries according to the intervals for which fc_{pmax} was computed. The initial value of sum_{eps} is set to 0.5 times $p_{eps}(n)$.

The last step is the decoupling in time of the weight function from the time window $-k..k$ to the actual time step according to (6). The "best guess forecast" with the maximum probability over the time window -k to k $fc_{pmax}(i,l)$, is now calculated by using the updated c_{wp} and fc_{pmax} from the pmt-filer:

$$\widetilde{c_{wp}}(i,l) = p_{eps}(i) \quad for \quad 1 \leq i \leq n_{eps} \tag{20}$$

$$fc_{pmax}(i,l) = min(|(fc_{eps}(i) \cdot \widetilde{c_{wp}}(i) - fc_{pmax}(i)|) \quad for \quad 1 \leq i \leq n_{eps} \tag{21}$$

The other parameters c_{sum}, fc_{sum}, w_{min}, w_{max}, $f\widetilde{c_{mean}}$, p_{eps} are then re-calculated according to (**??**), (**??**), (**??**), (**??**), (2) and the probability density function $p_e ps(n)$.

If the algorithm is run for a single site, the output contains a series of tables of probabilities for each bin with maximum and minimum percentages. The tables are created for each hour of the forecast length of e.g. 72 hours.

The values for the "best guess forecast", the minimum and maximum and the mean are also tabulated.

If the pmt-filter algorithm should be applied for parameter fields, the iteration needs to be computed in space rather than in time, or of course in both time and space. This means that the algorithm has to be applied in a 2-dimensional or 3-dimensional way. For the case of field calculations, the uncertainty of the forecast needs to be extended into the horizontal space, i.e. if one member is best at one grid point, it is required to be best at the next grid point as well. However, the principle is the same.

To summarise, the *best guess forecast*, which is the forecast with the maximum probability over an increase time interval $(-k, ..i, ..k)$ reflects the concept of the pmt-filter, namely, that it is better to trust a smaller group of forecasts over a longer time window, than to trust in the highest probability in each time step. The ensemble mean usually gives a very skillful, but also relatively smooth forecast. The mean is also often biased by outliers in extreme situations, whereas this is not the case for the *best guess forecast*.

In longterm statistics and when considering parameter fields, it was found that a statistically weighted ensemble mean scores better than the best guess. However, it has not yet been tested to apply the statistical weights as a fixed part of the weight function A to ensure that the skills of the individual members does not destroy the skill of the *best guess forecast* from the pmt-filter. Additionally, the *best guess forecast* showed a clear advantage over the ensemble mean and the statistically weighted mean for all observed extreme events, because it does not take outliers into account. That means, the larger the spread and the larger the uncertainty of the forecast, the more likely will the *best guess forecast* deviate from the mean and the better it scores relative to the mean.

It has been found that the correlation between the prediction skills of the *best guess forecast*, when iterating over -6 to $+6$ hours is highest. This knowledge is used to find patterns, where groups of members perform well over a certain time interval and transform this into weight coefficients. These coefficients are taken into account in the selection procedure of the *best guess forecast*. To improve this procedure and the learning algorithm, it will however be necessary to include long term statistics, climatic or weather specific weights and weights for the skills of the individual ensemble members. When working with multi-scheme ensembles, where the members differ in their physical parameterisations, which naturally differ in their skills at different weather situations, this difference should be taken into account in the weight matrix.

2.2 Sample application of the pmt-filter

In Figure 2 an example of a real world application of the pmt-filter is given. The probability plot displays aggregated wind power production in the the Western part of Denmark on the 17^{th} of March 2005 at 12 UTC and 72 hours ahead with an ensemble of 75 individual forecasts of wind power production. The ensemble data was generated by a short-range multi-scheme ensemble system. The multi-scheme approach is well suited for short-range applications, where the uncertainty lies mostly in the development of the fast physical processes (e.g. Stensrud, 1999, 2000; Möhrlen, 2004).

The white solid line displays the "best guess forecast" derived from the pmt-filter and the dashed line displays the mean of the ensemble. The gray shading on figure 2 shows the probability density of the forecasts and indicates how different forecast cluster together to "groups" of possible outcomes, as described above (Figure 1).

Figure 3 displays a box plot of the same forecast as in Figure 2. Even though the box plot describes the probability density of the forecast with the median, lower and upper quartiles, it can be seen in Figure 3 that the density of the forecast is not always enough information to interpret the evolution of an event over time.

In many real applications, the end-user needs to take actions upon forecasts. In such cases, a plain probability density output with median and quartiles at one point in time does not provide enough information to evaluate their risk. In some cases a user would choose the minimum, the mean, median or the maximum for the safest operation, or least economic loss. One of the largest problems an end user faces is the fact that his decision is to made over a number of hours or days and seldom at one specific point in time. Therefore, it is imperative to many end-users to get a probability distribution over a time range that is optimised to his needs. In many applications the decision is also more complex and manifold than a simple on/off signal. If, for example, the likelihood of a certain event, such as "rain" or "no rain" is changing over the course of the time interval an end-user requires to take action, then the decision process becomes too complicated for a human brain to combine past and future probabilities at each time interval to find the most likely solution.

It is particularly for those cases that the pmt-filter provides the necessary information. It can be seen in the example (Figure 2) that the path of the "best guess" derived from the pmt-filter (black solid line) deviates from the median especially when the uncertainty increases, i.e. when the ensemble spread increases. Because the pmt-filter searches for "groups" of possible outcomes forward and backward in time at each time step and selects by taking the past and the future into account, it is simulating the way an operator would evaluate the situation, but with more information than the operator could possibly use. It is interesting to see in this respect that the difference between the mean of the ensemble and the "best guess" is smallest when the uncertainty is small and largest, when the uncertainty is high (between 10 UTC to 00 UTC on the third day of the forecast).

The forecast in Figure 2 demonstrates this principle. In the first 24 hours from 12UTC to the next day, the uncertainty of the forecast is rather small and a decision making process is easy. After 12 UTC on the second day, however, it becomes more difficult to take decisions. When following the individual "groups" of possible solutions and their change over time a statistical filtering becomes absolutely necessary. This is when the pmt-filter is most beneficial, as it selects the "best guess" in each time step according to constraints made by the end-user. If no constraints are given, it searches for the most logical outcome with help of statistical parameters, i.e. the probability density is followed forward and backward to take the most likely path. In the next section, using an application to an extreme event, it will be shown how the constraints can be defined and how those constraints affect the results.

3 Application of the pmt-filter: The Danish Storm in January 2005

On the 7^{th} of January 2005 the hurricane ERWIN moved from the British Isles towards the southern part of Norway (German Weather Service , 2005 and DMI, 2005) and reached Denmark on the 8^{th} of January (see Figure 4).

The Danish Meteorological Institute (DMI) declared the storm as the best predicted storm in several years (DMI, 2005), The Danish transmission system operator's (TSO) interpretation of the forecast from their operational forecasting tool for wind power did not give the same information and hence resulted in the assumption that the storm will not affect the wind turbine production. If wind speeds of more than 25 m/s are measured over a time span of more than 15 min, wind turbines switch off production. In the wind energy community this is referred to a "cut-off".

The pmt-filter is applied to this event and to demonstrate the usefulness of an ensemble forecasting system in risk assessment. This example also demonstrates the necessity and importance of interpretation tools for the uncertainty in weather (ensemble) prediction, such as the pmt-filter.

Figure 5 shows a probability graph generated with the pmt-filter of a wind power production forecast over 72 hours of the western part of Denmark. The power production is given in % of installed wind power capacity. A generation of 100% would correspond to a power production of 2900 MW.

¿¿From the ensemble spread in Figure 5, it can be seen that some of the ensemble members did not reach wind speeds higher than 25 m/s, which causes wind turbines to switch off ("cut-off"). Hence, the operational forecast from the TSO was within the ensemble spread and the operator in the TSO had very little possibility to quantify the risk of a "cut-off" from the single deterministic forecast as it happened.

The ensemble forecasts however indicate that there was a risk for "cut-off" of large amounts of wind power. The ensemble hence adds value not only because the "best guess" (white solid line in Figure 5 computed by the pmt-filter) of the ensemble forecast indicates a significant drop of power production, but in addition because the end-user gets the possibility to evaluate and foresee the risks arising from large scale "cut-off" of wind power.

Today, risk analysis is becoming a more and more important issue under such conditions, because focus is no longer only on the danger to life, but also on insurance and liability disputes, which are one of the side effects of liberalisation of modern globalised economy (Carpenter, 2005).

This example demonstrates clearly that in such events, not only the weather forecasts, but also wind power forecasts derived from an ensemble of weather forecasts, are imperative information in such events. Part of the risk assessment is in this case the capability to evaluate the likelihood of reduced transmission capacity from wind turbine "cut-offs", damages on the electricity lines and reduced consumption on the grid.

With wind speeds above 40 m/s on the west coast of Denmark (see Figure 6), there was a serious risk of damages that would affect the operation of the electrical grid. The result was in fact that part of Denmark was without electricity for several hours and smaller fractions were without electricity for 2 days. In Denmark and the Northern part of Germany, such storms happen frequently and are a risk for the security of the electricity supply, because the wind power production exceeds at times the consumption. Such examples are the hurricanes Anatol and Lothar in 1999, Kerstin and Liane in 2000, Anna and Jannett in 2002 and Erwin, Ulf and Dorian in 2005. Most of these hurricanes move southwest or northwest of the British Isles towards Denmark and the northern part of Germany and the scale of these hurricanes are often difficult to predict.

The 8^{th} of January storm was on a relatively large scale and moved slower than previous storms. Its center was also not close to Denmark. However, for the TSO in the western part of Denmark, the 8^{th} of January storm was difficult to predict for two particular reasons. The recovery of wind turbines that had switched off, because of too high wind speeds, was unpredictable. This can be seen on the observations in Figure 4 (black dotted line). Older wind turbines often need to be manually restarted, thus, the turbines did not recover as the wind dropped below 25 m/s. Secondly, according to the prediction, the entire area had either full power production from the wind turbines or no power production.

It was however not very likely that the power production would result in something like 50%, because the wind was more correlated horizontally over the entire area than in most other storms. This was due to the spatial extend of the low pressure system in 2005 (see also figure 6). The wind speed peaked only at gale force at the northern part of Denmark (Hanstholm) with a sharp cold front passage at 16 hours UTC (see Figure 6). As Figure 4 demonstrates, this was forecasted extremely accurate by the ensemble mean and the “best guess” from the pmt-filter.

Risk assessment of weather is therefore no longer a meteorological problem, but has also significant impact on other aspects of life, such as the electricity supply. The same could also be demonstrated on many of the flooding events around the world.

4 Summary and discussion

A probabilistic multi-trend filter has been developed. The pmt-filter provides an alternative to previously developed ensemble classification methods (Atger, 1999, Yussouf et al., 2004, Ziehmann, 2001). In comparison to the Euclidean distance dissimilarity and Ward’s method for hierarchical clustering used by Yussouf et al. (2004) and Roulston et al. (2003), or the tubing approach (Atger 1999), and skill prediction as shown by Ziehmann (2001), the pmt-filter is not only a classification approach for verification of an ensemble, but has a selecting procedure inherent to generate a “best guess” from the distribution of the ensemble members.

As shown in the schematic of the pmt-filter in Figure 1, the strength of the pmt-filter lies in its applicability to practical applications, such as forecasting of wind power with an ensemble of forecasts.

In fact, there are a number of possibilities to generate statistical “best guesses” with the pmt-filter. One possibility has been shown by the example of a storm event, i.e. to rank the ensemble members at the forecast start with long-term statistical weights. As discussed in section **??**, other possibilities are to generate weather dependent weight coefficients. At present, phase adjustments of fronts are tested when online observations are available. Another possibility is to use bias corrections for each ensemble member. In both cases the adjustments can be taken into account for the future states and can be reduced at each time step with a function suitable for the problem by updating the weight matrix A.

Further developments will include two space dimensions and a second time dimension. The potential of this method is that it allows for previous forecasts and observations to be taken into consideration in the ensemble evaluation. However, members from such “older” forecasts should not be allowed to control the time evolution, unless they show good agreement with the "newer" forecasts or the observations. The "older" forecasts will also only be accepted as likely, if they follow a group of "newer" forecasts.

The strength of this approach is that it evaluates the data automatically in a manner that is similar to an experienced human operator, as it filters out the poor forecasts and thereby provides more accurate and "real" probabilities. The pmt-filter therefore also contains a inherent learning algorithm, when online observations are present.

It is also an efficient way to selectively reduce the amount of data. When compared to a set of static weight coefficients, it is argued that the potential improvement of the pmt-filter is higher when all possible members (also from previous forecasts) are taken into account with at least 1% weight.

To conclude, ensemble forecasts add essential information for the evaluation of extreme events. The forecasts alone are however not always sufficient information for decision making. The application of the pmt-filer as an alternative to standard statistical parameters allows operators, civil protectors, disaster management and control centres to evaluate a critical situation some time prior to the event and to make necessary arrangements in the case of extreme events. Here the pmt-filter has been applied to an extreme event in the electricity industry with wind power on their grid. However, the application of the pmt-filter is by no means limited to such applications. It can be applied to any decision making procedure that involves weather parameters, such as problems of flooding, hurricanes, weather derivatives, economics, stocks, to name only a few applications.

References

Alhamed, A., Lakshmivarahan, S.& Stensrud, D. J. (2002). Cluster analysis of multimodel ensemble data from SAMEX. *Monthly Weather Review*, *130*, 226-256.

Atger, F. (1999). Tubing: an alternative to clustering for the classification of ensemble forecasts, *Weather and Forecasting*, *14*, 741-757.

Carpenter, G. (2005): Windstorm Erwin / Gudrun January 2005. *Specialty Practice Briefing, An update from the Property Specialty. Issue No. 2* http://www.dmi.dk/dmi/8januarstormguycarprapport.pdf

DMI: Danmark ramt af landsdkkende storm. 10. januar 2005 http://www.dmi.dk/dmi/danmark_ramt_af_landsdaekkende_storm

Farina, L. (2002). On ensemble prediction of ocean waves. *Tellus A*, *54*, 148-158.

Georgakakos, K. P., Seo, D. J., Gupta, H., Schaake, J., & Butts, M. B. (2004). Towards the characterization of stream flow simulation uncertainty through multi model ensembles. *Journal of Hydrology*, *298*, 222-241.

German Weather Service (Deutscher Wetterdienst): Der Orkan ERWIN am 8. Januar 2005 http://www.dwd.de/de/FundE/Klima/KLIS/prod/spezial/sturm/Orkan_ERWIN_am_8_Januar_2005.pdf

Gneiting, T. & Raftery, A.E. (2005). Weather forecasting with ensemble methods. *Science*, *310*, 248-249.

Gong, X. & Richman, B. (1995). On the application of cluster analysis to growing season precipitation data in North America east of the Rockies. *Journal of Climate*, *8*, 879-931.

Gouweleeuw, B. T., Thielen, J., Franchello, G., De Roy, A. P. J. & Buizza, R. (2005). Flood forecasting using medium-range probabilistic weather prediction. *Hydrology and Earth System Sciences. 9*, 365-380.

Gupta, H., Sorooshian, S., Gao, X. G., Iman, B., Hsu, K. L., Bastidas, L., Li, J. L., & Mahani, S. (2002). The challenge of predicting flash floods from thunderstorm rainfall. *Philosophical Transactions of the Royal Society of London Series A*, *360*, 1363-1371.

Lang, S., Möhrlen, C., Jørgensen, J., Ó Gallachóir, B. O. & McKeogh, E. (2006). Application of a Multi-Scheme Ensemble Prediction System for wind power forecasting in Ireland and comparison with validation results from Denmark and Germany. *Proceedings of the European Wind Energy Conference, Athens 2006*.

Legg, T. P., Mylne, K. R., & Woolcock, C. (2002). Use of medium-range ensembles at the Met Office I: PREVIN - a system for the production of probabilistic forecast information from the ECMWF EPS. *Meteorological Applications*, *9*, 255-271.

Liu, Y., Wang, Y., Li, Y., Zhang, B. F., & Wu, G. F. (2004). Earthquake prediction by RBF neural network ensemble. *Advances in Neural Networks - ISNN 2004*, *3174*, 962-969.

McIntyre, N., Lee, H., Wheater, H., Young, A., & Wagener T. (2005). Ensemble predictions of runoff in ungauged catchments. *Water Resources Research*, *41*, doi:10.1029/2005WR004289.

Molteni, F., Buizza, R., Palmer, T. N., & Petroliagis, T. (1996). The ECMWF ensemble system: Methodology and validation. *Quarterly Journal of the Royal Meteorological Society*, *122*, 73-119.

Moradkhani, H., Sorooshian, S., Gupta, H. V., & Houser, P. R. (2005). Dual state-parameter estimation of hydrological models using ensemble Kalman-filter. *Advances in Water Resources*, *28*, 135-147.

Möhrlen, C. (2004). Uncertainty in wind energy forecasting. Ph.D. dissertation (University College Cork, Ireland).

Nakaegawa, T. & Kanamitsu, M. (2006). Cluster analysis of the seasonal forecast skill of the NCEP SFM over the Pacific-North America sector. *Journal of Climate*, *19*, 123-138.

Palmer, T. N. (2002). The economic value of ensemble forecasts as a tool for risk assessment: From days to decades. *Quarterly Journal of the Royal Meteorological Society*, *128*, 747-774.

Pellerin, G., Lefaivre, L., Houtekamer, P., & Girard, C. (2003). Increasing the horizontal resolution of ensemble forecasts at CMC. *Nonlinear Processes in Geophysics*, *10*, 463-468.

Roulston, M.S., Kaplan, D.T., Hardenberg, J., Smith, L.S. (2003). Using Medium Range Weather Forecasts to Improve the Value of Wind Energy Production. *Renewable Energy*, *28*(4), 585-602.

Rundle, J. B., Rundle, P. B., Donnellan, A., Li, P., Klein, W., Morein, G., Turcotte, D. L., & Grant, L. (2006). Stress transfer in earthquakes, hazard estimation and ensemble forecasting: Interferences from numerical simulations. *Tectonophysics*, *413*, 109-125.

Stensrud, D.J., H.E. Brooks, J. Du, M.S. Tracton, and E. Rogers, (1999). Using ensembles for short-range forecasting. *Mon. Wea. Rev.*, *127*, 433-446.

Stensrud,D.J. Bao,J-W. Warner, T.T., (2000). Using Initial Condition and Model Physics Perturbations in Short-Range Ensemble Simulations of Mesoscale Convective Systems. *Mon. Wea. Rev.* Vol. *128*(7), 2077-2107.

Taylor, J. W. & Buizza, R. (2003). Using weather ensemble predictions in electricity demand forecasting. *International Journal of Forecasting*, *19*, 57-70.

Tracton, M. S., & Kalnay, E. (1993). Operational ensemble prediction at the National Meteorological Center: practical aspects. *Weather and Forecasting*, *8*, 379-398.

Vialard, J., Vitart, F., Balmaseda, M. A., Stockdale, T. N., & Anderson, D. L. T. (2005). An ensemble generation method for seasonal forecasting with an ocean-atmosphere coupled model. *Monthly Weather Review*, *133*, 441-453.

Yussouf, N., Stensrud, D.J., Lakshmivarahan, S. (2004). Cluster analysis of multimodel ensemble data over New England. *Monthly Weather Review*, *132*, 2452-2462.

Ziehmann, C., (2001). Skill prediction of local weather forecasts based on the ECMWF ensemble. *Nonlinear Processes in Geophysics*, *8*, 419-428.

Figure captions

Figure 1. Qualitative demonstration of the pmt-filter algorithm for an arbitrary forecast quantity and an arbitrary forecast time. Shown are 6 ensemble members (A, B, C, D, E and F). The ellipsoids mark the grouping procedure.

Figure 2. Example of a probability forecast of aggregated wind power for the western part of Denmark on the 17^{th} of March 2005 generated with the pmt-filter. The dashed white line is the EPS mean, the solid white line is the *best guess* from the pmt-filter. The grey shading displays the probability density in % power production of wind turbines from the ensemble.

Figure 3. Box plot of the 72 hour wind power forecast shown in Figure 2. The solid black line displays the *best guess* from the pmt-filter and the dashed black line is the EPS mean.

Figure 4. Satellite image of the hurricane Erwin on the 8^{th} of January 2005 (source: DMI, METEOSAT-8, http://www.dmi.dk/dmi/index/nyheder/nyheder-2005/danmark_ramt_af_landsdaekkende_storm.htm).

Figure 5. Probability forecast generated with the pmt-filter for the hurricane Erwin on the 8^{th} of January 2005, showing the probability for a large-scale cut-off of wind power production (in % installed capacity) in the western part of Denmark. The graph shows the probability of wind power production from 6 UTC of the 7^{th} of January + 72 hours. The dashed white line is the EPS mean, the solid white line is the *best guess* from the pmt-filter, the white thin dotted line is a second *best guess 2* that takes "user constraints" (long term statistics) into account, the black dashed line shows the observations. The grey shading displays the probability density in % power production of wind turbines from the ensemble.

Figure 6. Measurements of wind speed and wind gusts at the 8^{th} of January over Denmark. The numbers show the highest measured wind speeds as 10-minute averages (upper number) and wind gusts (lower number) throughout the storm. (source: Danish Meteorological Institute, http://www.dmi.dk/dmi/index/nyheder/nyheder-2005/danmark_ramt_af_landsdaekkende_storm.htm)

Figure 2:

Figure 2:

Figure 3:

Figure 3:

Figure 4:

Figure 5:

Figure 6: