ECMWF Feature article

from Newsletter Number 107 – Spring 2006

METEOROLOGY

Recent developments in extreme weather forecasting

www.ecmwf.int/en/about/news-centre/media-resources
doi:10.21957/kl9821hnc7

This article appeared in the Meteorology section of ECMWF Newsletter No. 107 – Spring 2006, pp. 8–17.

Recent developments in extreme weather forecasting

Ervin Zsótér

The forecasting of extreme weather events has gained special emphasis in recent years. The gradual improvement of the NWP models has offered a better treatment of hazardous weather phenomena such as tropical cyclones and squall lines. In addition, an Ensemble Prediction System (EPS), such as that at ECMWF, can deliver more reliable and user-specific warnings of extreme events in a probabilistic way. However the extraction of the extreme weather-related information from the huge volume of EPS data is a complex and difficult task. A possible and efficient way of summarising the available information about extreme weather is to scale the EPS forecast distribution with respect to the model climate.

The Extreme Forecast Index (EFI) has been developed to indicate where the EPS forecast distribution is substantially different from the model climate (see *Lalaurette,* 2002 for more details). It is considered to be useful supplementary information to other EPS products, such as probability maps or EPSgrams. The real advantage of using the EFI lies in the fact that it is an integral measure referenced to the model climate that contains all the information regarding variability of a parameter in location and time. Therefore, the users can recognise the abnormality of the weather situation without defining different space- and time-dependant thresholds.

EFI products and recent developments

More than four years ago the EFI became available to users from Member and Co-operating States via the ECMWF website. In addition, from October 2003 the EFI has been archived in MARS. In the dissemination system and on the ECMWF website the following EFI products are currently available.

- 2m temperature EFI for D+1 to D+5 (computed for each 12 UTC).
- 10m wind speed EFI for D+1 to D+5 (computed for each 12 UTC similarly to 2m temperature).
- Wind gust EFI for D+1 to D+5 (computed for the daily maximum wind gust during each 24-hour period starting at 00 UTC).
- Total precipitation EFI.
  – 1-day (24 hours) accumulation for D+1 to D+5 (computed for each 24-hour period starting at 06 UTC).
  – 5-day (120 hours) accumulation for D+1 to D+5 and D+2 to D+6 (computed for the two 120-hour periods starting at 06 UTC).
  –10-day (240 hours) accumulation for D+0 to D+10 (computed for the whole forecast period).

In the last few years two main changes have taken place in the operational production of the EFI.

- Implementation of a revised formulation of the EFI (operational from October 2003) to make the index more sensitive to infrequent and unusual events located in the tails of the distribution.
- Replacement of the original pseudo model climate used by the EFI with the 30-year EPS control climate based on ERA-40 reanalyses to take into account the EPS model climate in a more realistic way. This used consistent model versions in the EPS forecast and the model climate over a long period.

These developments will now be discussed in more detail.

Revised formulation of the EFI

The first version of the EFI has been criticised for being too sensitive to a shift in the median of the forecast distribution, and for not being sensitive enough to shifts in the tails, where unlikely, but potentially damaging events are located. This characteristic of the EFI is due to it being modelled on the Kolmogorov-Smirnov test. In order to improve the EFI response around the distribution tails, we revised the EFI by introducing an extra weight term into the formulation according to the Anderson-Darling test (*Anderson & Darling,* 1952) – see Box A.

The statistical Anderson-Darling test is a modification of the Kolmogorov-Smirnov test which gives more weight to the tails. Figure 1 illustrates the difference between the original and the revised formulations, given the simple case which occurs when the EPS members have two values: (a) the record-breaking

 doi:10.21957/kl9821hnc7

extreme (Q100, forecast by proportion *r* of all the EPS members, where Qn refers to the n'th percentile of a Cumulative Probability Distribution) and (b) a specified percentile of the climate (results are given for the four cases where Q0, Q50, Q75 and Q90 are forecast by the remaining proportion 1–*r*). The impact of using the revised formulation is obvious. It results in higher values, demonstrating that the presence of EPS values in the tail of the distribution is sufficient to keep the EFI values high, even if other EPS members are close to the median (Q50).

Figure 2 gives an example of a 2m temperature EFI from October 2005 when the weather in Western Europe was affected by an intense warm spell. This case illustrates the relation between the original and revised EFI formulations. The revised EFI covers a significantly larger area and also tends to slightly decrease the maximum, compared to the original version. By analysing the tail characteristics in the area highlighted by the EFI we could also conclude that the revised EFI clearly gives a better indication of the existence of climatologically extreme EPS members (see also Figure 10(a) which will be discussed later).

Revised formulation of the Extreme Forecast Index A

The original Extreme Forecast Index (EFI_n) is defined as follows:

$$EFI_n = -(n+1)\int_0^1 \left(p - F_f(p)\right)^n dp \quad \text{if } n \text{ is even}$$

$$\text{and } \int_0^1 F_f(p)\,dp > \tfrac{1}{2}$$

$$EFI_n = (n+1)\int_0^1 \left(p - F_f(p)\right)^n dp \quad \text{otherwise}$$

where $F_f(p)$ denotes the proportion of EPS members lying below the *p* quantile of the climate record. The pre-operational version of the EFI used $n = 3$.

The revised Extreme Forecast Index (EFI_{AD}) scales departures from the reference climate Cumulative Distribution Function using an extra weight term in the quotient so that:

$$EFI_{AD} = \frac{2}{\pi}\int_0^1 \frac{p - F_f(p)}{\sqrt{p(1-p)}}dp$$

The weight $p(1–p)$ is used because it is a quadratic function that takes its maximum value for $p = 0.5$ and goes to zero at both ends of the probability range. In a different context, this quantity is also known as the "uncertainty" in the Brier Score decomposition and the variance of the binomial distribution.

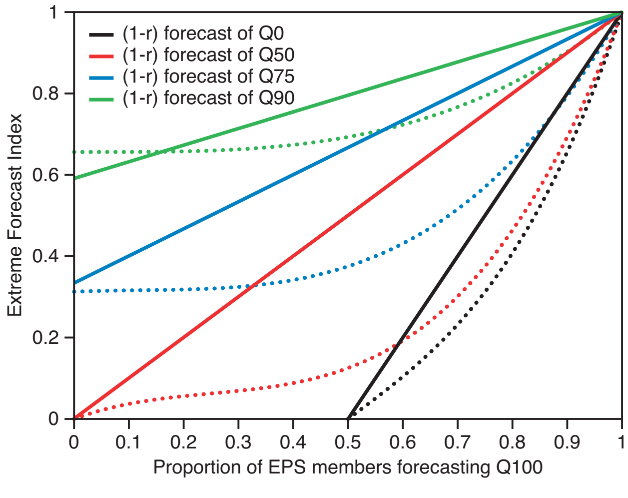

Figure 1 EFI variations for an EPS forecast with two values: a proportion r of EPS members have a forecast of record-breaking values (Q100), and a proportion (1-r) have a forecast of different percentiles of the climate distribution (see legend). Full lines refer to the revised EFI based on the Anderson-Darling formulation, while dotted lines show the original EFI configuration (i.e. the pre-operational version). Consider an example where 60% of the EPS members forecast the climate maximum (Q100). In this case, depending on the forecast of the remaining 40%, the original and revised EFIs (in this order) give the values of 0.1 and 0.2 for Q0 (i.e. 40% of the EPS members forecast the minimum value of the climate – Q0), 0.19 and 0.6 for Q50 (median), 0.43 and 0.73 for Q75 (upper quartile) and 0.72 and 0.84 for Q90 (last quintile).

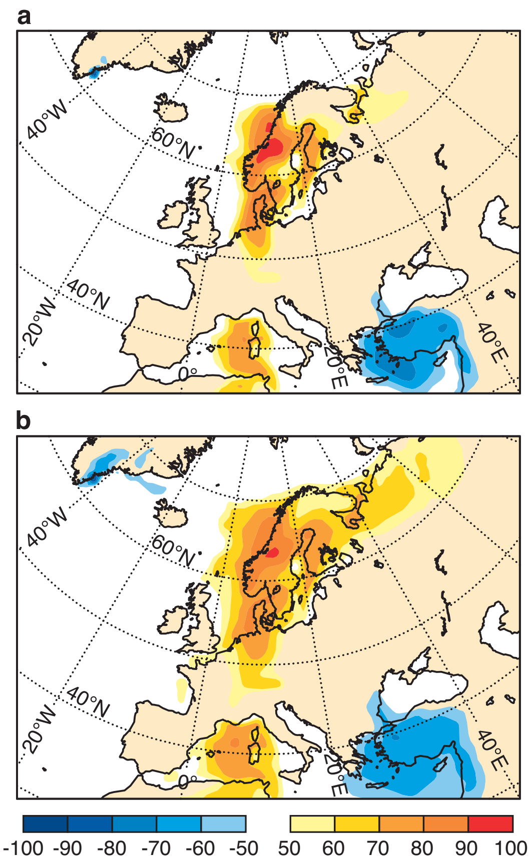

Figure 2 Comparison between (a) original EFI and (b) revised EFI for the t+108 EPS forecast of 2m temperature valid at 12 UTC on 31 October 2005. Positive values are shown with shades of orange (only from 50% to 100%) and negative with shades of blue (only from -50% to -100%).

The new EFI climate

General requirements

For the production of any climate-based forecast product with high quality, it is not sufficient to just have an outstanding forecast model. It is also essential to have a reliable model climate that reflects in every aspect the long-term characteristics of the associated forecast model. For example, different representation of orography or land-see mask in the climate could directly introduce a significant EFI bias over mountainous or coastal areas; this is related to the induced shift between the Cumulative Distribution Functions (CDFs) of the forecast and climate. Besides the requirement for consistent orography and resolution, it is also desirable to have the same physical parametrizations. This will introduce similar model error characteristics and therefore guarantee that any differences in the EFI values are indeed the result of dissimilar forecast situations.

WMO defines a climate as a collection of statistics over three decades (30-year period). For the model climate it is hard to satisfy this requirement, since there are no operational forecasts available for such a long period. Moreover the countless model upgrades that have taken place would inevitably ruin the consistency of such a visionary data set. To use model reanalyses, such as ERA-40, is not an ideal solution either, because of the inconsistency between the model cycles being used to generate the reanalyses and the operational forecasts. The ideal solution would be to rerun the model (preferably for 30 years) based on reanalyses. However, this would make extremely high demands on computational resources, particularly for the EPS, where numerous forecasts would need to be run for each day.

The old pseudo climate

At the time of the EFI construction there was no medium-range model climate available at ECMWF. Also it was not appropriate to develop the system of model re-forecasts since the ERA-40 had only recently been completed at ECMWF in 2003. Therefore a temporary solution, the so-called pseudo climate (*Lalaurette,* 2002), was used for the computation of the EFI. This pseudo climate was based on a relatively consistent set of EPS forecasts. It was generated for each calendar month, and the EFI climate for a particular day was determined by interpolating between pseudo climates of two adjacent months. In the climate files of each month, three calendar months of EPS data were put together. The ensemble analyses and the day-5 and day-10 forecasts were kept from each 12 UTC run of the EPS. They were considered to be different realisations of the model climate. This super-ensemble contained the order of 10,000 fields (~3 x 30 x (1 + 50 + 50) = 9,090).

 doi:10.21957/kl9821hnc7

The new EPS control re-forecast suite

The development of the medium-range model climate for the EFI started a few years ago, and the new system was implemented operationally on 1 February 2006 (with the latest main model resolution upgrade). The idea is based on the previously mentioned criteria of an ideal model climate with sampling from over a sufficiently long 30-year period. However, in order to limit the computational cost, instead of running a full or even partial (only with a few members) set of EPS re-forecast for 30 years and for the whole forecast range, it has been decided to run an EPS control version up to 48 hours. The computational cost is the equivalent of adding only six 10-day EPS control runs per day.

The EPS control re-forecasts have two components, the atmospheric and wave segments. The common EPS post-processing (surface parameters and some pressure levels of some values of potential temperature and potential vorticity) has been used and the time frequency is 6 hours. The data is archived in MARS with the stream specifications of *"Ensemble Forecast Atmospheric Hindcast" = efhc and "Ensemble Forecast Wave Hindcast" = ewhc.*

At the time of writing, the EPS control re-forecast suite runs on a daily basis (instead of a "once per every model upgrade" version), performing 30 model integrations each day from 1971 to 2000, starting from the interpolated ERA-40 analyses of the base day (which is ahead by about three weeks compared to the date of run in order to provide the necessary re-forecast files for the EFI computation, see later). We only prepare the forecasts starting at 12 UTC, having run the model up to 48 hours to be able to provide model estimates of the actual weather conditions for each day. An illustration of how the model climate is produced from the re-forecasts for one day is given in Box B.

The new EPS control model climate has the important characteristic that it provides climate elements (i.e. re-forecast files) for each day of the year; this is similar to the real climate constituted from real observations from a sufficiently long period (30-year). In this sense the EFI climate produced on the basis of these re-forecasts will be the "close as possible" model version of the truth.

The other key characteristic of this new system is that the model configuration used in the re-forecast suite is always the latest in operations, having the current resolution and physics (the only exception is the stochastic physics computation, which is not used in the EPS control re-forecast). Therefore the EFI climate constituted from the re-forecasts will always be as consistent as possible within itself and also with the actual EPS forecasts.

Preparation of the model climate for one day **B**

Here is an example of how the element for one day of the model climate is prepared for 2 of January.

The 30 control runs at 12 UTC on 1 January will provide the model climate for 2 January. The t+24 hour control re-forecasts are taken as the model climate elements for 12 UTC on 2 January for the 2m temperature and 10m wind speed. For the 24-hour wind gust climate (00 UTC on the 2nd to 00 UTC on the 3rd) the wind gust maximum is computed in the re-forecasts from t+12 to t+36. And finally for 24-hour total precipitation (climates for 5-day and 10-day total precipitation are also composed of the daily, 24-hour precipitation climate files) the forecast range of t+18 to t+42 (06 UTC on the 2nd to 06 UTC on the 3rd) is taken into account.

doi:10.21957/kl9821hnc7

The sample size criteria for the EFI climate

In order to develop an appropriate model climate for the EFI, the sample size of the data that constitute the EFI climate had to be considered with great care. To get a good estimate of the mean characteristics (such as the median) of the underlying climate distribution, we can rely on relatively small data sets. However, determining the tails of the distribution, where the rare events lie, based on a small sample could easily result in large sampling noise. In order to adequately sample the tails we need significantly more realisations of the model. For example to sample the 99th percentile we need at least the order of a hundred model reruns, not to mention the case of the hypothetical climate extremes. The new EFI formulation is more sensitive to the tails and moreover the most hazardous weather events are related to the last part of the distribution's tails. Therefore it is important to have climate CDFs which are good estimate of the hypothetical model climate distribution including the upper and lower tails.

Although the EPS control re-forecast system provides the climate for every day, the number of available forecasts for a single day is only 30. In order to create a realistic and detailed description of the underlying daily climate variability we need a significantly larger sample size. Therefore we have to combine many of these daily data sets. However this climate accumulation has to be done very carefully since, by gathering climate files from too longer period, we might sample climate representations of a period with significantly different climate characteristics.

Figure 3 shows how the EFI model climate behaves at a specific location when we combine different numbers of daily re-forecast files. We have tested different samples, from the shortest daily climate to the case when we gather fields from the next ±15 days. The re-forecast elements in the sample are considered as equal representations of the model climate, with no weighting applied at present. The number of files in each tested combination are 30 (±0 day, altogether 30 years times 1 day), 210 (±3 days, 30 years times 7 days), 450 (±7 days) and finally 930 (±15 days). The results presented are for Reading, with a 5-day forecast of 2m temperature from January 2006. Please be aware that in the old pseudo climate only the percentile values are archived, and therefore only these values now plotted. The climate versions based on ERA-40 are relatively similar for most of the probability range; the medians of the distributions cluster close to each other and the EFI values, being computed with the different climates, are close together. However, as expected, much larger differences occur in the tails. This is highlighted by the enlarged upper and lower ends of the distributions. As we increase the sample size by merging from 1 (±0) to 31 (±15) days of re-forecasts, the minimum (maximum) values get systematically lower (higher). At Reading in the middle of January this results in a difference of -4°C and +3°C respectively.

For operational production of the EFI we decided to base the model climate on the 31-day option that provides a distribution consisting of 930 re-forecast values (30 years times 31 days). With this choice we actually keep the original idea of having a month-long EFI climate; however this time they are generated specifically for each day, instead of using interpolation between climate files of different calendar months. The accumulation of model re-forecasts from the 15 days closest to the actual date is probably short enough for most of the areas to sample only the location and time related variability of the daily climate.

Having about 1,000 fields seems to be large enough to reasonably sample the infrequent events in the tails of the distribution. In the old EFI climate configuration, the vast number of EPS forecast fields (almost 10,000) that were used made it possible to have a very small sampling noise. Although it is not possible to simulate the sample size of the old system, the new reference EFI climate has a reasonably smooth spatial behaviour.

Regarding 2m temperature there is hardly any sampling noise, not even in the new climate. Figure 4 shows the maximum fields of the old pseudo climate (Figure 4(a)) and the new reference climate (Figure 4(b)) in January where both fields are in line with the topographical and climatological characteristics. However for wind and precipitation - which have high intrinsic variability - the level of sampling noise is higher close to the end of the distribution tails, especially in the new climate where the sample size is significantly smaller. It is found though, that the fields of the 98th and partially the 99th percentile values show already quite smooth behaviour even for precipitation.

doi:10.21957/kl9821hnc7

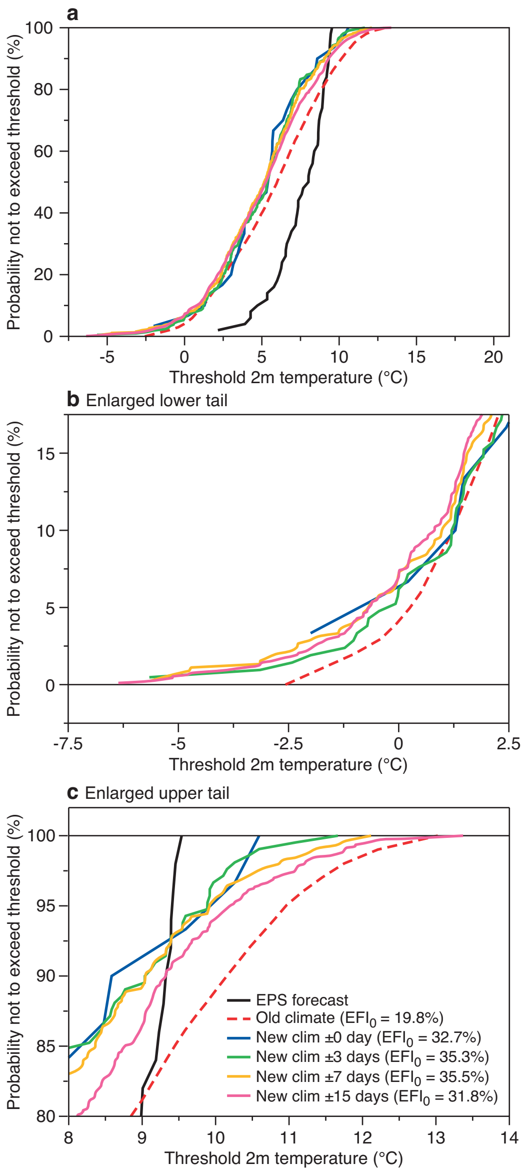

Figure 3 (a) Cumulative Distribution Functions for different EFI model climate distributions for a t+108 EPS 2m temperature forecast from 00 UTC on 16 January 2005 (solid black line) for Reading, UK. Besides the pseudo climate (dashed red line) the new reference climate based on ERA-40 is plotted with different combinations of ±0 (solid blue line), ±3 (solid green line), ±7 (solid orange line) and ±15 (solid magenta line) days of re-forecasts each year from 1971 to 2000. Enlarged versions of the lower and upper tails are given at (b) and (c). The EFI values (denoted by EFI0) computed by different climate references are shown in the legend.

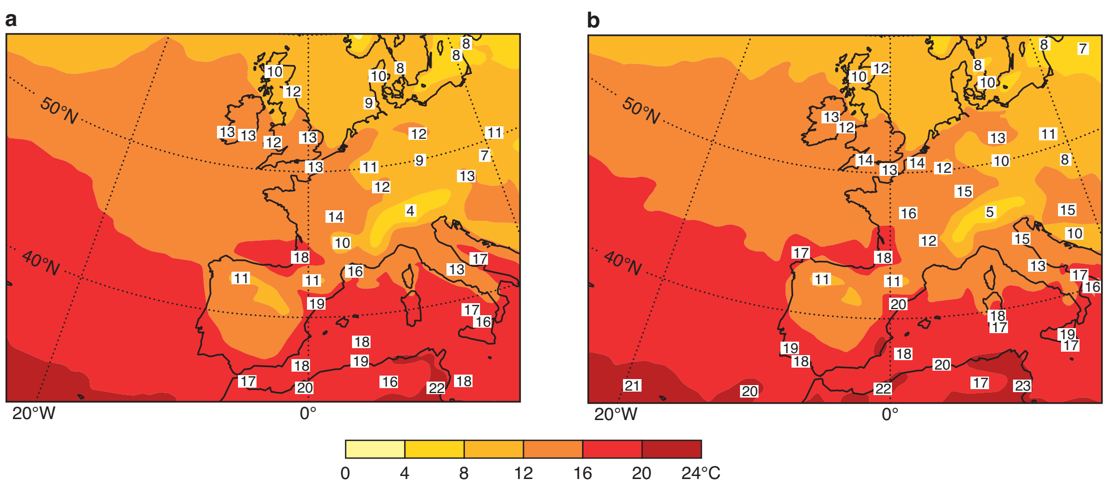

Figure 4 The 100 percentile (Q100) fields for 2m temperature of the (a) old pseudo climate and (b) new reference climate based on ERA-40 for January. The numbers with white blanking represent the local minima and maxima.

Impact of the new reference climate on the EFI

The relation between the original and latest versions of the EFI (i.e. how much they differ from each other) might also be of interest to the users. Since the last change in the EFI affected only the reference climate formulation, the characteristics of the change can be determine directly by examining the difference between the previous pseudo climate and the new reference climate (i.e. ERA-40 based model climate). In Figure 3 we have already seen a specific example for the city of Reading in January. In that case the old pseudo climate was shifted by almost 1°C to the positive side, and the EFI values computed with the old climate was almost 15% lower compared to the EFIs based on the new reference climate. As already indicated, the climate distributions and thus the EFI values can differ significantly depending on location and time. The difference between the EFIs using the pseudo climate and the new reference climate can be rather striking.

As an example, Figure 5 shows the EFIs computed using the two climates for a case from October 2005 (the same case as used for Figures 2). The difference varies significantly from place to place, and areas of potentially extreme weather can appear from nowhere or simply disappear in the EFI using the new climate compared to that using the pseudo version.

To demonstrate the global characteristics of the change we have chosen the example of January for 2m temperature again. Consider the difference between the two model climates as indicated by the deviations in the minimum, median and maximum values of the distributions. The temperature range covered by the new reference climate is larger for most of the locations (see Figure 6). The minimum value of the climate distribution tends to be lower (Figure 6(a)), while the maximum value higher (Figure 6(c)) in the new climate. This is in line with the fact that the ERA-40 based sampling is better by relying on a 30-year period in the new reference climate instead of a 3-year one in the pseudo climate, even though the sample size is drastically smaller in the new system (an order of 1,000 instead of 10,000).

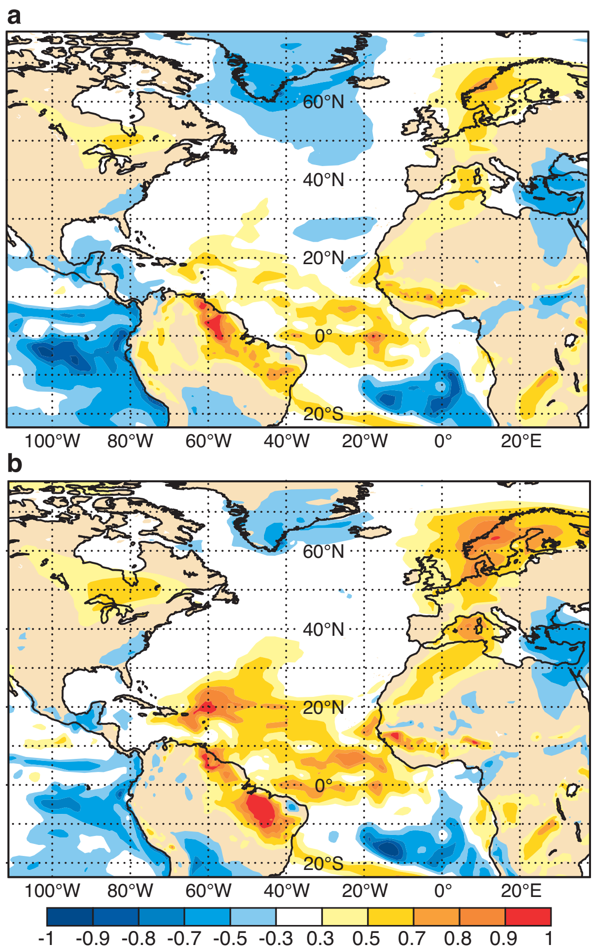

Figure 5 Extreme Forecast Index maps computed for the same t+108 EPS 2m temperature forecast (from 00 UTC on 27 October 2005) using the (a) pseudo climate and (b) new reference climate based on ERA-40. Positive values are plotted by shades of orange (only above 0.3 (30%)) and negative values with shades of blue (only below –0.3 (–30%)).

doi:10.21957/kl9821hnc7

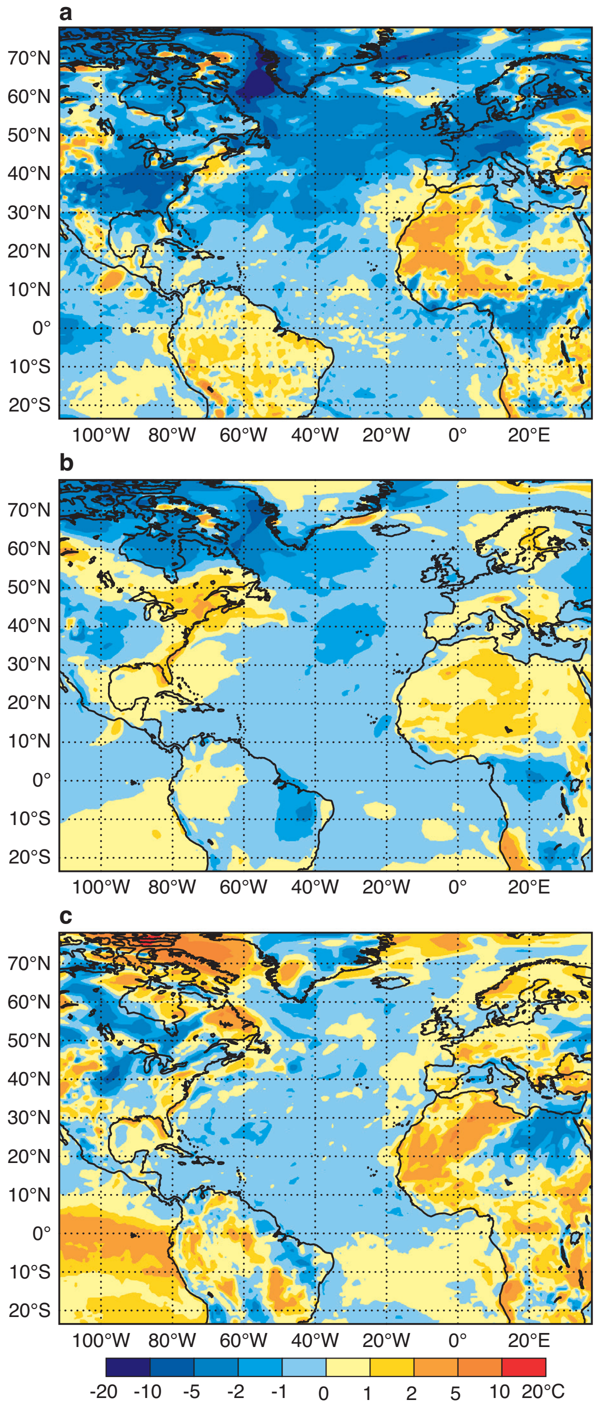

Figure 6 Difference between the new 2m temperature reference model climate and the old 2m temperature pseudo climate for January for the (a) minimum, (b) median, and (c) maximum values of the climate CDFs.

The overall difference between the two model climates can also be measured by the EFI itself (i.e. by computing the integral distance between the CDFs in the new reference climate and the old pseudo-climate). The EFI-measured differences are shown for 2m temperature, 10 m wind speed and 24-hour total precipitation in Figures 7(b), 7(c) and 7(d), this time for July. For 24-hour total precipitation we find significantly larger signals in the tropics; the EFI-measured differences can reach locally even 40-50%, whilst over the extra-tropics the agreement is much stronger where EFI values are usually below 5%. For the 10m wind speed the tropics still hold larger signals, though we find more areas with larger differences outside the tropics. The behaviour of the 2m temperature is different – the distribution of the EFI values is rather homogeneous over the globe. The EFI-measured deviation is above 20% over large areas and the maximum values are not too far from 100%. In addition the effect of the sea seems to be important. Over most of the extra-tropical areas the difference between the two climates are significantly larger.

There are significant differences in the way the two EFI climate versions are generated (50 EPS members with different lead-times within the 10-day period versus EPS control only up to 48 hours). This means that the explanation for the EFI deviations is probably related to the different climate characteristics of the period, from which the climate elements are gathered (3 years in the old pseudo climate versus 30 years in the new formulation). The relation between Figure 7(a) and Figure 7(b) supports this idea. The difference between the old and the new 2m temperature model climates in terms of the EFI shows very similar characteristics compared to the temperature anomaly of the 3-year period in the old climate, referenced to the 30-year period of the ERA-40 reanalyses (1971–2000). Although the agreement is far from perfect, the areas with positive shift in the EFI climate usually correspond to negative temperature anomaly; i.e. colder conditions in the 3-year period (old pseudo climate) compared to the 30-year period (new reference climate), and vice versa.

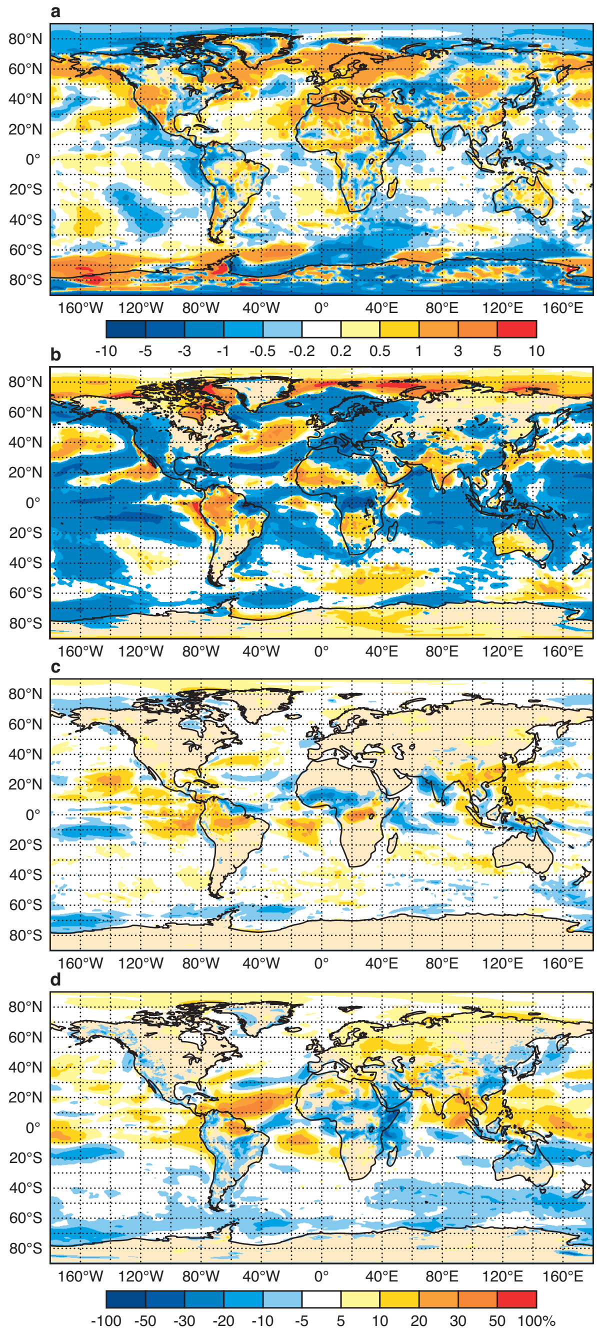

Figure 7 (a) 2m temperature anomaly for July as a three-year average (period of 2003–2005). Operational model analyses are taken as observations, while the reference climate is generated from the ERA-40 reanalyses (1971 –2000). The shades of blue (negative values) indicate colder than normal conditions, and the shades of orange (positive values) warmer ones. (b) Difference between the two model climates as measured by the integral distance (EFI) between the CDFs in the new reference climate and the old pseudo-climate for the 2m temperature in July. Positive values (shades of orange) indicate areas where the new climate has shifted to warmer conditions, while negative (shades of blue) values indicate generally colder conditions in the new reference climate based on ERA-40. (c), (d) The same as in (b) but for the 10m wind speed and 24-hour total precipitation.

doi:10.21957/kl9821hnc7

Ongoing developments

The multi-parameter extreme weather risk map

One of the reasons for the success of the EFI is that the information on parameters directly related to extreme weather (i.e. wind, precipitation and temperature) is provided in the form of a synoptic map without having to decide any arbitrary threshold that would not be relevant on a continental scale or throughout the whole year. The extreme weather risk map is a further step in synthesizing the available information related to extreme weather. Based on the EFI, two main categories of abnormal weather have been defined. The range of EFI values from 50% to 80% is regarded as unusual (warm, windy etc.) and for areas with EFI values above 80% the very unusual or extreme label is allocated (for temperature the negative range is also of equal interest). On the multi-parameter map the areas where one or more of these conditions are met are highlighted by different shadings or markers. We show an example of this combined extreme weather risk interpretation in Figure 8.

The extreme weather risk map is able to summarise the "far from normal" forecast characteristics in a simplified, easily understandable way. However the questions - how was the model behaving in the previous runs and whether the underlying abnormality was forecast in a consistent way – arise from the forecaster's point of view. The answer is given in the distribution plot associated with the extreme weather risk map.

In the background numerous distribution diagrams are prepared for each EFI parameter with the information on the climate and the available forecast distributions, including the corresponding EFI values (see Figure 9). We have extreme weather risk maps for each EFI timestep from the latest EPS forecast, and separate distribution diagrams for each EFI parameter. As the forecast lead-time increases the diagram contains fewer forecast distributions (since currently the EFI is computed only up to D+5). The system currently runs in test mode pre-operationally; the distribution plots are prepared in advance as postscript files. In order to keep the generated volume of data relatively low, it has been decided to create the diagrams based on 10° x 10° gridboxes as the first option to be tested. The gridbox diagram will hold the data from the gridpoint which has the highest EFI value in the shortest available EPS forecast (i.e. the location of the maximum EFI on the corresponding extreme weather risk map). The distribution diagrams can be accessed simply by clicking onto the desired gridbox of the extreme weather risk map.

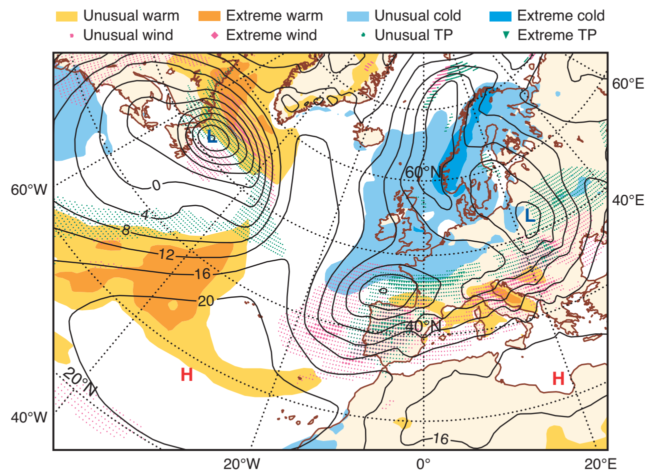

Figure 8 Extract from a global extreme weather risk map. The EFI fields for 2m temperature (different colour shadings), 10m wind speed (magenta markers) and total precipitation (green markers) are combined with the EPS mean of the 1000 hPa geopotential height (black lines). Unusual weather is assigned to EFI values above 50%, while extreme weather conditions are taken to be those above 80%.

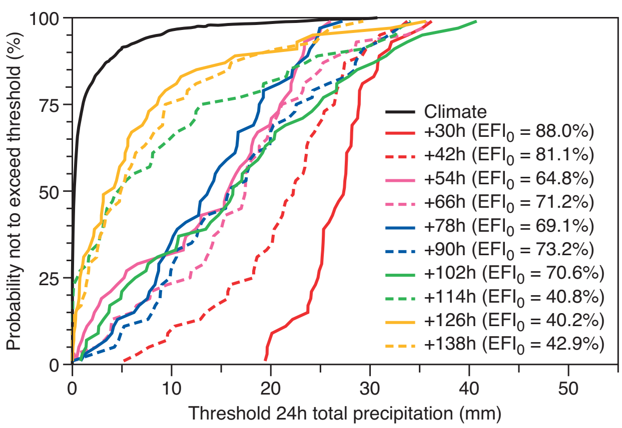

Figure 9 Example of the distribution plot attached to the multi-parameter extreme weather risk map. The thick black line represents the ERA-40 based model climate, while the other lines in different colours show the distribution of EPS runs verifying on the actual date, advancing backward in time. The corresponding EFI values are indicated in the legend for all the EPS forecasts.

doi:10.21957/kl9821hnc7

Additional indices for extreme weather forecasting

The EFI is a useful summary tool to interpret the extreme characteristics of the EPS distribution. Since the revised formulation has been implemented the EFI has become more sensitive to the tails, where the potentially hazardous weather events lie. In addition to using the revised formulation of the EFI, the level of extremity in the tails of the forecast distribution can be explored in a more direct way by measuring the probability distance in the tails. The "Shift in Probability Space" index (SPS) can provide some information on how "extreme" the tail of the EPS distribution is, compared to the climate – see Box C.

The Shift in Probability Space index can complement the EFI with interpreting the abnormality level in the EPS. However, neither the EFI nor even the SPS can give information about some of the characteristics of the EPS tail. There are rare, potentially devastating weather scenarios, when a significant fraction of EPS members predict climatologically (relative to the model climate) unprecedented values. The level of extremity could remain hidden if they are analysed only by either the EFI or SPS indices. In order to measure the abnormality of such situations, we can consider the "Shift of Tails" (SOT) index – see Box D.

"Shift in Probability Space" index (SPS) **C**

The distance in probability space:

$$SPS(p) = F_c(p) - p$$

where $F_c(p)$ is the value indicating where the p quantile of the EPS distribution ranks in the climate record (0 being the absolute minimum, 1 the absolute maximum).

It should be noted that the $SPS(p)$ has a similar form to the quantity present in the EFI formulation given by $p - F_f(p)$ (see Box A) except that in this case it is a function of the EPS forecast rank, and not of the climate rank (in the EFI definition p is the quantile of the climate distribution).

The maximum shift for $SPS(p)$ is $1-p$, when at least a proportion $1-p$ of the EPS is already outside of the climate range. For example in case of the last decile (p = 0.9, Q90) the maximum *SPS* that can be reached is 0.1 or 10%. For total precipitation and wind, only the upper part of the distribution is of real interest regarding extreme weather forecasting. However, for the 2m temperature, the negative shift of the lower distribution fractions is equally informative (for Q10, for example, with an absolute minimum value of –0.1 or –10%).

"Shift of Tails" index (SOT) **D**

The SOT is the scaled distance in the space of the meteorological variable:

$$SOT_{+}(p) = -\frac{Q_f(p) - Q_c(1)}{Q_c(p) - Q_c(1)} \qquad SOT_{-}(p) = -\frac{Q_f(p) - Q_c(0)}{Q_c(p) - Q_c(0)}$$

where $Q_c(0)$ is the minimum, $Q_c(1)$ is the maximum value found in the model climate record, whilst $Q_c(p)$ and $Q_f(p)$ are the p quantiles of the climate and the forecast Cumulative Distribution Functions (CDFs) respectively.

The *SOT+* is the index for measuring the abnormality level in the upper tail, while the *SOT_* is for the lower tail of the CDFs. The *SOT* value of –1 corresponds to the case of $SPS(p) = 0$, when the forecast CDF intersects the climate CDF exactly at the level of p. If $SOT(p) = -1$ for all p, then the climate and forecast CDFs are identical, which is also the case when *EFI* is 0. Nevertheless the SOT index only has a value of 0 if the p quantile in the forecast equals the maximum climate value, $Q_c(1)$.

In addition, if the tail of the EPS is beyond the range covered by the model climate, the SOT becomes positive and is proportional to the distance to the climate maximum in the meteorological variable space. These positive values will indicate climate outliers (EPS members out of the range of the model climate); and the SOT value will also give an indication of the degree of extremity (scaled by the relative distance to the climate CDF). For example $SOT(0.9) = 1$ would mean that at least five EPS members are outliers, all of them exceeding the climate maximum by the value of $Q_c100–Q_c90$ or more.

 doi:10.21957/kl9821hnc7

In line with the SPS, the use of the negative version of the SOT makes sense for the 2m temperature but not for wind and precipitation. Besides, the SOT is referenced to the model climate (as is the EFI), since the quotient depends on the climate range. On the other hand it is a disadvantage that the SOT can take any values. However, in practice the weather conditions which can realistically happen will reasonably bound the SOT. The cases of potentially abnormal weather – which are of interest to the users – will have SOT values close to or above 0% (which actually correspond to the case when a certain proportion of the EPS reaches the climate maximum). It is important to mention that, in order to be able to use a reliable SOT, it is necessary to have a smooth, appropriately detailed model climate, since the SOT is referenced to the tails of the climate.

Figure 10 gives two examples which demonstrate how the SPS and SOT indices can add extra information to the EFI. Figure 10(a) refers to the same case as the one shown in Figures 2. The chosen parameter is the 2m temperature with SPS and SOT values of p = 0.1 (Q10) and p = 0.9 (Q90). As expected, the SPS and SOT in general cover very similar areas as the EFI. However, they also highlight regions, especially the SPS, where the EFI is relatively low. In Figure 10(a), the white triangle in Scandinavia denotes the place of maximum SOT, where the EFI is in fact below its maximum. Figure 11 shows the corresponding forecast and climate distributions at that location. This highlights that, despite of the lower EFI, in this case 25% of the EPS members predicted climate outliers, namely a temperature above 6°C (please be aware that this point has a model orography of approximately 900 m). Although the SPS also takes its maximum value, showing a high risk of extremity, in this situation it would have reached the same value even if the last tenth of the EPS distribution had been 2°C lower.

Figure 10(b) is another informative example representing a wet period from early March 2006. On the western half of the map, the three extreme weather indices are consistent with each other. However, this is not the case for the eastern side. The SPS and SOT identify some sub-regions which are detached from the area of potential extreme weather as shown by the EFI alone. Consider for example the area indicated by the blue box. At the bottom of the box, just on the edge of the 50% EFI range, the SOT indicates more extreme EPS members, whilst in the middle there is no sign of any extreme members, even if the EFI is higher.

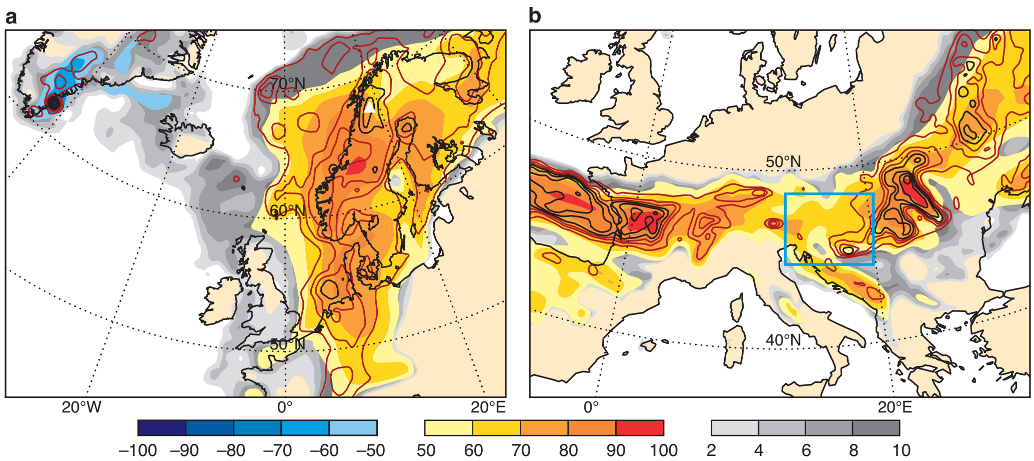

Figure 10 EFI, SPS and SOT indices based on (a) t+108 EPS forecast of 2m temperature valid at 12 UTC on 31 October 2005 and (b) t+06 to t+126 EPS forecast of 120-hour total precipitation valid for the period 06 UTC on 3 March to 06 UTC on 8 March 2006. Positive EFI values are shaded by tones of orange and negative ones by tones of blue (only 2m temperature). The Shift in Probability Space is plotted by grey shades for p = 0.9 and with 2m temperature for p = 0.1, as well (legend is shown only at p = 0.9). The Shift of Tails with p = 0.1 (only temperature) and p = 0.9 are plotted with dashed brown lines for values between 0 and –0.5 (areas where the forecast tail $Q_f(0.9)$ or $Q_f(0.1)$ are close to the climate maximum) , whereas values above 0 (areas where the forecast tail is beyond the climate range) are shown by black thick solid lines. The white triangle in (a) indicates a point in Scandinavia which is the location for the distributions shown in Figure 11. The blue box in (b) is referred to in the text.

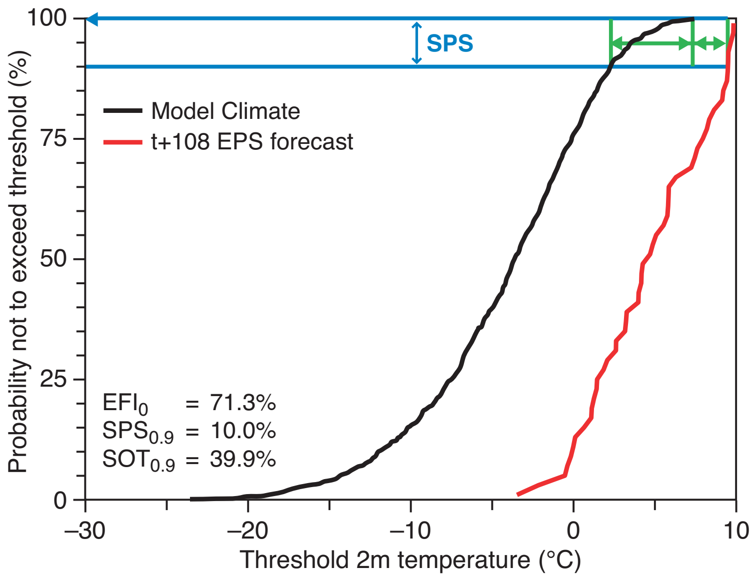

Figure 11 Cumulative Distribution Functions of model climate (black) and t+108 EPS forecast valid at 12 UTC on 31 October 2005 (red) of the 2m temperature at the location where the SOT index has its maximum in Scandinavia (67.7°N, 18.1°E). The location is marked by white triangle in Figure 10. The corresponding values for the EFI (denoted as EFI_0), $SPS_{0.9}$ and $SOT_{0.9}$ are also shown. The blue and green lines illustrate the computation of $SPS_{0.9}$ (blue) and $SOT_{0.9}$ (green).

Conclusions and future work

The demand for reliable early warnings of hazardous weather seems to have increased continuously in recent years. The Extreme Forecast Index together with the proposed Shift in Probability Space and Shift of Tails indices can provide useful information in forecasting extreme events. However, we still need to gather more experience about the capabilities and usefulness of these indices, though evaluation of these products is particularly difficult as there is no comprehensive observed climate available. This means that the definition of extreme events is ambiguous. Regarding the EFI some preliminary verification has been done already, but more detailed, comprehensive evaluation is needed in the future.

In order to make the EFI more reliable, the old pseudo model climate has been recently replaced by the more realistic reference model climate based on ERA-40. Although this formulation already favours the criteria of having a 30-year climate and using the latest model configurations, this system is not yet perfect. We have EPS control re-forecasts only up to 48 hours, so these short-range model forecasts have to be used to constitute the EFI model climate for all the required lead-times. This seems to be a very good approximation for short lead-times of the EPS. However, the EPS variability at larger time steps cannot be simulated with the same accuracy by only sampling from short-range control forecasts.

The effect of different model error characteristics between the control and the EPS members might to some extent deteriorate the similarity between the current EFI reference climate (constituted of 930 EPS control reruns) and the hypothetical EPS model climate. It is important to note that the stochastic physics probably play a major role in developing these differences, since it is an additional factor (not applied to the EPS control) to increase the spread and it is significantly more active for surface parameters. This effect is also probably responsible for the EPS members too often falling outside of the climate range which is reflected in the positive SOT values.

The ongoing project to develop EPS re-forecasts for the calibration of ensemble forecast will have an effect on the EFI climate preparations. Also the development of the ECMWF seamless forecast system (i.e. introducing VAREPS and merging the medium and monthly forecast ranges) will inevitably require further consideration about how to provide forecasts of extreme weather.

As we have shown, at present only the EFI is computed and archived operationally for a few basic surface weather parameters. In addition the extreme weather risk map is still in a pre-operational phase while the other extreme weather indices (SPS and SOT) are only in test status.

doi:10.21957/kl9821hnc7

Based on experience and user recommendations we plan to produce the EFI and selected SOT or SPS fields for some new parameters, such as minimum and maximum temperatures. In addition, consideration will be given to further improving and implementing into operations the extreme weather risk map supported by the related distribution diagrams. The development of new forecast products which are referenced to the model climate are also considered as future activities.

The various developments that are planned will broaden the range of tools that can be used in the field of extreme weather forecasting.

The developments described in this article build upon the work carried out by François Lalaurette during his period at ECMWF.

Further reading

Anderson, T. & **D. Darling,** 1952: Asymptotic theory of certain goodness of fit criteria based on stochastic processes. *Ann. Math. Statist.,* **23,** 193–212.

Lalaurette, F., 2002: Early detection of abnormal weather conditions using a probabilistic extreme forecast index. *Q. J. R. Meteorol. Soc.,* **129,** 3037-3057. Also available as *ECMWF Tech. Memo. No. 373,* 27 pp.

doi:10.21957/kl9821hnc7