WORLD METEOROLOGICAL ORGANIZATION

COMMISSION FOR BASIC SYSTEMS

MEETING OF THE CBS EXPERT TEAM ON ENSEMBLE PREDICTION SYSTEMS (ET-EPS)

GENEVA, SWITZERLAND, 14-18 NOVEMBER 2011

FINAL REPORT

(Back Row) Jean Nicolau, Ken Mylne, Pierre Eckert, Peter Chen, Masayuki Kyouda, Gerhard Wotawa

(Front Row) Young-Youn Park, Alice Soares, André Méthot, Jing Chen, David Richardson, Yuejian Zhu

EXECUTIVE SUMMARY

The meeting of the CBS Expert Team on Ensemble Prediction Systems (ET-EPS) took place at the WMO Headquarters, in Geneva, Switzerland, from 14 to 18 November 2011, under the chairmanship of Mr Ken Mylne (UK).

The meeting reviewed recent progress in implementing global EPS, regional EPS, convective-scale EPS and specialized EPS (e.g. Typhoon EPS). Significant advances have been made by several WMO Members in EPS, whose outputs provide the most important basis for a probabilistic approach to weather forecasting, in all time scales, including short-range forecasting. These products, used in conjunction with high-resolution deterministic NWP outputs, represent an enhanced forecasting strategy, especially for predicting severe weather events.

While a limited number of GDPFS centres run EPS operationally, many NMHSs are exploring various applications, and others are seeking to build capacity for their forecasters to access and effectively use EPS products in their forecasting process as well as to deliver services that are based on probabilistic forecasting methods. In particular, EPS application to support early warning of severe weather is of the highest priority, in contributing to disaster risk reduction. In this context, the meeting reviewed recent progress of EPS applications and integration into operational weather forecasting, post-processing and calibration, and the propagation of weather forecast uncertainty into specialized applications (e.g. EPS in atmospheric transport and dispersion, etc.). Collaboration with Public Weather Services (PWS) on communicating forecast uncertainty to users, with emphasis on disaster risk reduction and to ensure effective pull-through of the benefits to societal impact, was discussed.

The meeting was briefed on key research activities of THORPEX-TIGGE, achieved and anticipated, and their implications and benefits for operational EPS production, access and use. It developed the EPS operational requirements for further research and studies.

The application of EPS to predict severe or high-impact weather events is among the most important topics, including in particular the propagation of the weather forecasts into impact models. Operational flood forecasting systems, including coastal flood forecasting due to waves and storm surges, are increasingly moving towards the adoption of EPS to drive their predictions. The availability of several global ensemble weather prediction systems has been providing an opportunity to explore new dimensions in flood forecasting and the potential to provide early warnings. The Team reviewed recent progress, on flash floods and coastal flood forecasting and other specialized applications such as Atmospheric Transport and Pollution.

Many NMHSs have difficulties in making use of ensemble fields made available in real-time by the producing centres for downloading. In particular, many NMHSs, especially those of developing countries and LCDs, do not have the computing facilities or capability to generate products from ensemble fields. Therefore, EPS continues to be a critical component of the guidance data provided to SWFDP subprojects from the global centres. The subprojects in Southern Africa, the South-west Pacific and in Eastern Africa have been supplied with EPS data by ECMWF, the Met Office UK and NCEP. The subproject in Southeast Asia, planned to commence in April 2012, will be supplied with EPS data by CMA, JMA, KMA and ECMWF (products available for all WMO Members). The use of EPS data is also included in plans for the SWFDP subproject being initiated in the Bay of Bengal. The meeting agreed that the training workshops carried out as part of the SWFDP subprojects have proven one of the most effective ways of providing training on the use of EPS, as the training has been provided in conjunction with ongoing access to operational EPS data.

The meeting reviewed and updated the *Guidelines on EPS and Forecasting*, which are provided in Annex IV to this report. The meeting also reviewed the status of implementation of the Lead-Centre for EPS Verification.

Noting that a first version of the new *Manual on the GDPFS* (WMO-No. 485) will be presented to CBS-XV, in 2012, for consideration, the meeting agreed that any proposed amendments to the *Manual on the GDPFS* should be reflected in the new version. The meeting requested the chairperson, in coordination with the Expert Team members and in liaison with the chairperson of the CBS Coordination Group on Forecast Verification, to review the draft text for the new Manual, and submit to the Secretariat a revised version by mid-March 2012.

The meeting reviewed the team's Terms of Reference (ToR) as adopted at CBS-Ext.(10), and proposed revised ToR as given in Annex VI.

GENERAL SUMMARY OF THE WORK OF THE SESSION

1. OPENING

1.1 The meeting of the CBS Expert Team on Ensemble Prediction Systems (ET-EPS) was opened by its chairperson, Mr Ken Mylne (UK), at 09.30 hours on Monday, 14 November 2011, at the WMO Headquarters, in Geneva, Switzerland. Mr Mylne introduced Mr Geoffrey Love, Director of the Weather and Disaster Risk Reduction Services (WDS) Department of WMO, to address the meeting.

1.2 Mr Love, on behalf of the Secretary-General of the WMO, Mr Michel Jarraud, welcomed the participants to the meeting, to WMO and to Geneva. He noted that, while a limited number of GDPFS centres run EPS operationally, many NMHSs are exploring various applications, and others are seeking to build capacity for their forecasters to access and effectively use EPS products in their forecasting process as well as to deliver services that are based on probabilistic forecasting methods. Mr Love pointed out that many NMHSs have difficulties in making use of ensemble fields made available in real-time by the producing centres for downloading. In particular, many NMHSs, especially those of developing countries and LDCs, do not have the computing facilities or capability to generate products from ensemble fields. Therefore, EPS continues to be a critical component of the guidance data provided by global centres through the Severe Weather Forecasting Demonstration Projects (SWFDPs). He highlighted that the training workshops carried out as part of the SWFDPs have proven one of the most effective ways of providing training on the use of EPS, as the training have been provided in conjunction with ongoing access to operational EPS data. In addition, GIFS-TIGGE has been providing new promising products for near-operational trial, via relevant RSMCs, using the SWFDP framework. Mr Love stressed the need for an effective communication of probabilistic forecasts and uncertainty to users, with emphasis on disaster risk reduction and to ensure effective pull-through of the benefits to societal impact.

1.3 The meeting recalled that, as part of its Terms of Reference (ToR) adopted at CBS-XIV, in 2009, the ET-EPS should develop specifications for the introduction of probabilistic information into products from RSMCs.

1.4 Mr Mylne added his welcome to participants in the meeting, and provided introductory remarks regarding the work of the Expert Team. He recalled that ECMWF and NOAA/NCEP have been running EPS for almost 20 years, while the Expert Team has started its activity in 2001. He noted that there was a very heavy programme of work for and challenging deliberations at this meeting, especially regarding the applications and integration of EPS in combination with deterministic NWP into operational weather forecasting, the practical application of EPS in existing projects (e.g. SWFDPs), and the development of guidance for forecasters. He stressed the need for developing guidance that is adequate for application by the different types of WMO membership.

2. ORGANIZATION OF THE MEETING

2.1 Adoption of the agenda

2.1.1 The meeting adopted the agenda, which is found in Annex I.

2.2 Working arrangements

2.2.1 The meeting agreed on the organization of its work, including the working hours. All pre-session documents can be found via the Documentation Plan (INF. 1) which is posted on the WMO Web site linked to the banner for the meeting at:

http://www.wmo.int/pages/prog/www/BAS/CBS-meetings.html

2.2.2 Noting that a number of participants were new to the ET-EPS, they briefly introduced themselves, to facilitate interactions throughout the meeting. The list of participants is provided in Annex II.

3. REPORT ON THE OUTCOMES OF CBS-Ext.(10) AND CG-XVI RELATED AND/OR RELEVANT TO THE ET-EPS

3.1 The meeting was presented with a report on the outcomes of WMO Constituent Bodies, particularly focused on the 2010 extraordinary session of the Commission for Basic Systems (CBS-Ext.(10)) and the Sixteenth World Meteorological Congress (Cg-XVI), relevant to EPS. The meeting was informed that, while noting significant advances made by several WMO Members in operational EPS, Cg-XVI encouraged:

(a) WMO Members to develop probabilistic forecasting to support early warning of severe and high impact weather as among the highest priority, including propagating probabilistic forecasts of important meteorological and other weather-related parameters into impact models, such as in flash floods and coastal flood forecasting, in contributing to disaster risk reduction.

(b) Suitable training be provided to trigger a fundamental change in thinking by both weather forecasters and users (e.g. disaster management organizations) whereby alerts of severe weather would become more probabilistic in nature to represent the risks associated with severe and high-impact weather. Cg-XVI requested CBS and other relevant technical commissions to collaboratively address this issue. CBS-Ext.(10) agreed that periodic training of weather forecasters is needed in the use of EPS products, specifically in the prediction of dangerous weather phenomena, to improve warning programmes and services.

3.2 In this context, the meeting agreed that the annual training events carried out within the SWFDP framework provide an appropriate framework for contributing to capacity building and for helping developing countries to make best possible use of existing EPS products for improving warnings of hazardous weather conditions and weather-related hazards. At the same time, the meeting recalled that CBS-Ext.(10) recommended to investigate the feasibility of modelling a Virtual Laboratory for SWFDP along the lines of the CGMS-WMO Virtual Laboratory for Satellite Meteorology. The meeting noted that this issue will be addressed at the upcoming meeting of the SG-SWFDP, in February 2012.

3.3 The meeting noted that CBS-Ext.(10) suggested that GDPFS and PWS collaborate in addressing effective communication to users of the meaning of low probabilities of high-impact events and of warnings, which should become more probabilistic in nature to represent the risks associated with high-impact weather and related phenomena.

3.4 The meeting agreed to discuss these issues under the relevant agenda items.

4. PROGRESS OF EPS IMPLEMENTATIONS

4.0 The meeting noted that significant advances have been made by several WMO Members in EPS, whose outputs provide the most important basis for a probabilistic approach to weather forecasting, in all time scales, including short-range forecasting. These products, used in conjunction with high-resolution deterministic NWP outputs, represent an enhanced forecasting strategy, especially for predicting severe weather events, with longer forecast lead-times. The meeting reviewed recent progress (since its previous meeting, in October 2009) in implementing global EPS, high-resolution regional EPS, convective-scale EPS and specialized EPS (e.g. Typhoon EPS). Full reports are available on the WMO Web site at http://www.wmo.int/pages/prog/www/DPFS/Meetings/ET-EPS_Geneva2011/docPlan.html. A summary of the major progress achieved in implementing EPS by EPS centres and their future plans is given in Annex III.

4.1 Global EPS

4.1.1 The meeting noted that the Global Ensemble Prediction Systems (GEPS) have improved during the last couple of years (since its previous meeting, in 2009), in terms of resolution, ensemble size, length of integration and frequency of forecast cycles. The horizontal resolution has increased to 32-70km for most of the EPS Global Producing Centres, while Météo-France has implemented a system with variable horizontal resolution (15km over France to 90 km at the antipodes (e.g. over New Zealand)). The number of vertical levels has also increased to between 28 and 70 levels. Most of the centres increased ensemble size and most systems have more than 20 members. The length of integration is 10-15 days at most centres. At most of the centres, the frequency of the outputs is 6 hours, however for short-range forecasting, some centres provide outputs every 3 hours.

4.1.2 The meeting noted that while singular vectors and Breeding Vector (BV) methods are still used in generating initial perturbations, an increasing number of centres are now implementing variations on the Ensemble Kalman filter (EnKF) including the Ensemble Transform Kalman Filter (ETKF) and Ensemble Transform and Rescaling (ETR) methods. Several centres are now integrating their ensemble generation and data assimilation systems using Ensemble Data Assimilation (EDA) systems or hybrid data assimilation systems in which the ensemble is used to provide background error statistics for data assimilation.

4.1.3 The meeting noted that models error uncertainties are simulated by a range of stochastic physics perturbations, including Stochastic Kinetic Energy Backscatter (SKEB), Random Parameters (RP) and Stochastic Total Tendency Perturbations (STTP). Some systems use a Multi-Model or Multi-Physics approach and Perturbations of Physics Tendencies (PPT).

4.1.4 A summary of the main characteristics of the Global Ensemble Prediction Systems implemented at the participating centres at this meeting (as of November 2011) is provided in Table 4.1.

Table 4.1 – Summary of main characteristics of the Global Ensemble Prediction Systems implemented at the participating centres at this meeting (as of November 2011)

	ECMWF, Europe	Meteo-France	UK Met Office	KMA, Korea	CMA, China	JMA, Japan	NCEP, USA	CMC, Canada
Name		PEARP	MOGREPS-G /15		GEPS	One-week EPS	GEFS	GEPS
Model Name		ARPEGE	UM	UM		GSM	GFS	GEM
Assimilation Method	4D-Var	4D-Var	Hybrid 4D-Var	4D-Var	SSI	4D-Var	GSI	196 members EnKF
Horizontal Resolution	T639/T319 (32/64 km)	Variable TL538 with a stretched coeff of 2.4 (15km→90 km)	N216 (60km)	N320 (40km)	T213 (60km)	T319 (60km)	T190 (70km)	66km
Vertical Resolution (levels)	62	65	70	70	31	60	28	40
Initial Times	00,12	06, 18	00, 12 (From 2012 M-G also 06, 18) [2)]	00,12	00,12	12	00, 06, 12, 18	00, 12
Lead Time	15 days	72h (06), 108h (18)	M-G 3d M-15 15d	10 days	10days	9 days	16 days	16 days
Output Frequency	3h to 144, then 6h	3h to 54h then 6h	M-G 3h M-15 6h	6h	6h to 120h,12h	3,6,12-hour	6 hours	6 hour

					to 240			
No. of Members (+control)	50+1	34+1	23+1 (From 2012 M-G 11+1) 2)	23+1	14+1	50+1	20+1	20+1
Coupled Ocean	From D10	No	No	No	No	No	No	No
Multi-Model	No	No	No	No	No	No	No	Yes
Initial Perturbations	SV+EDA	SV+EDA	Localized ETKF	Localized ETKF	BGM	SV	ETR	EnKF
Model Perturbations	SKEB+ SPPT	Multi Physics	RP, SKEB	RP, SKEB	No	SPPT	STTP	SKEB & PPT
Surface Perturbations	No	No	SST	No	No	No	No	No

Notes: 1) Only systems represented at the meeting are included.
2) From 2012 MOGREPS-G will change from 24 members twice per day to 12 members 4 times per day; products will be generated from 24 members by time-lagging of the last 2 cycles.

4.2 LAM-EPS

4.2.1 The meeting noted that operational Regional Ensemble Prediction Systems (REPS) is playing a rapidly increasing and important role in a number of NMHSs and Consortia (e.g. COSMO, Aladin, UM). Currently, the horizontal resolution ranges typically from 7km to 33km, with up to 70 levels. The ensemble size ranges from 15 to 24 members and the length of integration ranges from 2 to 5.5 days. At most of the centres, the frequency of the outputs is 3 hours. In addition one centre is now running a convection-permitting EPS with a horizontal grid-length of 2.8km and hourly outputs for a 1 day forecast – this is included here for completeness, but is discussed in more detail in section 4.3 below.

4.2.2 The meeting noted that a variety of methods is used in generating initial perturbations in REPS, including global localized Ensemble Transform Kalman Filter (ETKF), Breeding Growth Method (BGM) and extended Breeding Vector (BV) method with Ensemble Transform and Rescaling (ETR). It should be noted that a number of these use downscaling techniques where the regional perturbations are downscaled from the global EPS providing the boundary conditions; in some cases the initial analysis is also downscaled. Models errors are addressed by a range of stochastic physics perturbations, including Stochastic Kinetic Energy Backscatter (SKEB) and Random Parameters (RP). Some REPS use a Multi-Model or Multi-Physics approach and Perturbations of Physics Tendencies (PPT). Most of the centres are implementing perturbations of surface fields (e.g. SST).

4.2.3 A summary of the main characteristics of the Regional Ensemble Prediction Systems implemented at the participating centres at this meeting (as of November 2011) is provided in Table 4.2.

Table 4.2 – Summary of main characteristics of the Regional Ensemble Prediction Systems implemented at the participating centres at this meeting (as of November 2011)

	UK Met Office	COSMO	COSMO	CMA, China	NCEP, USA	CMC, Canada
Name	MOGREPS-R	LEPS	DE-EPS	REPS	SREF	REPS
Forecast Model	UM	COSMO (param. Conv)	COSMO (expl. Conv)	WRF	NAMB/W RF	GEM
Assimilation Method	4D-Var	No	Nudging	3D-Var	GSI	No
Initial Analysis	Regional 4D-Var	Global	High resolution	Regional 3D-Var	Various	Global EnKF
Domain	Europe & N Atlantic	Europe	Germany	China	North American	North America
Horizontal Resolution	18km	7km	2.8km	15km	30km	33km

Vertical Resolution (levels)	70	40	60	31	40-60	40 (plus lid nesting – top boundary)
Initial Times	06, 18 (From 2012, 03, 09, 15, 21Z)[2)]	12	0, 3, 6, 9, 12, 15, 18, 21	00,12	03, 09, 15, 18	00, 12
Lead Time	54h	132h	24h	60h	87h	3 days
Output Frequency	3h	6h	1h	3h	3h	6 hour
LBC Source	MOGREPS-G	ECMWF EPS	SREPS	CMA GEPS	various	Canadian GEPS
No. of Members	23+1 (From 2012 11+1)[2)]	16	20	14+1	21	20+1
Multi-Model	No	No	No	No	Yes	No
Initial Perturbations	Global Localized ETKF	No (ECMWF EPS)	From 4 Global models	BGM	BV and ETR	Glob EnKF
Model Perturbations	RP, SKEB	No	No	Multi Physi.	Yes	SKEB & PPT
Surface Perturbations	SST	No	No	Yes	Soil moisture	No

Notes: 1) Only systems represented at the meeting are included.
2) From 2012 MOGREPS-R will change from 24 members twice per day to 12 members 4 times per day; products will be generated from 24 members by time-lagging of the last 2 cycles.

4.3 Convective-scale EPS

4.3.1 The meeting noted that at many advanced centres and consortia, NWP development to support high-impact weather prediction is now focused on high-resolution global models and convection-permitting (or convective-scale) models (grid spacing: 1-3 km). EPS is highly relevant to convective-scale NWP because convective instability adds a new scale of forecast uncertainty not resolved by the lower resolution models, and with much shorter timescales. In addition to convection itself, models on this resolution have greatly enhanced capability for forecasting other aspects of local weather, such as low cloud and visibility of interest to aviation. Many of these phenomena are significantly affected by topographic forcing which may give enhanced predictability when that forcing can be resolved by the models (e.g. convective initiation or valley fog). Convective scale EPS has the potential to provide information on the predictability of all these weather elements.

4.3.2 Physical processes leading to convection are highly non-linear, so that the explicit modelling of convective cells over one or a few hours should already be seen as medium-range forecasting, with very limited deterministic predictability for the individual cells. Thus, convection-permitting ensembles already focus on the shortest-range (0-24 hours), as opposed to the short-range (0-3 days). Furthermore, error growth in convection-permitting models does not necessarily behave in a similar way as in convection-parameterization models. Different error growth may arise from the strong non-linearities and the different role of physical processes.

4.3.3 A few centres (e.g. Météo-France, UKMO, and COSMO) have been experimenting with convection-allowing ensembles with horizontal resolutions of around 2.5km. None of these are operational at the current time, except the COSMO-DE-EPS. The emphasis has been put on the prediction of heavy precipitation events.

4.3.4 The meeting noted that uncertainty on the convective scale has a very large dimensionality, and a very large ensemble would be required to fully sample this. Therefore, these systems have been addressing uncertainty in the location of convective precipitation by neighbourhood techniques whereby the probability of heavy precipitation at a location can be estimated by considering whether the (deterministic) model has precipitation within a neighbourhood around that location. Since the size of the ensemble is limited by computing resource, it is intended to use the neighbourhood technique in combination with the ensemble.

4.4 Specialist EPS (e.g. Typhoon EPS)

4.4.1 The meeting noted that both global and regional EPS have been operated in the prediction of high-impact weather events. The meeting noted that JMA has recently improved its Typhoon EPS, by implementing a stochastic physics scheme and by setting the initial perturbation (IP) target area around the central position of TC forecasts as a circular region (in contrast to the previous rectangular-area settings) and reducing the IP amplitude. These revisions have improved the spread-skill relationship of TC track forecasting. JMA operates the Typhoon EPS four times a day (00, 06, 12, and 18 UTC) with a forecast range of 132 hours. There is a plan to merge this with the JMA global EPS.

5. PROGRESS OF EPS APPLICATIONS AND OPERATIONAL INTEGRATION

5.1 Integration of EPS in core operational forecasting

5.1.1 While noting that the majority of operational forecast production in many NMHSs has continued to be based on deterministic models, with the ensembles providing peripheral and supplementary information, the meeting learned that work is now underway by a number of centres to further integrate EPS in core operational forecasting.

5.1.2 The meeting recalled the establishment of the North American Ensemble Forecast System (NAEFS) in 2004 by the Meteorological Service of Canada (MSC), the National Meteorological Service of Mexico (NMSM), and the US National Weather Service (NWS). Downscaling products for parts of North America have been implemented. The meeting noted that within the NAEFS, ensemble producing centres (currently MSC and NWS): (1) exchange in real-time their raw forecast data and bias corrected forecast; (2) statistically post-process (include downscaling) all ensemble members; and (3) jointly with other Met Services (currently NMSM) develop and produce end products based on the combined ensemble of forecasts. The participating centres collaborate in the development of post-processing algorithms and software and share a common procedure to generate the basic products of bias-corrected forecasts, the corresponding weights and climatological percentile values. The products for probabilistic forecast (10%, 90%, 50%, mean, mode and spread) have been generated after statistical bias correction for all ensemble members. These products are freely available in digital form through the NOMADS Web site (http://nomads.ncdc.noaa.gov/), including easy selection of regional cut-outs. Downscaling probabilistic products have been generated in NDGD (National Digital Guidance Database) grid by using Real Time Mesoscale Analysis (RTMA) as proxy truth. Some of the end products are developed jointly (such as the North American week-2 temperature and precipitation anomaly forecast), while others are provided by individual participating centres. Since the January 2010 earthquake, meteograms (EPSgrams) and tailored EPS weather maps zoomed over the Caribbean area are produced in real-time at MSC-CMC and sent to Météo-France’s office in Martinique to support Haiti.

5.1.3 The meeting was informed that in order to integrate EPS more fully in the operational production, UKMO has been working to post-process deterministic and EPS forecasts in a more integrated fashion. The two post-processing techniques are being used:

(a) Site-specific forecasts
The UKMO uses a database system to store forecast data for specific locations (sites) both in the UK and around the globe. Site-specific extraction from model grids is carried out by a routine named SSPS (site-specific processing system) which selects the most appropriate model grid-point and corrects for differences between model orography and station height. Exactly the same SSPS routines are used to extract data from EPS grids (both MOGREPS and ECMWF EPS) and the deterministic models. Where site observations are available, a

Kalman filter (KF) is applied to correct site-specific model biases in forecasts, which further reduces the impact of the difference in resolution between EPS and deterministic models. Creation of the deterministic “best-data” value is now based on a Best Data Blending technique, which is a lagged blending updating technique, and incorporates a lot of EPS data. This is accompanied by probabilistic tables which store percentile values, Ensemble Mean and Spread. Probabilities for any threshold required by a user can be interpolated from these percentile points. If necessary, the ensemble spread is adjusted to ensure that the distribution is consistent with the deterministic value. This combination of deterministic and probabilistic best-data allows all automated products to use consistent data, including uncertainty as either error bars, a full distribution or probabilities of specific events.

(b) Gridded data
Gridded NWP data for the UK are stored on a 2km *UKPP* (UK post-processing) grid (outside the UK a 5km EuroPP grid is used for Europe and a 10km GlobalPP grid is under development). Forecasts from both short-range high resolution models and longer-range lower-resolution global models are downscaled to the same 2km grid. Ensemble data are now being experimentally downscaled to the same 2km grid, using high-resolution topographical information to make simple temperature and wind-speed adjustments.

5.1.4 The meeting discussed the portability of such systems and techniques, focusing especially on developing countries. It noted that high-resolution gridded fields will not be appropriate for all countries due to the high computing demands. It recommended sharing knowledge and experience to evaluate the applicability of such systems and techniques to other countries/regions.

5.2 New developments in post-processing and calibration

5.2.1 The meeting stressed the importance and need of post-processing to improve EPS outputs. To provide better predictions, some centres conduct statistical post-processing to their EPS products for bias correction and calibration of the pdf (probability density function). Bias correction is the most widely applied procedure, and an adaptive Kalman Filter approach is commonly used. The use of reforecast data can allow more complete training of statistical post-processing and make it more effective, especially for extreme events. At some centres, combination of an EPS with the corresponding high resolution deterministic prediction is operationally implemented, and statistical downscaling technique, with high resolution analysis as the reference, is used to provide forecast guidance at local scale. Some more sophisticated techniques for calibrating the first and second moments of the pdf are also being implemented. The meeting agreed to provide further explanation and advice on these post-processing techniques in the *Guidelines on EPS and Forecasting* (agenda item 8.1).

5.3 Communication of forecast uncertainty to users (PWS)

5.3.1 The meeting noted that a Public Weather Services (PWS) representative has been invited to attend this meeting to discuss collaboration with the PWS programme on communicating forecast uncertainty to users, with emphasis on disaster risk reduction and to ensure effective pull-through of the benefits to societal impact. However, due to unforeseen circumstances, no PWS representative participated in this meeting, so the paragraphs below reflect the views of the members of the ET-EPS.

5.3.2 The meeting noted that a major block to the publication and use of probabilistic forecasts has been the belief among many scientists and managers that people do not understand probabilistic information. There is now a considerable body of academic research which shows that a large majority of people do understand this information, and make better decisions when presented with it.

5.3.3 A common justification for use of deterministic forecasts is that “people just need to make decisions”. A good example of the above belief is the 2010 report PWS-21 “Guidelines on early warning systems and application of nowcasting and warning operations” (WMO/TD No 1559) which discusses risk management and defines risk as the *product of Hazard Probability and Vulnerability*, but goes on to state: “Probability is a concept and skill that most people have problems understanding, as many cannot handle statistical concepts or effectively factor probabilities into their decision-making.” Recent evidence does not support this view.

5.3.4 Two separate experiments conducted at Exeter University in the UK, using samples of several hundred undergraduate students studying a wide range of academic disciplines, have shown that candidates presented with temperature forecasts including uncertainty information in different formats make significantly better decisions than those presented with simple deterministic forecasts. People presented with the uncertainty information in graphical form were able to assimilate it and make their decisions more quickly than those presented with the same information in tabular form.

5.3.5 The meeting noted that a mass-participation experiment conducted by the UK Met Office in collaboration with the Universities of Bristol and Cambridge through an on-line game has also shown that people presented with probabilistic information in a variety of formats make better decisions than those given simple deterministic forecasts. The game format, in which participants were asked to make decisions to help an ice-cream seller maximise his sales on the basis of temperature and rainfall forecasts, was used to attract over 8000 separate people to participate. The majority of these were in the UK and spanned a wide range of ages and academic attainment. Participants were randomly assigned different types of forecast presentation, and each candidate asked the same questions based on the forecasts. Early results show not only that people make better decisions with probabilistic forecasts, but that they make the best decisions with the most complex presentations. For rainfall forecasts they made better decisions, and showed greater understanding of their confidence in their decisions, when the probability of precipitation was presented numerically (e.g. 20%) than when a simpler Low/Medium/High scale was used. Similarly they performed better with graphical presentations which gave more detail; the best performance came when a graphical format was accompanied by a numerical probability. For temperature forecasts they performed best when presented with a fan chart showing 5 percentile points, the most complex presentation used. Importantly, in simple situations where the decision was easy even with a deterministic forecast, participants did not make worse decisions when presented with probabilistic presentations, so they were not confused by the information.

5.3.6 The meeting also noted that experiments conducted by psychologist Susan Joslyn in the USA have also shown that people make better decisions when presented with uncertainty information in weather forecasts. In general, this study indicates that forecasters make better decisions with uncertainty products which mean more advisories have been posted when event is likely, or less advisories when event is unlikely. For general public or non-experts, people understand that forecasts involve uncertainty. They are able to estimate the uncertainty on their own when no uncertainty estimate is provided. The good uncertainty forecast could improve people’s understanding of future weather events when acknowledging and specifying the uncertainty. For decision advises, there is advantage by using uncertainty information by comparing deterministic information. This study concludes that people realize there is uncertainty in weather forecasts, and are reluctant to act when they have no information on the confidence of the forecast. People make better decisions when forecasts include uncertainty estimates. Uncertainty estimates increase trust in the forecast as they acknowledge the uncertainty.

5.3.7 The meeting agreed that the evidence shows that while not everyone may be able to make better decisions with uncertainty information, a very large number can. It would not make sense to withhold the useful extra information from those who can benefit from it because of a minority who are less able to exploit it.

5.3.8 There is evidence that many people want a simple headline weather forecast which they can assimilate quickly; one option is to provide the uncertainty information to those who choose to look beyond the headline for more detail. Nevertheless, for example, for precipitation the headline may be best presented as a Probability of Precipitation figure, with or without an accompanying graphical symbol.

5.3.9 The latest research also provides useful guidance on the relative effectiveness of different presentations of uncertainty information in helping people make better decisions. It is recommended that the PWS programme review and update the guidelines on the presentation of probabilistic forecasts in the light of this new research.

5.3.10 Given the above results, the meeting strongly recommends that PWS should consider wider provision of uncertainty information to both civil responder services and the general public. The use of probabilistic information in severe weather warnings is discussed in item 5.4. In addition to aiding better decision-making, providing this uncertainty information has the following advantages:

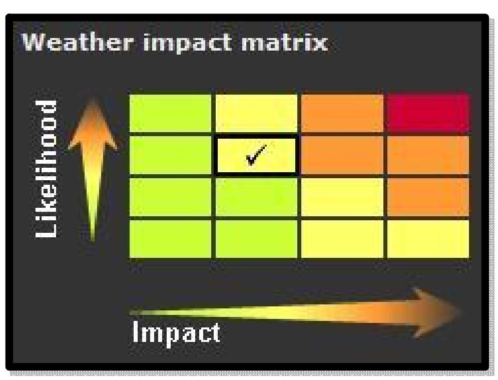
- Scientific integrity through more honest representation of the capability and limitations of prediction systems;
- Allows public users of forecasts to decide their responses according to their personal sensitivities and vulnerabilities;
- Defensible position for the NMHS where accusations of wrong forecasts are made, since all available information is shared publicly.

5.3.11 It should be noted that the studies of public understanding have been conducted in developed countries. Even the widest study using an on-line game, which included candidates with a wide range of academic achievement, was nevertheless a selective sample of those who chose to play the game and complete it to the end. It can therefore be argued that the results are not applicable to all countries or to the entire population. Despite this, there is now very strong evidence that very many people can benefit from having well-presented uncertainty information, and it does not make sense to withhold that information because some others may not be able to fully exploit it.

5.4 Customer-focused applications

5.4.1 The meeting considered the use of probabilistic information in severe weather warnings as one of the customer-focused applications. The application of EPS in atmospheric transport and dispersion is addressed in item 5.5. Other applications, including probabilistic forecasts of meteorological and other weather-related parameters into impact models, are discussed under agenda item 7.1. The following paragraphs describe risk-based national severe weather warning services implemented in some of the participating NMHSs in this meeting.

5.4.2 The UK Met Office has recently introduced a new risk-based National Severe Weather Warning Service (NSWWS). This is based on the concept of a risk matrix taking account of both the probability and impact of severe weather.



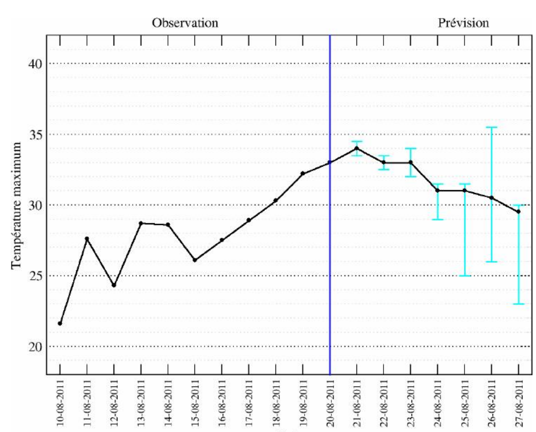
5.4.3 The headline presentation of a warning is a “traffic light” colour from the range Green, Yellow, Amber or Red (as widely used in the European MeteoAlarm system). As illustrated in the matrix above, a Yellow warning may result from either a higher probability (likelihood) of a moderate impact, or a low probability of a high impact event. Amber warnings result from similar but higher combinations, while a Red is limited to a high probability of a high impact event.

5.4.4 A similar concept is used in Meteo-France's *Vigilance* system. For D/D+1, this system is based on 4 colours (each of them representing a level of risk for 9 different parameters): green (no special vigilance required); yellow and orange (potentially dangerous and unusual meteorological phenomena have been forecast); and red (an absolute vigilance is necessary; dangerous and exceptionally intense meteorological phenomena are forecast; keep regularly informed about meteorological developments and conform to the advice or orders given by authorities). When an orange or red level is assessed, direct exchanges between Météo-France and safety services are systematically set. At D+2/D+3 an early warning is now provided to safety services. This warning is based on a well-calibrated probability that an orange or red level may be issued at D/D+1. It is assessed by the forecaster over the 22 administrative regions of France in terms a 4 levels risks scale (high, medium, low and no-risk). The calibration is done through the past statistical correspondence between the risk level and the probability threshold.

5.4.5 The meeting noted that assessment of the level of risk and impact of an expected event is to a large degree subjective, but initial guidelines on impact levels for different types of weather event (strong winds, heavy rain, etc) have been defined based on forecaster experience in analysing the meteorological situation and discussions with civil responder groups who are the primary users of the severe weather warnings. Climatology provides a useful reference for impact thresholds, but the meeting noted that relevant thresholds for impact-based warnings may vary according to the season and also due to antecedent conditions – for example, thresholds for strong wind impact may be lower in summer when there are leaves on the trees, or if the soil is saturated so trees are more likely to fall. Given an appropriate set of impact-related thresholds, EPS can be used to estimate the probabilities of exceeding these thresholds to complete the risk matrix shown in 5.4.2. The UK Met Office showed promising early results with a first-guess warning system exploiting the risk matrix with a regional EPS.

5.4.6 The meeting agreed that warnings should normally be related to significant societal impact. For example, for Tropical cyclones (TC), which are major high-impact storm systems, it is indispensable to predict their approach and landfall at an early stage, even if their probabilities are very low, to aid prevention of disasters. In the case of JMA, and its RSMC Tokyo - Typhoon Center, the TC track forecast up to five days ahead is represented with probabilistic information in the form of geographical circles with a 70% probability of enclosing the TC centre. A worst-case scenario based on the highest impact is also presented. This will encourage early attention by the public and preparation for Typhoons. With the extension of the forecast range from three to five days mainly using outputs from the JMA Typhoon EPS, JMA will consider alternative methods such as an elliptical form, based on verification results. A similar process is used in RSMC La Reunion.

5.4.7 Probabilistic temperature forecasts are used by MeteoSwiss in order to warn public health authorities about heat waves. Since 2011, these are provided graphically in the following form:



The evolution starts with the recent observations and continues with the temperature forecast provided as a first guess by the ensemble median and adjusted by the forecasters. The spread is given by the ensemble. The recipients of these forecasts much appreciate this type of presentation which helps them take preventative actions against dehydration.

5.5 EPS in Atmospheric Transport and Dispersion

5.5.1 The meeting was briefed on applications of EPS in atmospheric transport modelling (ATM) and the modelling of oil spill drift. Regarding ATM, the meeting noted that recent events (European volcano crisis in 2010, and the nuclear accident in Japan, in 2011) triggered significant attention for this technology. The potential of EPS to improve ATM results is increasingly explored and investigated. The meeting noted that some centres have started or are planning to start experimenting with ATM models coupled with weather forecasts from EPS systems.

5.5.2 The meeting recognized that working on ATM ensemble systems is a good approach and recommended that it should be further explored and developed. The meeting agreed that it should continue to deal with ATM EPS applications under the agenda item “Progress of EPS applications and operational integration”, in the same way as it deals with other applications in the EPS area, and that the subject is implicitly covered in item (h) of the revised Terms of Reference (ToR).

5.5.3 The meeting recognized that ensemble ATM applications require substantially more and different input data (model-level data) from EPS systems than the other applications. The centres operating EPS systems need to be aware of that additional user requirement, and are encouraged to work with the ATM community. Since EPS data sets are usually so large that they cannot be stored permanently, and since it may even be problematic to transfer all the data, the meeting noted the option to develop methods to reduce data needs, for example by utilizing cluster approaches.

6. GIFS-TIGGE DEVELOPMENTS AND PLANS

6.1 GIFS-TIGGE

6.1.1 The meeting was briefed on the key outcomes of the ninth meeting of the GIFS-TIGGE Working Group, held in Geneva, from 31 August to 2 September 2011. The meeting noted the developments for a new sub-seasonal to seasonal prediction project, and the links that will be established to maximize compatibility between this new project archive and the existing TIGGE archive.

6.1.2 The meeting noted that recent TIGGE developments focus on calibration of EPSs, combination of multi-models and research related to the development of probabilistic forecasts. The meeting noted that results on multi-model benefits vary and depend on the application and on the methods employed, for example whether the models are bias corrected prior to combination. The meeting agreed that the multi-model combinations for operational use should be made after bias correction of the individual models, and therefore suggested further research studies on these aspects. The meeting also suggested further research studies focused on surface variables. The meeting stressed the need for research studies on the trade-off between numbers of EPS members versus resolution, and consequential benefits of multi-ensemble combinations.

6.1.3 The meeting noted that GEOWOW (GEOSS interoperability for Weather, Ocean and Water) is an EU-funded FP7 3-year project beginning in September 2011. GEOWOW will make a significant European contribution to the Global Earth Observation System of Systems (GEOSS) by improving the overall quality of the current GEOSS Common Infrastructure (GCI), addressing access to data, usability and interoperability. The project will enhance the accessibility of the TIGGE archive at ECMWF for the wider user community, in particular the ability to efficiently access long time series of forecast data at user-specified locations, and will support the development of TIGGE-LAM facilities at ECMWF. GEOWOW will promote the wider use of TIGGE data for research across a range of GEO Societal Benefit Areas and show how the TIGGE archive can be used to develop ensemble products for different applications. GEOWOW will demonstrate the potential use of such ensemble products, with a focus on severe weather, in close liaison with the SWFDP.

6.1.4 MRI/JMA has developed a Web site (http://tparc.mri-jma.go.jp/TIGGE/tigge_SWFDP.html) which displays risks of high-impact weather (e.g. heavy rainfall, extremely high/low temperature, and strong wind) using the TIGGE data from four global NWP centres (ECMWF, JMA, NCEP and UKMO) and a brief description about these products and guidelines on how to use them. The genesis potential of high-impact weather is calculated by comparing ensemble members with climatological PDFs calculated from each of the four TIGGE models. The Web site is automatically updated every day and includes forecast up to 15 days ahead. The meeting reviewed these guidelines and agreed that these provide a good explanation on the interpretation of GIFS-TIGGE products. The meeting noted that prototype products based on TIGGE data suffer from a 2-day delay which is a barrier to operational use, but are useful examples of what could be done. As these products go out to 15 days, they could also provide good information on the trends. New products would need dialogues between users and product developers. The meeting was informed that SWFDP RSMCs have been invited to evaluate prototype products available on the Web site and assess requirements for near-real-time products.

6.1.5 The meeting noted that the GIFS-TIGGE Working Group recommended that training on new GIFS products should be integrated in any training conducted by the ET-EPS. While noting that in the past the ET-EPS used to conduct training and capacity building activities, the meeting realized that EPS guidance has now been covered through the SWFDP, and therefore recommended that training on the interpretation of the new products developed under GIFS should be integrated in SWFDP training events.

6.2 TIGGE-LAM

6.2.1 The meeting noted that TIGGE-LAM is intended to provide an equivalent to the TIGGE archive for LAM-EPS. Scope for this is much less globally because of limited overlap of ensemble domains, but the TIGGE archive centres agreed several years ago to archive a limited set of High Priority Parameters for ensembles in their own regions of the globe. The area with greatest overlap was Europe where the TIGGE archive centre is ECMWF.

6.2.2 The meeting noted that progress on implementing TIGGE-LAM has been very slow for two reasons: (a) lack of devoted resources; and (b) technical problems due to GRIB2 coding. The meeting learned that CMA is now archiving high priority parameters from the CMA Regional EPS System and ECMWF will archive European products as part of GEOWOW. Efforts were being made to get similar archiving going in other parts of the world.

6.2.3 With the focus of LAM ensemble development moving rapidly to convection-permitting ensembles, the meeting recommended that TIGGE-LAM implementation should now accommodate these developments, and noted that this has already been considered within GEOWOW.

7. EPS IN SEVERE WEATHER FORECASTING, INCLUDING THE SWFDP

7.1 Probabilistic forecasts of important meteorological and other weather-related parameters into impact models (e.g. flash floods and coastal flood forecasting)

7.1.1 The meeting noted that the ECMWF Extreme Forecast Index (EFI) is one of the products that is widely used to provide alerts for potential severe weather events. Following feedback from users, the current EFI products are being extended to include more parameters (e.g. snowfall and ocean waves) and later forecast ranges (currently only provided to day 5; new products will include information up to day 15). The meeting learned that KMA has been working on the development of an EFI of daily minimum and maximum temperature, wind gust, and daily accumulated rainfall amounts to support one-week forecast.

7.1.2 The meeting also noted that products for tropical cyclone genesis have been implemented on the ECMWF Web site. A new TC tracker is under pre-operational testing, which extends the predicted tracks from 5 days up to 10 days ahead.

7.1.3 The meeting noted the increasing application of EPS to predict severe or high-impact weather events and in the propagation of the weather forecasts into impact models. The uncertainty in the weather forecast can be propagated through to uncertainty in impact by coupling ensemble members to impact models and generating a distribution of impact predictions. Examples of applications include:

- risk-based first-guess severe weather warnings, which provides guidance to forecasters related to the impact of severe weather (e.g. Met Office UK);
- risk models combining hazard probability from EPS with models of societal impact (e.g. strong wind impact on transport networks implemented at Met Office UK);
- hydrology (e.g. MAP D-Phase for small catchments; and EFAS for large catchments. NWS of US for all time and space scales. Experimental hydrological EPS is implemented at Météo-France);
- Storm surge (e.g. Met Office UK and Netherlands have implemented; Météo-France will do it next year);
- Waves (e.g. implemented at the ECMWF; and a multi-model jointly implemented by NCEP and US Navy);
- Marine Pollution (e.g. implemented in Météo-France);
- Atmospheric Transport (e.g. implemented experimentally at Met Office UK, and with a multi-analysis ensemble for backtracking in Austria);
- Energy (wind, solar, etc.) and Trading;
- Forest fire (e.g. implemented in Canada);
- Aviation (i.e. icing, CAT, low visibility; routing; runway management).

7.1.4 The meeting strongly endorsed this approach, particularly where the primary source of uncertainty is from the weather forecasts. For ATM, the meeting stressed that further work is required to understand the complete picture, particularly understanding the uncertainty in the source term.

7.2 Practical application of EPS in existing projects (e.g. SWFDP)

7.2.1 The meeting noted that many NMHSs have difficulties in making use of ensemble fields made available in real-time by the producing centres for downloading. In particular, many NMHSs, especially those of developing countries and LCDs, do not have the computing facilities or capability to generate products from ensemble fields. Therefore, EPS continues to be a critical component of the guidance data provided to SWFDP subprojects from the global centres. The subprojects in Southern Africa, the South-west Pacific and in Eastern Africa have been supplied with EPS data by ECMWF, the Met Office UK and NCEP. The subproject in Southeast Asia, planned to commence in April 2012, will be supplied with EPS data by CMA, JMA, KMA and ECMWF (products available for all WMO Members). The use of EPS data is also included in plans for the SWFDP subproject being initiated in the Bay of Bengal.

7.2.2 The meeting learned that products from multiple ensemble systems (PEARP/MF, EPS/ECMWF, MOGREPS/UKMO, EPS/CMC, GEPS/NCEP) and also combined products, produced from the TIGGE database are made available via the RSMC La Réunion Website. Access to these products is provided to members of the Tropical Cyclone Committee (RA I) and to participating NMHSs in the SWFDP – Southern Africa, under password access and a usage agreement.

7.2.3 The meeting agreed that the training workshops carried out as part of the SWFDP subprojects have proven one of the most effective ways of providing training on the use of EPS, as the training have been provided in conjunction with ongoing access to operational EPS data.

7.2.4 The meeting also agreed that the SWFDP framework represents a systematic approach for building capacity and for transferring knowledge and skills to NMHSs, especially to weather forecasters. Its approach has been used to implement a series of ready-for-production enhancements to the forecasting process and to provide benefits to other scientific and technological developments that are intended for operational implementation (e.g. transfer of relevant promising research outputs from GIFS-TIGGE into operations through trials at the SWFDP).

8. DEVELOPING GUIDANCE FOR FORECASTERS

8.1 How to complete and update the Guidance started at the last meeting

8.1.1 The meeting reviewed and updated the *Guidelines on EPS and Forecasting*, based on the discussions under the previous agenda items. The updated *Guidelines on EPS and Forecasting* are provided in Annex IV. Noting that further developments are still required in order to complete it, the meeting noted the availability of other sources of guidance which may be of use to WMO members, in particular the ECMWF User Guide (http://www.ecmwf.int/products/forecasts/guide/) which provides comprehensive guidance on the use of ECMWF systems including detailed advice on the use of EPS, and the COMET training materials (http://deved.meted.ucar.edu/nwp/pcu1/ensemble/)

8.1.2 The meeting requested the chair of the meeting to work with the WMO Secretariat to investigate the possibilities for arranging for editing of the *Guidelines on EPS and Forecasting* and then to publish them as a WMO publication and to distribute it to all WMO Members.

8.2 Role of EPS in deterministic forecasting

8.2.1 The meeting discussed the role of EPS in deterministic forecasting under agenda item 5.1. It also noted that this issue is partially addressed in section III of the *Guidelines on EPS and Forecasting* (see Annex IV).

8.3 Initial guidance on convective scale EPS

8.3.1 Based on the discussions under agenda item 4.3, the meeting developed initial guidance on convective-scale EPS and included it as an additional section in the *Guidelines on EPS and Forecasting* (see section VIII in Annex IV). The meeting noted that due to the very high computing costs of convective scale EPS, and the early stage of development of such systems, forecasts from convective-scale EPS are unlikely to be available to WMO Members outside the producing nations for a number of years.

9. VERIFICATION

9.1 The meeting recalled that JMA, as the Lead Centre for EPS Verification, has been operating two Internet sites since 2004: (a) an FTP site for the EPS producing centres to upload the statistical data for their EPS verification reports; and (b) a Web site (http://epsv.kishou.go.jp/EPSv/) that shows the original verification statistics, their update and contents information, and their visualized figures, indicating that it is possible to diagnose the skill of EPS of each EPS producing centre through comparison with those of other EPS. The original verification statistics are replicated at the Web site and available to confirm after the registered centre puts them on the FTP site.

9.2 The meeting noted that as of October 2011, eight EPS producing centres are registered in the FTP site, and seven are uploading their verification statistics to the FTP site. However, exchanged parameters are still quite limited. The exchange of CRPS, which is defined as a new score in the latest amendment of the *Manual on GDPFS*, has started up quickly.

9.3 The meeting recalled that the verification grid spacing used in calculation of the statistics has been 2.5°x2.5° (lat/lon). Noting that for deterministic NWP verification, Cg-XVI approved for inclusion in the *Manual on the GDPFS* (WMO-No. 485) that all parameters shall be verified against the centre's own analysis on a regular 1.5° x 1.5° grid, the meeting realized that the same principle should be applied to the EPS verification.

9.4 To ensure consistency between results from different centres, the meeting agreed that a common climatology shall be used for those scores requiring a climatology. All centres shall use the climatology provided via the lead centre, which is the same as the one used for deterministic NWP verification.

9.5 The meeting noted that following the request by the 2006 extraordinary session of the Commission for Basic Systems (CBS-Ext.(06)), the Web site ceased to be password protected in December 2009, in order to make the EPS verification results available to all WMO Members. At the same time, a new page (http://epsv.kishou.go.jp/EPSv/Exchange/reportstatus.html) was established in order to show the reported verification parameters and their available periods for each centre. In July 2011, following the approval of the amendments to the *Manual on the GDPFS* by Cg-XVI, the Web site and its operating system was updated to include: (a) a scheme to decode and publish CRPS tables; (b) links to lead to published CRPS-related data; and (c) a Note to show the details of each center's report. Recently, the CRPS tables, which include statistics of both ensemble and control, and its published data become available on the Web site. The CRPS-base skill score are also visualized as a part of the published data in order to clearly show not only ensemble skill but also impact of ensemble use on single use.

10. REVIEW OF EPS ASPECTS IN THE MANUAL ON THE GDPFS

10.1 The meeting noted that the World Meteorological Congress (Cg-XVI, May 2011) agreed that there are fundamental changes under way in the Basic Systems and that the review of the *Manual on the GDPFS* (WMO-No. 485) should be done with the existing system of world, regional and national centres of the GDPFS, and the future evolution of the GDPFS in mind, such as the inclusion of all WMO operated meteorological centres that provide operational data-processing and forecasting services (including those coordinated by CBS and joint CBS-other Technical Commissions and/or WMO Programmes, as well as joint WMO-other international organizations centres). The review should take into account the developments in WIGOS and WIS, lessons learnt from SWFDP, and anticipated results of, and operational implications from the WWRP/TIGGE project "Global Interactive Forecast System". In this context, the meeting was informed of the outcomes of the Expert Meeting on the Revision of the Manual of the Global Data-Processing and Forecasting System (GDPFS) that was held in Geneva, Switzerland, from 19 to 21 October 2011.

10.2 Noting that a first version of the new *Manual* will be presented to CBS-XV, in 2012, for consideration, the meeting agreed that any proposed amendments to the *Manual on the GDPFS* should be reflected in the new version. Draft text for the new *Manual* on global ensemble prediction, coordination of EPS verification and limited area ensemble prediction, including proposed amendments, is given in Annex V. The meeting requested the chairperson, in coordination with the Expert Team members and in liaison with the chairperson of the CBS Coordination Group on Forecast Verification, to review the draft text for the new Manual, and submit to the Secretariat a revised version by mid-March 2012.

11. REVIEW THE TERMS OF REFERENCE OF THE ET-EPS

11.1 The meeting reviewed the team's Terms of Reference (ToR) as adopted at CBS-Ext.(10), and proposed revised ToR as given in Annex VI.

12. ANY OTHER BUSINESS (AOB)

12.1 There were no other issues raised during the meeting.

13. CLOSING OF THE MEETING

13.1 The meeting of the CBS Expert Team on Ensemble Prediction Systems (ET-EPS) closed at 16:30 on Friday, 18 November 2011.

ANNEX I

AGENDA

1. **OPENING**

2. **ORGANIZATION OF THE MEETING**
   - 2.1 Adoption of the agenda
   - 2.2 Working arrangements

3. **REPORT ON THE OUTCOMES OF CBS-Ext.(10) AND CG-XVI RELATED AND/OR RELEVANT TO THE ET-EPS**

4. **PROGRESS OF EPS IMPLEMENTATIONS**
   - 4.1 Global EPS
   - 4.2 LAM-EPS
   - 4.3 Convective-scale EPS
   - 4.4 Specialist EPS (e.g. Typhoon EPS)

5. **PROGRESS OF EPS APPLICATIONS AND OPERATIONAL INTEGRATION**
   - 5.1 Integration of EPS in core operational forecasting
   - 5.2 New developments in post-processing and calibration
   - 5.3 Communication of forecast uncertainty to users (PWS)
   - 5.4 Customer-focused applications
   - 5.5 EPS in Atmospheric Transport and Dispersion

6. **GIFS-TIGGE DEVELOPMENTS AND PLANS**
   - 6.1 GIFS-TIGGE
   - 6.2 TIGGE-LAM

7. **EPS IN SEVERE WEATHER FORECASTING, INCLUDING THE SWFDP**
   - 7.1 Probabilistic forecasts of important meteorological and other weather-related parameters into impact models (e.g. flash floods and coastal flood forecasting)
   - 7.2 Practical application of EPS in existing projects (e.g. SWFDP)

8. **DEVELOPING GUIDANCE FOR FORECASTERS**
   - 8.1 How to complete and update the Guidance started at the last meeting
   - 8.2 Role of EPS in deterministic forecasting
   - 8.3 Initial guidance on convective scale EPS

9. **VERIFICATION**

10. **REVIEW OF EPS ASPECTS IN THE MANUAL ON THE GDPFS**

11. **REVIEW THE TERMS OF REFERENCE OF THE ET-EPS**

12. **ANY OTHER BUSINESS (AOB)**

13. **CLOSING**

ANNEX II

LIST OF PARTICIPANTS

Ken **MYLNE** (Chair) Ensemble Forecasting Applications Manager Met Office FitzRoy Road Exeter Devon EX1 3PB United Kingdom	Tel: Fax: E-mail:	+44 (0)1392 886070 +44 (0)1392 885681 ken.mylne@metoffice.gov.uk
André **METHOT** 2121 Trans Canada Highway Dorval, QUEBEC H9P 1J3 Canada	Tel: Fax: E-mail:	+1 514 421 4654 +1 514 421 4657 andre.methot@ec.gc.ca
Jing **CHEN** Operational Division CMA Numerical Prediction Center 46 Zhongguancun Nandajie Street BEIJING 100081 China	Tel: Fax: E-mail:	+86 10 68407423 +86 10 58995418 chenj@cma.gov.cn
Jean **NICOLAU** Météo-France 42 avenue Coriolis 31057 TOULOUSE France	Tel: Fax: E-mail:	+33 5 6107 8210 +33 5 6107 8609 jean.nicolau@meteo.fr
Gerhard **WOTAWA** NMC Vienna Central Institute for Meteorology and Geodynamics Hohe Warte 38 A-1190 VIENNA Austria	Tel: Fax: E-mail:	+43 1 36026 2007 +43 1 3691233 Gerhard.wotawa@zamg.ac.at
Masayuki **KYOUDA** Numerical Prediction Division Japan Meteorological Agency 1-3-4 Ote-machi, Chiyoda-ku TOKYO 100-8122 Japan	Tel: Fax: E-mail:	+81-3 3211 8407 +81-3 3211 8407 kyouda@naps.kishou.go.jp
Young-Youn **PARK** Numerical Model Development Division Korea Meteorological Administration 45 Gisangcheong-gil, Dongjak-gu SEOUL 156-720 Korea (Rep. of)	Tel: Fax: E-mail:	+82-2-2181-0513 +82-2-2181-0907 yypark@kma.go.kr
Pierre **ECKERT** Météo Suisse 7 bis avenue de la Paix 1211 GENEVE 2 Switzerland	Tel: Fax: E-mail:	+4122 716 2839 +4122 716 2829 pierre.eckert@meteosuisse.ch

Yuejian **ZHU** Ensemble Team Leader Environmental Modeling Center NOAA – National Centers for Environmental Prediction 5200 Auth Road, Suitland, MARYLAND MD 20746-7566 USA	Tel: Fax: E-mail:	+1 301 763 8000 ext 7052 +1 301 763 8545 Yuejian.zhu@noaa.gov
David **RICHARDSON** Head Meteorological Operations Section ECMWF Shinfield Park, READING RG2 9AX United Kingdom	Tel: Fax: E-mail:	+44 1189 49 94 20 +44 1189 86 94 50 david.richardson@ecmwf.int
WMO REPRESENTATIVE		
Alice **SOARES**	Tel: Fax: E-mail:	+4122 8308449 +4122 7308128 asoares@wmo.int
Peter **CHEN**	Tel: Fax: E-mail:	+4122 7308231 +4122 7308128 pchen@wmo.int

ANNEX III

SUMMARY OF THE MAJOR PROGRESS ACHIEVED IN IMPLEMENTING EPS AND FUTURE PLANS

European Centre for Medium-Range Weather Forecasts (ECMWF)

ECMWF runs an operational ensemble prediction system (EPS) twice daily (from 00 and 12 UTC). The EPS uses the same forecast model as used for the ECMWF operational deterministic forecast, but at lower resolution: T639 (32 km) up to day 10, then T319 (64 km) from day 10 to day 15; 62 levels throughout. The atmosphere is coupled to an ocean model from day 10 onwards. The ensemble comprises one control forecast and 50 perturbed members. Initial perturbations are based on a combination of singular vectors (SVs) and an ensemble of data assimilations (EDA); the EDA perturbations replaced the evolved singular vectors (June 2010). Model uncertainty is represented using stochastically perturbed physical tendencies and (since November 2010) a spectral stochastic backscatter scheme. The changes to initial and model perturbations improved the ensemble spread throughout the forecast range. The EDA introduced initial perturbations for the whole of the tropics for the first time.

The current EFI products are being extended to include more parameters (e.g. snowfall and ocean waves) and later forecast ranges. A new tropical cyclone tracker is under pre-operational testing, which extends the predicted tracks from 5 days up to 10 days ahead. Products for tropical cyclone genesis have also been implemented on the ECMWF website. New EPS cluster products were implemented operationally in November 2010 (described in detail in ECMWF Newsletter 127). As well as the daily clustering of the EPS members, the new system includes a second component, which associates each cluster with one of four fixed climatological regimes, providing information about the likelihood of transitions between regimes.

ECMWF has introduced a set of six headline scores to measure the trend in performance over the next 10-year strategy period. Three of these are for the EPS (the other 3 are for the deterministic forecast). The primary headline score for the EPS is the continuous ranked probability skill score (CRPSS) for 850 hPa temperature over the extra-tropical northern hemisphere. Two supplementary scores monitor the performance for surface weather (precipitation) and severe events (EFI for extreme winds). ECMWF provides the agreed set of verification scores to the Lead Centre for EPS verification.

Météo-France

The Météo-France global Ensemble Prediction System (*PEARP*, which stands for “Prévision d’Ensemble ARPège”) is based on the global spectral model ARPEGE. The main features are:

- Initial perturbation: the coupling with the 6-members 3D-Var FGAT ensemble assimilation running parallel to the main 4D-Var Arpege assimilation cycle (Berre *et al.*, 2007). One goal
- of this assimilation ensemble with perturbed observations is to feed the 4D-Var with error statistics of the day. But another, equally important goal, is to provide perturbations for ensemble. These perturbations are combined to the different sets of singular vectors computed over 7 different areas in order to provide the perturbed initial conditions for the ensemble.
- Model perturbation: each ensemble member uses randomly a particular physical package among the 10 available.
- 35 members.
- Run twice a day at 06Z with a lead time of 72h in addition to the current one (18Z up to 108h).
- Resolution: The model used for the ensemble has a spectral truncation of TL538 with a stretching coefficient of 2.4. This corresponds to a resolution of about 15 km over France and 90km over New-Zealand.
- TIGGE: data are daily sent to TIGGE database on a 1.5°x1.5° global grid.

- Use: a lot of products are developed PEARP outputs (probabilities, percentiles, spaghettis, tracks, strike-probabilities, clustering, …). Particular attention is paid to the short range probabilistic forecast of precipitation and strong winds.

LAM-EPS: As PEARP have a LAM-EPS resolution over Europe (around 15km), Météo-France didn't develop a real LAM-EPS and the future convective scale EPS will be directly coupled to the global PEARP system.

Météo-France is developing a convective scale ensemble prediction system for operational NWP, using the AROME model at 2.5km resolution over western Europe. Several aspects of the ensemble setup are being examined :

- the selection of large scale lateral boundary conditions from a global ensemble
- the perturbation of initial conditions from an AROME ensemble data assimilation system
- the impact of synthetic observations designed to enhance dispersion for events of interest
- the added value of a fine-scale ensemble over other approaches (large-scale ensemble, lagged ensemble, multimodel)
- the potential of stochastic physics and surface conditions.

The future ensemble operational suite

Met Office UK (UKMO)

The Met Office runs two versions of its global ensemble, MOGREPS-G to 3 days and MOGREPS-15 to 15 days. Both run twice per day at N216L70 (~60km) grid resolution with 24 members using identical initial conditions. Initial condition perturbations are provided using an ETKF (Ensemble Transform Kalman Filter) with both horizontal and vertical localization to optimize the scaling of perturbations across the globe. Model errors are addressed by a range of stochastic physics perturbations including SKEB2 (Stochastic Kinetic Energy Backscatter) and Random Parameters. Perturbations to the SST (sea-surface temperature) field have recently been implemented giving a significant improvement in near-surface spread.

The Met regional ensemble (MOGREPS-R) runs over the North Atlantic and Europe at ~18km grid resolution with 24 members. This is a downscaling ensemble which takes its initial condition perturbations downscaled from the MOGREPS-G global ensemble initialized 6h earlier, and adds them to a higher resolution regional analysis. LBCs are also provided from MOGREPS-G.

Both MOGREPS-G and –R will soon be upgraded to run 4 times per day with 12 members per cycle; 24-member ensemble products will be generated by a lagged-average combination of the latest two cycles. This is to facilitate better integration of EPS with deterministic NWP in the operational forecast production, and will also support the planned new convective-scale EPS, MOGREPS-UK, which will also run 4 times per day with 12 members. MOGREPS-15 is not affected and will continue to run twice per day with 24 members.

Consortium for Small-scale Modelling (COSMO)

COSMO LAM EPS: COSMO LEPS

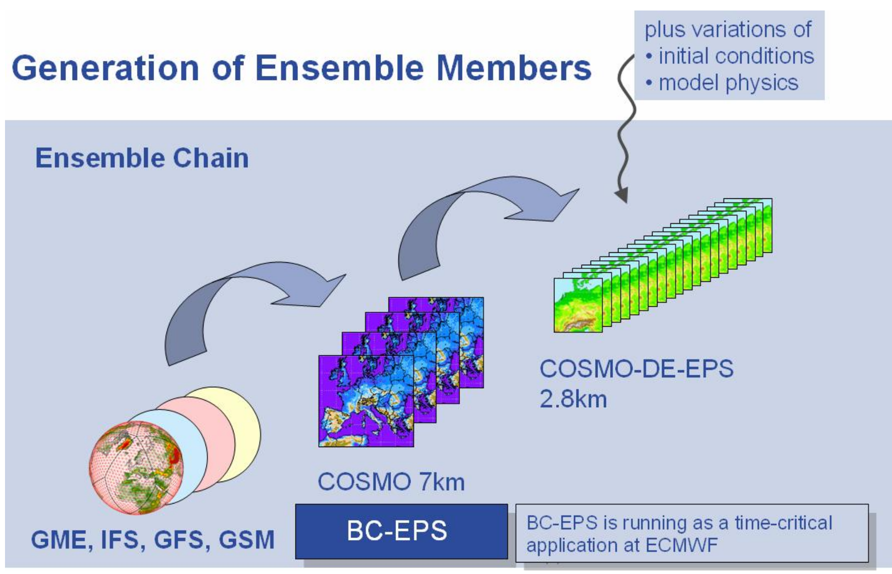
COSMO LEPS is a well established system which is used as a dynamical downscaling of the ECMWF EPS. The 102 members of the 00z and 12z EPS are clustered at day 4 and 5 so that 16 representative members are kept. They are used as initial and boundary conditions for 16 COSMO-7 runs (non-hydrostatic, parameterised convection). Perturbations in physics are introduced by employing either Tiedtke or Kain-Fristch convection scheme (randomly chosen) and perturbations in the turbulence scheme and in physical parameterisations.

Recent improvements have been reached by including the soil fields provided by the deterministic COSMO-DE. The long term verifications show that the scores become constantly better, especially for the higher thresholds. Finally the COSMO-LEPS will be relocated over the Black sea in view of the **2014 Sochi Olympic Games** as part of a Forecast Demonstration Project.

Convection-permitting ensemble COSMO-DE-EPS

The ensemble members are computed with the help of the non-hydrostatic model COSMO-DE. Deep convection is there described explicitly. The grid spacing is 2.8km and the domain covers Germany with some surroundings. In ensemble mode, the choice of grid spacing and domain remains unchanged. Concerning the ensemble perturbations, the focus is on three sources of uncertainty: the lateral boundary conditions, the model physics and the initial conditions.

To quantify the uncertainty due to boundary conditions, the COSMO-DE-EPS is embedded in a chain of ensembles starting with four global deterministic models for each of those a COSMO model with a resolution of 7 km is run. The models are used as boundary condition for 5 different COSMO models at 2.8 km resolution using different physics. 8 runs per day (every 3 hours) are integrated.



The uncertainty in the physics is represented by varying distinct parameters of the parameterisation schemes that are expected to have a significant impact on forecast results spanning the parameterisations of the soil and the vegetation, boundary layer processes, as well as microphysics.

The system is run in a pre-operational phase since December 2010. The first verification results show the following findings:

For precipitation:

- The ensemble provides additional value to COSMO-DE for all accumulations, lead times, precipitation thresholds,…)
- The ensemble underdispersiveness is relatively small
- Ensemble members may be treated as equally probable
- Additional calibration has good potential

Forecasters‘ feedback has also been collected:

- what they prefer to use:90%-quantile of precipitation
- precipitation probabilities for an area (10x10 grid points)
- they appreciate early signals for heavy precipitation
- indication that deterministic run may be wrong
- they criticize: jumpiness between subsequent runs
- lack of spread in T_2M and VMAX

The future plans are to produce the initial conditions by LETKF and to obtain the lateral boundary conditions by the ICON (new global model) EPS.

Korea Meteorological Administration (KMA)

With the total replacement of KMA NWP system with a new system based on the UK Met Office UM, the medium range ensemble has also been replaced by a new ensemble base on MOGREPS in 2010. The operation of the new ensemble was launched in March 2011 with a horizontal resolution of N320 (40km), and 50 vertical layers. The number of vertical levels has increased to 70 in May 2011. KMA also plans to launch a regional ensemble based on MOGREPS after 2012.

Configurations of the old and new global ensembles at KMA.

EPS name	GBVEPS	MOGREPS(N320 L50 M24)	MOGREPS(N320 L70 M24)
Operation period	Mar 2001- Nov 2010	23 Mar 2011-22 May 2011	23 May 2011 -
Machine	SX5/6, CrayX1E	Cray XE6	Cray XE6
Model	JMA GSM	MO UM Ver 7.5	MO UM Ver 7.7
Data assimilation	3DVar(self-cycle)	Analysis from the global model (4dVar, N320L50, ver 6.6)	Interpolated analysis from the global model (4dVar, N512L70, ver7.7)
Initial perturbation	Breeding + factor rotation	ETKF(updated in 2009, 92 local centres, 4 vertical bands)	ETKF(updated in 2011, 92 local centres, 4 vertical bands)
Model error	NO	RP2, SKEB2	RP2, SKEB2
Membership	17	24	24
Resolution	60km, 40levels	40km, 50levels	40km, 70levels
Forecast length	10days	10days	10days
Initial time	00, 12 UTC	00, 12 UTC	00, 12 UTC
No of CPUs per member	20CPU	240 MPI cores X 2 OMP threads	240 MPI cores X 3 OMP threads
Elapsed time(control)	40mins	55 minutes	60 minutes

China Meteorological Administration (CMA)

Global Ensemble Prediction System (GEPS)

CMA runs an operational global ensemble prediction system (GEPS) twice daily (from 00 and 12 UTC). The forecast model of GEPS is T213 with 31 levels in the vertical, and forecast length is up to day 10. The GEPS has one control run and 14 perturbed members. Initial perturbations are generated by breeding growth method (BGM). Currently, no model perturbation is applied to the GEPS.

A series of typical ensemble products (spaghetti plot, ensemble mean/spread, and probabilities of exceeding certain thresholds) has been using at National Meteorological Centre of CMA, and also are available on the NMC Web site.

Since June 2011, the ensemble products have been distributed to all local weather services office as guidance products, in which the Station Whisker-Box diagram for 39 major cities are included besides the conventional ensemble products.

Regional Ensemble Prediction System (REPS)

CMA Regional EPS at CMA has been operationally running since 1 June 2011.

1. Configuration of REPS
   - Model(WRF-ARW dynamical core)
   - 15km grid space, 31 vertical levels
   - 60h forecast at 00UTC and 12UTC
   - 15 members
   - BGM initial perturbation method
   - multi- physics

2. REPS products
Ensemble mean/spread and probabilistic products for surface parameters (precipitation accumulated over 3h, 6h,12h and 24h, 2m-temperature, 10m-wind, etc.).

3. Forecast data is being put into CMA TIGGE archive center every day.

Japan Meteorological Agency (JMA)

Global EPS

JMA launched its operational EPSs for One-week Forecasts and 5-day tropical cyclone (TC) track forecasts in March of 2001 and 2008, respectively. At present, 11 and 51 initial conditions are integrated by using a low-resolution version of the JMA global NWP model (GSM) for producing an ensemble of 132-hour forecasts in the Typhoon EPS and 9-day forecasts in the One-week EPS, respectively. The EPS model and the initial perturbations (IPs) generator are shared between two EPSs. IPs are generated by Singular Vectors method. Furthermore, a formulation of uncertainties associated to physical processes is employed in the EPS model as a model ensemble method.

In the past two years, four updates were conducted into the One-week EPS. In December 2010, a stochastic physics scheme was introduced as a model ensemble method in order to represent model uncertainties. In March 2011, the amplitude of the initial perturbation was reduced, its seasonal variance was removed, and the IP was expanded into the southern hemisphere. These revisions have made the ensemble spread more appropriate and improved the forecast skill.

For the Typhoon EPS, the IP target area around the central position of TC forecasts was set as a circular region in contrast to the previous rectangular-area settings and the IP amplitude was revised in May 2010. These revisions improved the spread-skill relationship of TC track forecasting. The stochastic physics scheme was implemented in December 2010 to enable use of the same scheme as the One-week EPS.

Graphical data of the One-week EPS for support to WMO Members

JMA has been providing its pilot project service of EPS guidance products, aiming at improving the EPS and increasing the availability of its products, to WMO Members since May 2006. The service includes probability maps of 24-hour accumulated precipitation, spaghetti diagrams for 500hPa geopotential height, and time-series point guidance (EPSgrams) for major cities in Asia. It is reported that some EPS products are used in daily forecast operation up to 3 days at Department Weather Forecast of Agency on Hydrometeorology under Ministry of Emergency Situations of the Kyrgyz Republic.
JMA participates in regional subprojects of SWFDP in both RA V and Southeast Asia, as a global centre. All products are provided in graphical format on the dedicated webpage of JMA for SWFDP.

EPS verification result

The verification results of the One-week EPS are published in annual WMO Technical Progress Report on GDPFS. Furthermore, the monthly verification data are available on the Web site of the WMO/CBS Lead Centre for EPS Verification (http://epsv.kishou.go.jp/EPSv/). In addition, the verification results of the Typhoon EPS are published in Annual Report on Activities of the RSMC Tokyo – Typhoon Center (http://www.jma.go.jp/jma/jma-eng/jma-center/rsmc-hp-pub-eg/annualreport.html).

US National Centres for Environmental Prediction (NCEP)

NCEP Global Ensemble Forecast system (GEFS)

NCEP GEFS is scheduled for upgrading in January 28 (tentative schedule) 2012. The main changes include:

1. The horizontal resolution will be increased from T190 (70km) to T254 (about 55km on equator) for 0-192 hours and the same for 192-384 hours (T190), the vertical resolution will be increased from 28 to 42 hybrid levels.
2. Increasing the size of lower level initial perturbations (near surface) by approximately 20% (maximum inflation factor), and gradually eliminate this inflation by near 500hPa.
3. Tuning Stochastic Total Tendency Perturbation (STTP) scheme to optimum for new GFS model.
4. The ensemble memberships, output frequencies, output resolutions are unchanged from this implementation.

The improvements of this upgrade could be summarized as: 1) improvement overall probabilistic forecast skills (for example of 500hPa height and 850hPa temperature); 2) significant improving tropical storm (TC) track forecast by 15-20% (summer 2011 season).

Short-Range Ensemble Forecast (SREF) system

Next SREF is planning for upgrading in March/April 2012 time frame (tentative schedule). The new configuration could be summarized as following:

I. Model Change:
    a. 4-model system becomes 3-model system (adding new model NMMB and getting rid of old models Eta and RSM)
    b. Model's horizontal resolution increases from 32km to 16km
II. IC Diversity Improvement:
    a. Use more diversity of control analyses: from 2 to 3

   b. Improve IC perturbation by blending larger-scale ETR (Wei et al. 2008) and smaller-scale BV (Toth and Kalnay 1997)
   c. Change 2-D mask to 3-D mask to control IC perturbation vertically

III. Physics Diversity Improvement:
   a. More diversity by adding wide range of various physics schemes and possibly adding stochastic parameterization physics scheme, too.

IV. Ensemble Product Improvement:
   a. Precipitation bias correction (probability-matching method, Ebert 2001 and 2004)
   b. Clustering (Tracton 1995, personal communication; Alhamed et al. 2002)
   c. Statistical downscaling to 2.5km by using high resolution RTMA (Cui et al. 2012)
   d. Many new ensemble products including Max/Min temperature, 10-25-50-75-90% probabilistic forecasts, best/worst member and associated weighted mean (Du and Zhou 2011), extreme weather probability as well as wind, energy, fire weather and convection-specific probabilistic products (Bright et al. 2005, 2006 and 2009).

Canadian Meteorological Centre (CMC)

Global EPS

Since 1996, CMC produces medium range ensemble forecasts operationally. The official forecasts of the Meteorological Service of Canada for days 6 and 7 are produced with the outputs of the GEPS. It is worth mentioning that this forecast system is also an integral part of the collaboration with the National Centres for Environmental Prediction (NCEP) in the North American Ensemble Forecasting System (NAEFS).

The data assimilation consist of a 6-hour cycle using 4 times 48 configurations of the GEM model providing (192) trial fields over 6-hour time windows. It is a 4-D data assimilation cycle using Kalman filter (so-called Ensemble Kalman Filter). The lid was recently lifted from 10 hPa to 2 hPa, and the system is now using a newly implemented GEM model's dynamical vertical discretization.

From the 192 Ensemble Kalman filter analysis, 20 are randomly selected to provide initial conditions for 20 members (forecast model). The models are integrated out to 16 days, twice a day (00z and 12z), at 0.6 degrees horizontal resolution over 40 levels with a lid at 2 HPa (nearly 40km).

The initial condition uncertainties are represented by the perturbed ensemble Kalman Filter data assimilation, while forecast model uncertainties are reproduced by various model physic's configuration, with a single dynamic core (GEM model). Also stochastic perturbations are added to tendencies in the parameterized physical processes and back-scattering energy parameterisation is used in order to augment the spread of the ensemble.

The model's lid was recently lifted from 10 hPa to 2 hPa, and the system is now using a newly implemented GEM model's dynamical vertical discretization. The latest upgrade is detailed at the following link (the Wednesday August 17, 2011): http://collaboration.cmc.ec.gc.ca/cmc/cmoi/product_guide/docs/lib/op_systems/doc_opchanges/technote_geps_20110906_e.pdf

EPS-LAM

The Canadian Regional EPS (REPS) is producing forecasts up to day 3 with the limited area (LAM) version of the Canadian Global Environmental Multi-scale model, GEM-LAM. It has been implemented in operations in September 2011

Initial conditions for the REPS at 00 and 12 UTC are interpolated from the Global EPS Ensemble Kalman Filter analysis. Similarly the lateral and top boundary conditions are provided by the GEPS and they are updated every three hours. The top boundary conditions are given by a piloting method.

The REPS consists of twenty (20) GEM-LAM members at 33km horizontal resolution and 28 levels, with a lead time of 72 hour. On the contrary to the Global EPS, the REPS uses a single GEM-LAM model configuration, which is equivalent to the Canadian 33km Global deterministic forecast model. However, it uses stochastic perturbations of physical tendencies and surface parameters in order to augment the spread of the ensemble.

The first operational implementation is detailed at the following link (the Wednesday September 22nd, 2011): http://collaboration.cmc.ec.gc.ca/cmc/CMOI/product_guide/docs/lib/op_systems/doc_opchanges/technote_reps_20111004_e.pdf.

ANNEX IV

UPDATED GUIDELINES ON EPS AND FORECASTING

1. Introduction

Ensemble Prediction Systems (EPS) are Numerical Weather Prediction (NWP) systems which allow us to estimate the uncertainty in a weather forecast as well as the most likely outcome. Instead of running the NWP model once (a deterministic forecast), the model is run many times from very slightly different initial conditions. Often the model physics is also slightly perturbed, and some ensembles use more than one model within the ensemble (multi-model EPS) or the same model but with different combinations of physical parameterization schemes (multi-physics EPS). Due to the cost of running an NWP model many times, the EPS is normally run at around half the horizontal resolution of the equivalent deterministic NWP model. The EPS normally includes a control forecast which uses the ensemble resolution model but without any perturbations to the analysis or model. The individual NWP solutions which make up the ensemble are often referred to as the ensemble *members*. The range of different solutions in the forecast allows us to assess the uncertainty in the forecast, and how confident we should be in a deterministic forecast. The uncertainty in a weather forecast can vary a lot from day to day according to the synoptic situation, and the EPS approach provides an estimate of this day-to-day uncertainty. The EPS is designed to sample the probability distribution function (pdf) of the forecast, and they are often used to produce probability forecasts – to assess the probability that certain outcomes will occur.

These guidelines are intended to provide some general advice to forecasters and forecast providers on the effective use of EPS, and on what EPS can and cannot be expected to provide. A general working knowledge of the principles and use of NWP is assumed. For those requiring more detailed information, the ECMWF User Guide (http://www.ecmwf.int/products/forecasts/guide/) provides comprehensive guidance on the use of ECMWF systems including detailed advice on the use of EPS; the COMET training materials (http://deved.meted.ucar.edu/nwp/pcu1/ensemble/) also provide training on the use of EPS.

In general, it is strongly recommended that uncertainty should be communicated as part of every forecast. Guidance on the communication of uncertainty is given in the PWS Guidance (WMO TD-1422).

Examples shown in these guidelines are mostly taken from the UK Met Office's MOGREPS EPS systems, or the ECMWF EPS, but the principles described apply to any EPS.

2. Why should we use EPS?

NWP systems using the latest numerical models of the atmosphere are very powerful systems to aid the forecaster in producing weather forecasts. Many models now provide a good enough representation of the weather that they can also be used to provide basic automated weather forecasts from Direct Model Output (DMO), although in general it is recommended that some post-processing should be used to calibrate automated forecasts. DMO provides a better representation of some weather elements than others, for example surface temperature is often quite well-resolved (at least away from steep surface orography) whereas precipitation is often much less well resolved.

However, despite these advances, it is well-known that forecasts from even the very best models can often go badly wrong. This is most obvious in forecasts several days ahead and is due to the *chaotic* nature of the atmosphere. We forecast the weather by starting the model from an analysis of the state of the atmosphere based on the latest observations which are taken all around the world. The model then calculates how the atmosphere will change and evolve from this initial analysis state over the coming days. Chaos theory means that the way the atmosphere evolves is very sensitive to small errors in that initial analysis, so that a tiny error (often too small for the forecaster to even notice) can grow into a large error in the forecast. Even with the best observations we can never make a perfect analysis, so we cannot make perfect forecasts. This is why we run EPS (ensembles).

In an ensemble forecast we make very small changes (perturbations) to the analysis, and then re-run the model from these slightly perturbed starting conditions. If the different forecasts in the ensemble are all very similar to each other then we can be confident of our forecast, but if they all develop differently, and for example some develop a major storm while others develop a much weaker depression, then we will be much less confident. However, by looking at the proportion of the ensemble members which predict a storm, we can make an estimate of how likely the storm is. When we look at shorter-range forecasts of 1 or 2 days ahead, the general pattern of the weather is usually much more predictable, but we can still find important differences between ensemble members when we look at the local detail of the weather which may be important to many forecast users. Also, occasionally the larger-scale evolution can be uncertain even at short-range – this is most likely to happen during the development of major storms, so it is important to take account of the EPS even in short-range forecasts.

3. Types of EPS

Ensemble Prediction Systems for use in weather forecasting come in three main types, global, regional and convective-scale, and as with deterministic NWP models, they address different time-scales in the forecast. These will be outlined briefly below. Within each of these categories there are many variations, such as the way in which perturbations are created and the variations in the models used within the models – however the principles of how the ensembles are used remains the same, and these details are not covered here. (It may be noted that ensembles are also used for long-range forecasting and climate prediction. The principles are very similar, but these will not be considered in these guidelines which focus on forecasts of up to 15 days, which is the period over which it is often possible to forecast daily weather.)

3.1 Global EPS

Global EPSs are normally designed and used for medium-range forecasting of 3-15 days ahead. They use global NWP models and are run at relatively low resolutions with typical grid-lengths of between 30 and 70km. Although they are primarily designed for use in the medium-range, their global coverage means that they can also be used to provide short-range EPS forecasts in regions of the globe where no other EPS are available, and may be the only available option for many WMO members. In this context they are used extensively to provide products to support the WMO SWFDP (Severe Weather Forecasting Demonstration Project) projects.

Forecasters using global EPS should always remember that the relatively low grid resolutions will limit the detail they can expect in the forecasts. Global EPS will often not be able to resolve details such as the full strength of wind speed in a storm.

3.2 Regional EPS

Regional, or LAM (Limited Area Model) EPSs use regional models over smaller areas and are focussed more on the short-range forecast of 1-3 days ahead. They use higher grid length resolution than the global EPS, typically between 7 and 30km, which allows them to forecast more local detail in the weather and also to better resolve intense weather systems. Nevertheless the forecaster should still remember the limitations of resolution, so for example a regional EPS should not be expected to predict details of small-scale systems like thunderstorms.

A regional EPS has to take its lateral boundary conditions (the weather systems moving into the area from outside the domain) from a global EPS. Some regional EPS systems use a high-resolution regional analysis and calculate corresponding high-resolution perturbations, but others simply take the initial conditions and perturbations from the same global EPS which provides the boundary conditions – this is normally referred to as *downscaling*. In a downscaling EPS the forecast needs to run for a number of hours before the model can “spin up” the higher resolution detail.

3.3 Convective-scale EPS

Convective-scale NWP, with model grid-lengths of 1-4km run over relatively small domains, is now available in a number of more advanced NWP centres. These models, sometimes referred to as *convection-permitting* are able to resolve some of the detail of large convective systems, and thus can attempt to predict details such as the location and intensity of thunderstorms. While this offers great potential for improved forecasts, convective systems evolve very rapidly and have short predictability timescales, so the forecasts can rapidly be affected by chaos. EPS is therefore highly relevant to convective-scale NWP because convective instability adds a new scale of forecast uncertainty not resolved by the lower resolution models, and with much shorter timescales.

In addition to convection itself, models on this resolution have greatly enhanced capability for forecasting other aspects of local weather, such as low cloud and visibility of interest to aviation. Many of these phenomena are significantly affected by topographic forcing which may give enhanced predictability when that forcing (slopes, coastlines, vegetation, albedo, ...) can be resolved by the models (e.g. convective initiation or valley fog). Convective scale EPS has the potential to provide information on the predictability of all these weather elements.

At the time of writing in 2011 convective scale EPSs are under development at various centres. DWD runs the COSMO-DE-EPS with a resolution of 2.8 km in preoperational mode since December 2010. The UK Met Office and Météo-France have plans to introduce such systems in the near future, and research is being conducted in other countries.

Due to the very high cost of running convective-scale EPS they are unlikely to be available outside the producing nations for many years, and experience of them is still very limited. They are discussed only briefly in these guidelines.

The much higher resolution of convective-scale EPS is expected to allow better resolution of many weather phenomena than is possible with global and regional EPS, for example local winds forced by topography and possibly elements like low cloud and visibility, especially where such phenomena are forced by local details of the topography or land surface.

For precipitation the models are likely to better resolve the intensity and spatial scales of local precipitation, especially in convective precipitation. However to sample the full range of uncertainty in convective precipitation would require very large ensembles with hundreds or thousands of members, which will not be affordable in the foreseeable future. It is therefore strongly recommended that convective-scale EPS is post-processed using techniques such as neighbourhood processing (where it is assumed that a feature such as a convective shower may be realistic but may be misplaced and occur anywhere around the neighbourhood within, say, 10 grid-lengths of where it appears in the model) to provide a more realistic spatial distribution of probabilities. Similar techniques may also be appropriate for other variables, to take account of the small size of the ensembles.

4. Standard EPS Products

This section describes some of the standard EPS products which are generated from most EPS systems, and briefly how they may be used.

4.1 Basic Direct Model Output Product Generation

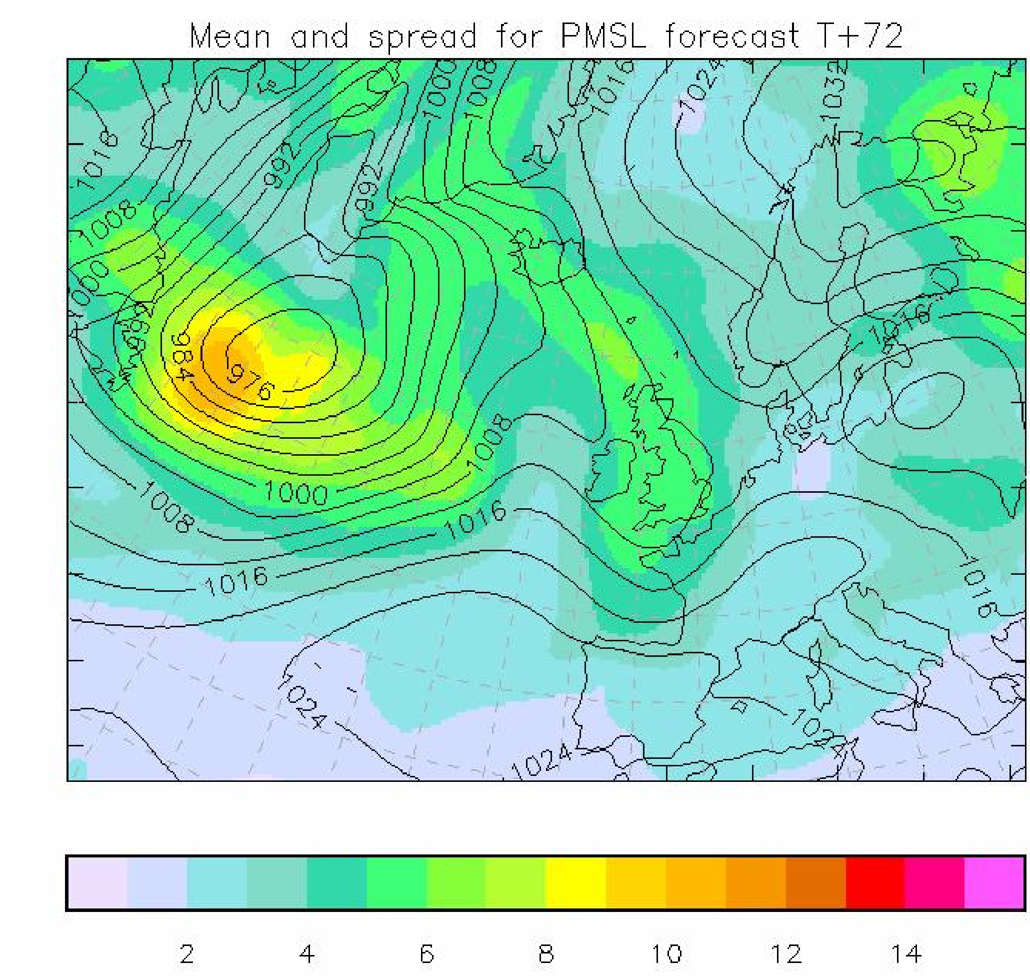
A range of basic products are produced from most EPS systems directly from model output fields. These typically include:

4.1.1 Ensemble Mean

This is a simple mean of the parameter value between all ensemble members. The ensemble mean normally verifies better than the control forecast by most standard verification scores (RMSE, MAE, ACC, etc) because it smoothes out unpredictable detail and simply presents the more predictable elements of the forecast. It can provide a good guide to the element of the forecast which can be predicted with confidence, but must not be relied on its own as it will rarely capture the risk of extreme events.

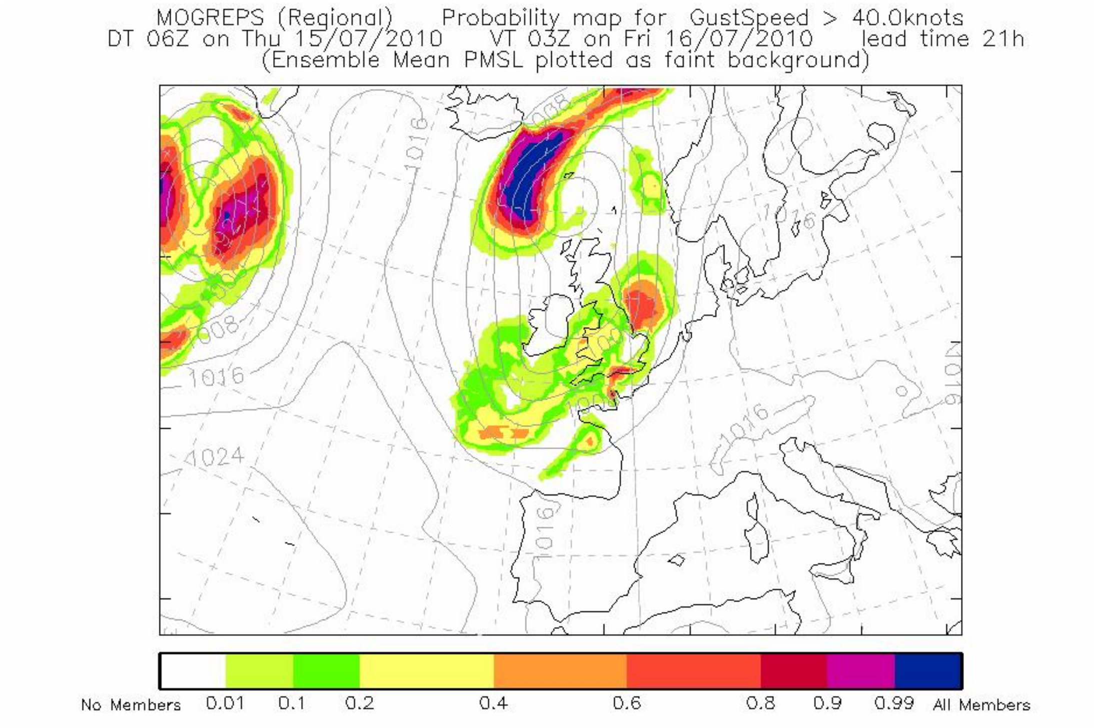
4.1.2 Ensemble Spread

This is calculated as the (non-biased) standard deviation of a model output variable, and provides a measure of the level of uncertainty in a parameter in the forecast. It is often plotted on charts overlaid with the ensemble mean. The example below shows both ensemble mean PMSL as black contours and spread of PMSL as colour shading. The areas of strong colours indicate larger spread and therefore lower predictability.



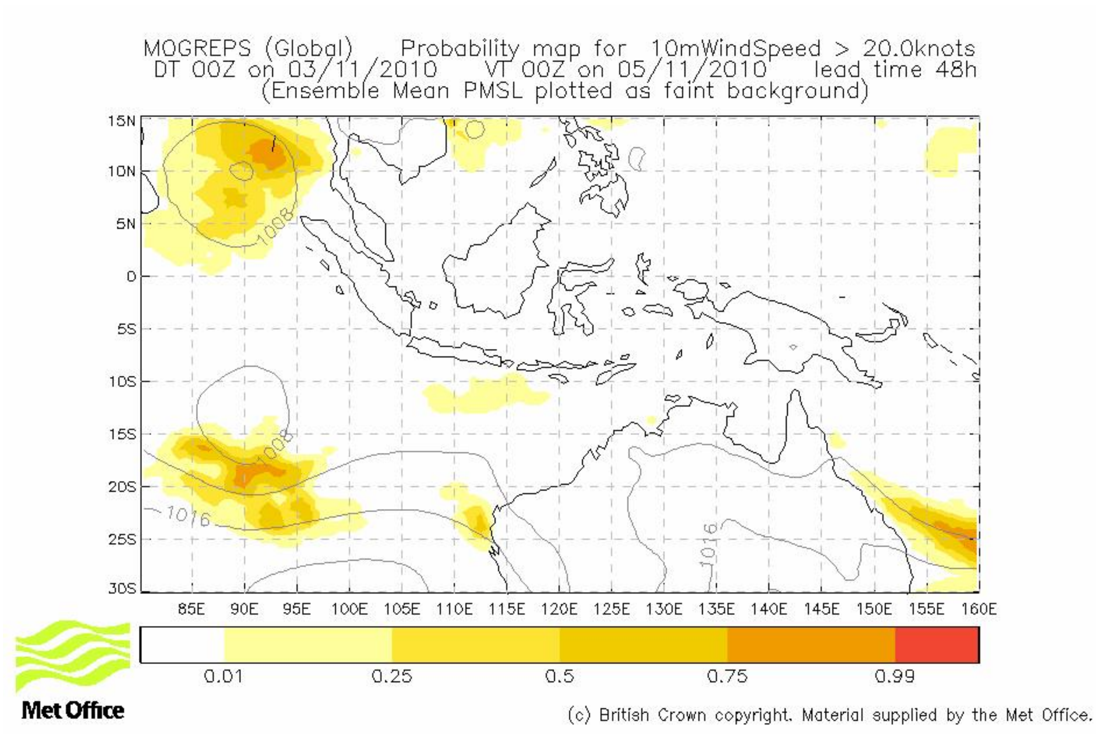
4.1.3 Basic Probability

Probability is frequently estimated as a simple proportion of the ensemble members which predict an event to occur at a particular location or grid-point (e.g. 2m Temperature less than 0 Celsius, or more than one standard deviation below normal). The example shown below shows the contoured probability of wind gusts exceeding 40kt. The ensemble mean pressure at mean sea-level (PMSL) is also included as grey contours.



It should be noted that this definition of probability is not a true Bayesian probability as would be defined by a statistician, but provides a useful estimate for practical purposes. It makes an assumption that the model accurately reflects the climate distribution of occurrence of an event. Probability forecasts produced in this way should always be verified over large samples of cases to determine the extent to which forecast probabilities relate to observed frequencies.

The second example shown below is one of those produced for the SWFDP project in the South Pacific.

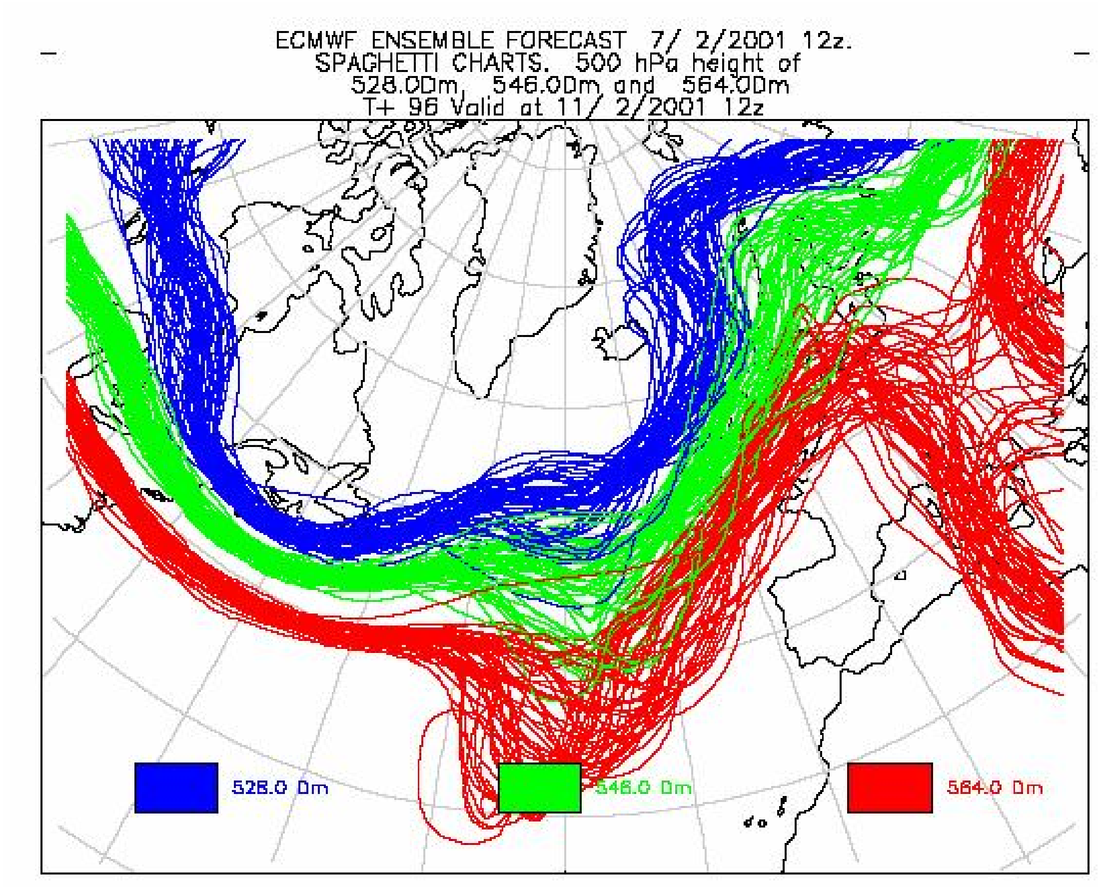


4.1.4 Quantiles

A set of quantiles of the ensemble distribution can provide a short summary of the uncertainty. Commonly used quantiles are the Maximum and Minimum of the ensemble distribution, and the 25th, 50th (median) and 75th percentiles. Others often used include the 5th, 10th, 90th and 95th percentiles.

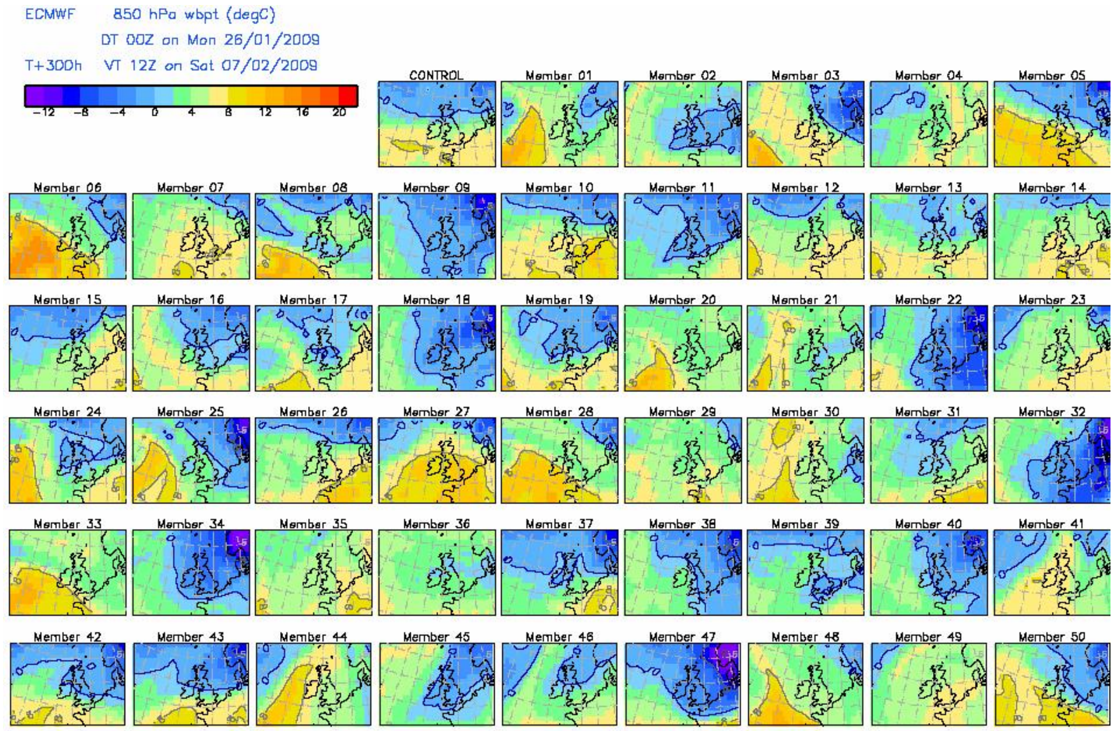
4.1.5 Spaghetti Maps

Charts showing a few selected contours of variables (e.g. 528, 546 and 564Dm contours of 500hPa geopotential height) from all ensemble members can provide a useful image of the predictability of the field. Where all ensemble member contours lie close together the predictability is higher; where they look lie spaghetti on a plate, there is less predictability.



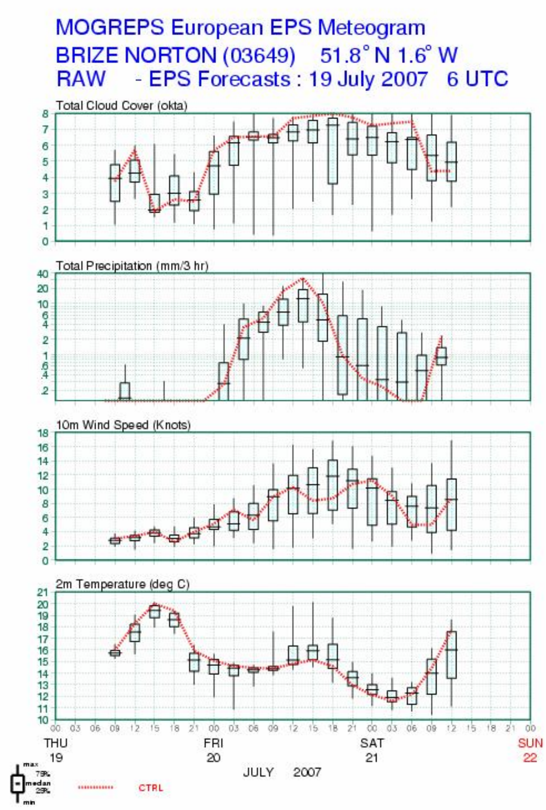
4.1.6 Postage Stamp Maps

A set of small maps showing contoured plots of each ensemble individual member allows the forecaster to view the scenarios in each member forecast, and assess the possible risks of extreme events. However this presents a lot of information which can be difficult to assimilate.



4.1.7 Site-specific Meteograms

Model output variables can be extracted from the grid for specific locations. There are many presentations which can be used to represent the forecast at locations, such as plume charts, Probability of Precipitation etc. One of the most commonly used is the Ensemble Meteogram (or EPSgram) which uses a box and whisker plot to illustrate the main percentile points of the forecast distribution for one or more variables.



5. General Comments Applying to all use of EPS

This section presents a number of general principles which apply to all use of EPS. Following sections provide more detail on the specific use of EPS for particular types of forecast production.

- An EPS best represents the uncertainty in resolved variables
    - Upper-air usually more skilful than surface
        - Surface parameters are affected by sub-grid scale uncertainty not resolved by the model
    - As resolution and model performance increases, the ability to predict surface weather parameters is continually improving
- An EPS is only as good as the model(s) it uses.
    - If a model is unable to represent certain phenomena, the EPS will also be unable to represent it.
        - A good example is that most ensembles cannot resolve convective storms, which is one of the reasons why some centres are developing ensembles at convective scale.
    - An EPS will share any systematic biases of the model used.
- How to combine deterministic forecast with ensemble/probabilistic?
    - Relative capabilities of ensemble members compared to hi-res/control
    - See the "Guidelines on using information from EPS in combination with single higher resolution NWP forecasts (February 2006)"
- A common question is whether a forecaster can improve the distribution by re-weighting members (e.g. the high-resolution control forecast if included) or by rejecting some members?
    - Forecasters may think that some members are unrealistic
    - Can we eliminate some members on the basis of recent observations or pick a "best member"?
        - PERHAPS, for certain aspects of the forecast over very short-period forecasts and for local forecasts over a small area
            - Over a large area or the full model domain, the control forecast will always be the most skilful.
        - NOT for longer period forecasts
    - This type of approach is subjective and difficult.
    - It is strongly recommended that forecasters should use the whole EPS distribution in a probabilistic approach.

- Strengths and weaknesses of the models/ensembles available to the forecaster should be known. Documentation should be easily available to the forecaster.
    - Verification of multiple thresholds to be available
    - Summary doc of strengths and weaknesses by season
- Be careful with "end of chain" diagnostic parameters (e.g. precipitation, cloudiness,…). For instance look at distributions of indices in convective situations.
- Forecasters should not always rely on direct model output of weather variables, but should also consider analysis of better resolved diagnostics which may aid interpretation of the EPS forecast (e.g. synoptic features,, environment/precursors/potential for high-impact weather developments such as moisture convergence, low level jets, development regions, convective diagnostics etc).

The use of EPS (and other probabilistic tools) opens the possibility of issuing two different types of forecast, fully probabilistic, or deterministic with supplementary uncertainty information (for instance confidence). Which type we use affects who makes decisions from the forecast. In general the use of fully probabilistic forecasts allows each user to tailor their decision to their specific needs (e.g. using cost-loss estimation), and is therefore strongly encouraged.

6. Use of EPS in Deterministic Forecasting

In general it is strongly recommended that probabilistic forecasts provide the best and most complete weather forecast for customers, and should be encouraged, especially at longer lead-times. However it is recognized that many customers demand a simple deterministic forecast, and where a deterministic forecast is to be produced, the use of an EPS can often provide a more reliable forecast than a single deterministic NWP run. This is particularly true for forecasts more than 1-3 days ahead, and can help reduce jumpiness from run-to-run of the forecast system at any time range.

Several indicators from the EPS can be used to optimize the deterministic forecast. The ensemble mean will on average score the best by many standard verification scores, but it must be remembered that it will tend to smooth out the smaller scale unpredictable detail, and will rarely capture the intensity of important high-impact weather systems. The ensemble mean should not therefore be used on its own if the use of the forecast is concerned about potential severe weather impacts. Other useful guides to the most likely forecast can be the median (central point in the pdf) or mode (most likely value in the pdf) – these are easier to identify for single weather parameters than for the complete forecast picture.

If a deterministic forecast is to be issued, it may sometimes be augmented by a statement of the confidence of this forecast to take some advantage of the uncertainty information available. The confidence will not always be the same for all elements of the same forecast. Confidence indices, if used, are best provided separately for each variable. The confidence level should be based on the spread of the ensemble, but also considering the known forecast skill limitations.

The best approach to issuing a deterministic forecast will depend on the predictability as indicated by the ensemble spread. The spread could be analyzed using various products such as, spaghetti plots, and map depicting variance at the synoptic scale and then, at the lower scales, using meteograms, quantiles, cluster analysis, etc. :

- Small spread in the ensemble (good predictability)
    - In this case it may be reasonable to offer more detail in the forecast.
    - Take the control, the high resolution control, the ensemble mean or the median as a guide (with due regard for the need for calibration or bias correction).
    - Spread may often differ between model variables so small spread in one parameter does not guarantee confidence in all aspects of the forecast.
        - Good synoptic scale predictability does not always mean predictability in surface weather variables such as temperature or convective precipitation.
        - Forecaster should still take account of uncertainty in parameters not resolved by the model.
- Large spread in the ensemble (poor predictability)
    - Avoid giving too much detail in the forecast
    - Ensemble mean should be considered but if the ensemble covers a range of scenarios the ensemble mean will not provide a realistic scenario
    - So in that situation, take most representative member of the ensemble (e.g. most populated cluster or mode of pdf) as a guide to the most probable outcome
        - Note that the most representative ensemble member may not give the most probable value for each weather element (e.g. most probable temperature at a location may not be correlated with the most probable precipitation amount.)
    - The uncertainty assessment
        - Encourage users to follow forecast updates.
    - Take into account extremes of the EPS and of the high resolution control
        - Make a careful evaluation of the possible evolutions of the synoptic situation and their potential impacts.
        - Take into account the behavior of models.
            - The high-resolution control may be better able to represent certain high-impact events.

- In the short range (12 - 18 hours), it may be possible to take into account the latest observations (3-6 hours into the forecast) in order to choose a scenario or a member of the ensemble
  - For example, a rapidly evolving cyclone may be best predicted by the member with the best position after a few hours *but ONLY* in the very short-range!
  - Be aware that future evolution will be influenced by features coming from upstream. This makes member selection for forecasts beyond ~24h impossible.
  - Also the consistency of the latest runs with respect to the previous is a factor to take into account.
- In the longer range, while probabilistic forecast are best suited, if a deterministic forecast is to be produced, the use of the ensemble mean or median could yield more reliable forecasts, with less jumpiness between runs of the forecast.

6.1 Decision Making from deterministic forecasts

Weather forecasts are only useful when people make decisions from them. It is often argued that it is easier to make a decision from a deterministic forecast than a probabilistic one. However when the forecaster issues a deterministic forecast the underlying uncertainty is still there, and the forecaster has to make his/her best guess at the likely outcome. Unless they fully understand the decision that the user is going to make based on the forecast, and the impact of different outcomes, then the forecaster's "best-guess" may not be well-tuned to the real needs of the user.

- The choice of making a deterministic forecast for a specific event to occur should not be taken without some knowledge of the needs of the end user. An optimal decision cannot be made without the cost-lost ratio of the user. This ratio can be assessed by a survey or a direct discussion with the end user.
- When appropriate the forecasters should convey the risks and impacts associated with worst-case scenarios alongside the most likely outcome

7. Scenarios

A useful way to summarise the uncertainty in a weather forecast can be to describe a small number of possible outcomes, or scenarios, rather than giving the full detail of a probabilistic forecast. For some customers used to receiving deterministic forecasts, this may be more acceptable. Ideally the EPS can be used to estimate the relative likelihood of the different scenarios presented. In most cases, to avoid confusion, the best approach may be to issue a most-likely scenario based on the advice above on issuing deterministic forecasts, plus a single alternative scenario – this may often be a *worst-case* scenario, perhaps reflecting a low probability but high-impact possibility suggested by the most extreme ensemble members. However care should be taken not to give the impression that either scenario will be correct – the truth could easily lie somewhere in between (or even be different again!)

Useful tools to aid in issues alternative scenarios are postage-stamp maps (4.1.6) which show the forecaster all the individual forecasts in the ensemble, or clustering (9.3) which automatically groups the ensemble members and provides the forecaster with an objective assessment of the possible scenarios.

8. Full Probabilistic Forecasts

Wherever possible, the use of a full probabilistic approach is recommended in issuing forecasts. This provides a full representation of the uncertainty information provided by an EPS, and also allows users to tune their decision-making to take account of their particular applications.

Probabilistic forecasts can be expressed in a number of ways, and need not always use the word *probability*:

- A forecast of a weather variable provided with error bars which vary according to the ensemble spread.

- A fuller representation of the ensemble distribution showing a number of percentile values, as used in the standard meteogram product.
- Probabilities of specific (well-defined) events occurring, expressed as numbers or as contoured shading on a map.

When a forecast is presented as a probability, it is very important to express very clearly *what* the probability is for, so that both the forecaster and the user is clear and understands. We often talk about the probability of an *event* occurring, and it is this event which must be defined. Often the event will be for a threshold value to be exceeded (e.g. more than 50mm of rain, or temperature below 0 Celsius). Ideally it will be something which has an important impact for which someone will have to take a decision (e.g. the probability that ice will form on roads so that road treatment will be required). It is also important to define when and where the event is forecast for:

- Exact time, or time period which the forecast refers to.
- Exact location or area which the forecast applies to.
    - If it is an area, does it mean a forecast that the threshold will ne exceeded somewhere in the area, or everywhere in the area?

A good test of whether an event is well-defined is to ask yourself whether you could easily measure whether the event does happen or not (in other words, could you verify the forecast). If you cannot easily say, then you may need to define the event better.

The following bullets provide a number of issues which should be considered when basing probabilistic forecasts on EPS outputs:

- Calibrated, bias corrected forecast can be directly issued to the end user (low cost).
    - This approach allows for the possibility of issuing automated forecasts for many locations and users.
    - Methods for bias correction and calibration are discussed in section 9.
- Direct model output (DMO) from ensembles should be used with care, as it may not provide reliable probabilistic forecasts, but will often nevertheless provide valuable information. In some cases use of DMO may be essential where there is no calibration system in place – calibration is difficult for certain variables such as precipitation, or where adequate observations are not available.
- To generate probabilistic forecasts of outcomes dependent on more than one weather element, it is important to calculate this outcome for each ensemble member and then combine members to create the probabilities. This retains consistent correlations between different weather variables and also different locations (e.g. the correlation in temperature between two locations). Calibration or post-processing may spoil this consistency.
    - This principal also applies when using the ensemble to drive downstream impact models (e.g. hydrological models) where the downstream model should be run for each ensemble member and then the probability of the downstream impact calculated.
- In “usual” situations, forecasters should not try to change the probabilistic forecasts issued by the EPS (DMO or post-processed). The forecasts can be issued directly to the public. Forecasters should target their attention to “unusual” situations.
- In “unusual” situations, probabilistic forecasts can be adapted by the forecasters using experience, analogues, conceptual models,…. Forecasters may be able to correct for some known system biases or model weaknesses. The corrections should be made by using the guidelines mentioned in section I.
- Studies have shown that the general public is able to make better decisions, when presented with uncertainty information in forecasts than with a deterministic forecast. When uncertainty information is not provided people make their own assumptions.
- Probabilities have to be presented in a comprehensive graphical form. Examples and guidelines are given in the PWS document PWS-18, WMO/TD No. 1422. (Homework for Alice: include a few beautiful graphics).
- Probabilities of events relevant to specific applications should be defined. This includes, for example, application in agriculture in which the occurrence of dry spells or rainy periods influences irrigation, seeding, harvest…

- Risk is a combination of impact and likelihood of a phenomenon which can be produced by the EPS. It gives an objective and valuable decision basis to the forecasters in order to assess different warning levels. Impact has to be agreed with the relevant authorities (PWS customers). Climatology usually provides a good reference to establish the thresholds of phenomena which produce impact. The thresholds can be adapted taking into account the recent evolution of the various environmental parameters (recent rainfall accumulations affect soil saturation, leaf-cover on vegetation, snow cover etc.).
- It is recommended that where probabilities are indicated for significant high-impact weather, a forecaster-written comment or warning should be added.

9. Post-Processing

The aim of this guidance is to provide explanation and advice for post-processing using statistical dynamical and other approaches to improve EPS outputs. There are numerous approaches and the paragraphs below capture some of the most common. Some methods are quite generic and may be best applied by EPS producers at source, while others are quite specific to applications and may be better applied specifically for individual users.

9.1 Statistical post-processing

Generally speaking statistical post-processing is needed in order to correct systematic errors in models and thereby add value to direct NWP model output. These errors are particularly important for surface parameters (e.g. 2m temperature, 2m humidity, 10m wind speed, precipitation, total cloudiness, …) and are linked to local conditions.

More precisely, statistical post-processing can be used to:

- Remove systematic biases
- Adjust ensemble spread
- Quantify uncertainty not represented directly by the EPS
- Predict what model does not represent explicitly (e.g. low visibility)

In general statistical methods are easier to apply to some types of model output variable than others. Temperature is often relatively easy, for example, as it is a continuous variable and varies relatively smoothly in model fields, and most importantly temperature errors are often approximately normally distributed. Precipitation, by contrast, is particularly difficult because precipitation fields often have much multi-scale structure which is poorly represented by models, especially on the small scales. Its climatological distribution, and hence the distribution of forecast errors is bounded at zero at one end and often highly skewed, making it much more difficult to represent statistically. The problem can sometimes be reduced by transforming the distribution to make it more quasi-normal, but in general post-processing methods for precipitation are much less effective than for other variables.

9.1.1 Bias correction of the First Moment of the PDF (Probability Density Function)

This post-processing is similar to MOS (Model Output Statistics) methods applied for single models, but with some important differences. For ensembles, it is well known that a traditional MOS which is trained specifically for each forecast lead-time will lead to a significant decrease of the ensemble spread at longer lead times. Instead, it is recommended to use a pseudo-perfect prognosis approach. This method is based on the use of MOS statistical models computed over the first 24h of the forecast and then applied to the corresponding steps during at all forecast lead-times.

Adaptive methods such as the Kalman-filter are recommended to allow the corrections to be automatically updated to account for model changes (upgrades) and changes in the season.

In the case of single-model ensembles (i.e. the same model is used for all of the members, even where model perturbations are implemented) the same statistical model should be trained using the control forecast and applied to all members of the ensemble.

In the case of multi-model or multi-physics ensembles (i.e. where different models are used to build the pdf, or systematically different model versions are applied, e.g. different parameterization schemes) specific statistical models should be trained and applied for each model version.

In either case the development of these statistical models need a training set of model outputs (predictors) and observations (predictands). In the case of adaptive methods such as the Kalman filter this training set is updated continuously from the daily forecasts.

The “observations” can be either site-specific observations or may be the best available set of analyses. In the case of site observations the statistical post-processing will lead to local forecasts (i.e. at each site specific point where observations are available). When analyses are used the end product is a bias-corrected and downscaled gridded forecast.

It should be noted that when different weather variables are independently bias-corrected, some of the correlation between variables represented by the different ensemble members may be lost. For this reason forecasters may prefer to view direct model outputs.

9.1.2 Calibration of higher moments of the PDF

Bias removal for the second moment of the pdf is often known as “calibration”. It aims to improve the reliability of the probabilistic forecast. Therefore this kind of post-processing is specific to ensemble prediction systems and is particularly important to optimize probability forecasts. As for the first moment bias correction, calibration is based on local conditions and requires high quality observations or analyses as a reference.

A number of methods are under development which attempt to calibrate both the first and second moments of the pdf to optimize the complete distribution.

- A method developed at the University of Washington is now considered as one of the best to deal with this issue. This method, called “Bayesian Model Averaging” is based on specific statistical assumptions (e.g. normal distribution for temperature).
- EKDMOS (Ensemble Kernel Distribution Model Output Statistics) is another technique which has been implemented in the USA.

The above methods are commonly applied to variables such as temperature and wind-speed. Variables such as precipitation are more difficult to correct due to the nature of the pdf and the local variability of observations. Some specific approaches are under development, but post-processing methods are at present less successful and may not improve significantly over raw model outputs.

It must be noted that there are limitations to the potential of statistical post-processing especially in the case of severe events. Commonly calibration will improve the statistical reliability of probabilistic forecasts (the match of forecast probabilities to frequency of observations of the event) but reduce the resolution of the forecasts (the ability to discriminate whether an event will occur or not). Sometimes it is found that calibration will improve forecasts of common events, but degrade the probabilities of more extreme events. The main reason for this is that observations of these kinds of events are rare, and the statistical distributions are trained to the more common events. Therefore calibration cannot be expected to provide significant improvement over the raw forecasts in this case.

Some attempts have been made to develop post-processing explicitly for prediction of more extreme events, for example first-guess severe weather warning systems. In these cases the systems can be calibrated specifically to optimize the reliability for extreme thresholds.

Nevertheless, human expert interpretation remains particularly important for assessment of the risk of extreme events.

9.2 Downscaling

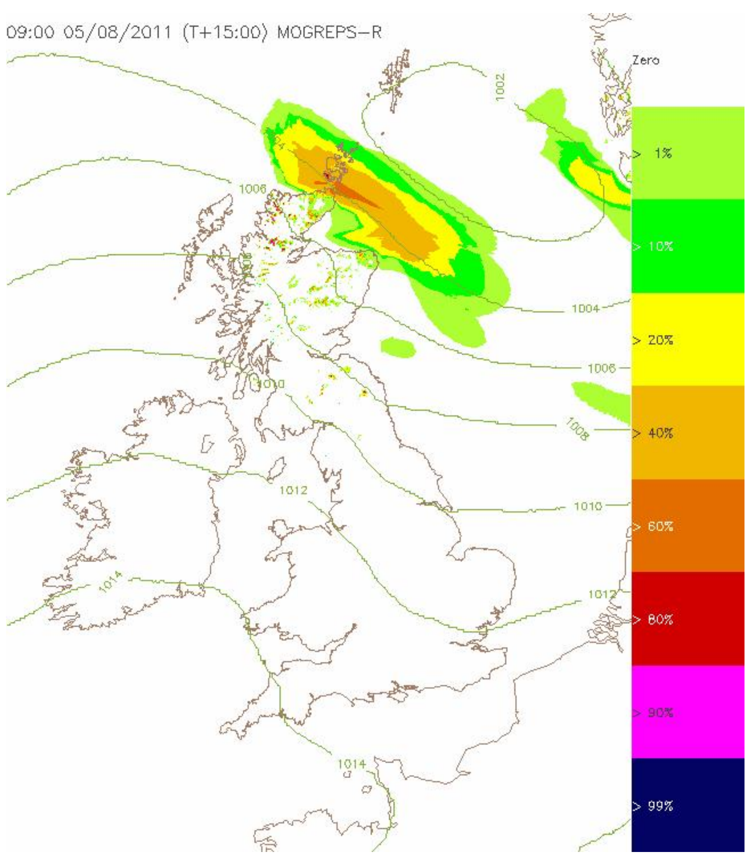
A number of methods may be used to add some local detail to forecasts generated with lower resolution models, and these techniques may be applied to EPS forecasts just as with deterministic NWP.

9.2.1 Dynamical Downscaling

Dynamical downscaling may be defined as the use of a higher resolution limited-area NWP model to add detail forced by topographic detail and to resolve fine-scale processes such as convection. Ideally all ensemble members will be downscaled, but where cost constraints prevent this, a selected set of members may be downscaled. In dynamical downscaling, the initial conditions, boundary conditions and perturbations are taken directly from the lower resolution EPS members. Care must be taken to ensure that the downscaling is appropriate to ensure good performance of the high-resolution model, e.g. appropriate ratios of grid sizes, rate of updating of boundary conditions etc. The model performance should be carefully tested over the domain. Many LAM and convective-scale EPSs are dynamical downscaling systems from global ensembles.

9.2.2 Topographic downscaling using simple physical models

For some parameters such as 2m temperature and 10m wind speed a simple downscaling can be applied using a relationship to the surface topography. For example in surface temperature forecasts the lapse rate may be used to downscale the low resolution EPS field to a higher resolution grid using a gridded topography. The example below shows probabilities of strong wind downscaled from a regional EPS using a high-resolution orography field, and shows how probabilities of winds over the mountains in Scotland can be detected which were missed in the DMO version of the chart.



9.2.3 Site-specific extractions

Forecasts for specific locations may be generated by extracting data from model grids. In the simplest implementations data are simply taken from the nearest model grid-point, or are interpolated between the nearest grid-points by linear interpolation. Various methods are used to improve on these approaches, using similar techniques to the downscaling methods. In particular corrections to surface temperature and wind speed should be made to account for the difference between model orography and the true altitude of the site. An intelligent grid-point selection system which chooses the most representative grid-point can also be better than a simple interpolation, especially near coastlines where it may be better to choose the nearest land-point to represent a land-location, rather than for example the nearest grid-point which may be over the sea. This approach may also be beneficial near steep orography.

A one-dimensional model could also be used for specific forecast applications (e.g. 1D fog models for airports).

9.2.4 Statistical Downscaling

Downscaling of surface fields may also be done by building a statistical relationship between low-resolution model fields and high-resolution analyses. There are two approaches which may be followed:

9.2.4.1 Using Analysis Differences

The statistical relationship may be developed by comparing high-resolution gridded analyses with the corresponding analysis fields on the EPS model grid. This provides a downscaling vector which may then be applied to EPS forecast fields to provide bias-corrected and downscaled forecast fields on the high-resolution grid.

9.2.4.2 Kalman Filter

A Kalman filter approach may be applied at each grid-point of the high-resolution grid to build a statistical relationship with the lower-resolution EPS analysis fields. This Kalman filter may then be applied to the EPS forecast fields to provide bias-corrected and downscaled forecast fields on the high-resolution grid.

9.2.5 High Impact Weather Diagnostics

A number of methods are available to diagnose specific high-impact weather phenomena from NWP models, and these can be applied equally to EPS. A good example is Severe Convection diagnostics. These often use a number of model multi-level model outputs to diagnose the instability and potential for severe convection, and provide probabilities for phenomena such as large hail, tornadoes and convective wind gusts.

9.2.6 Downscaling by combination of low-resolution EPS and high-resolution control forecast

Low resolution ensemble perturbation fields (difference between the perturbed member forecast and the control forecast) can be added to a high resolution control forecast fields to provide a high-resolution probabilistic forecast.

9.3 Clustering techniques

Classification processes can be used to synthesize the huge amount of information contained in ensembles. Different kinds of classifications can be implemented :

- Clustering attempts to group together members which are most similar in their evolution over a defined geographical region of interest. Several standard clustering algorithms are available and may produce different groupings under. The clustering outcome also depends on the variables chosen.

- The “tubing” classification identifies a central cluster of the members closest to the ensemble mean and those members most significantly different from the ensemble mean (tube extremes). Tubing is useful to identify the most likely outcome and also the possible scenarios most different from that solution.
- Classification of forecasts by matching ensemble members to a defined set of flow regimes, for example the Grosswetterlagen types defined for central Europe. This method may provide the clustering which best matches a synoptic forecaster’s expectations.

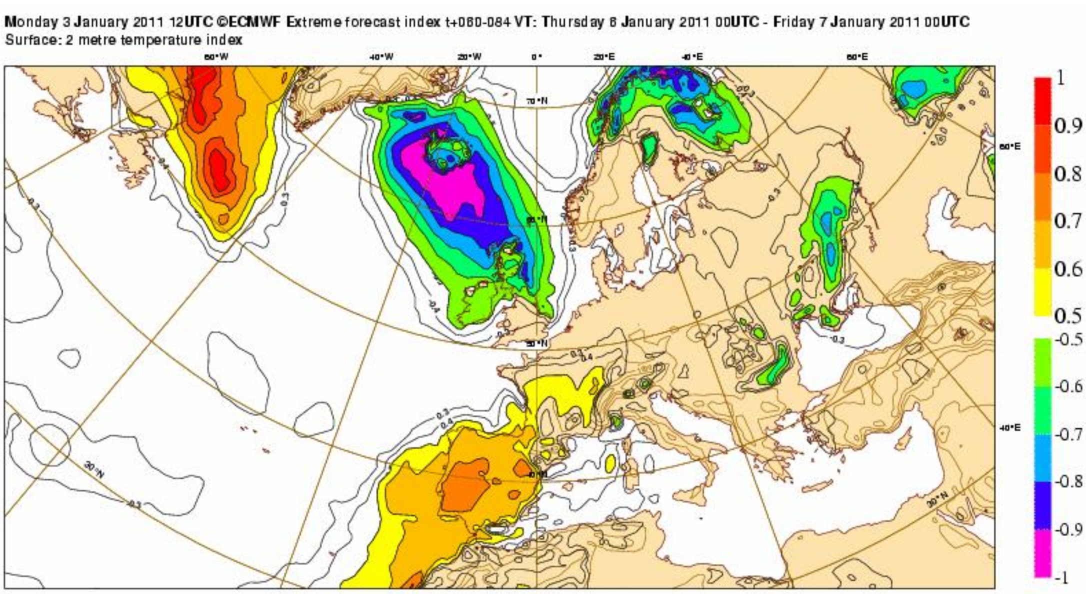
9.4 Use of Reforecasts

Research has shown that calibration of ensemble forecasts using historical sets of “reforecasts” – forecasts run with the same model or EPS from sets of historical cases, initiated from reanalyses – can be very effective in improving the quality and reliability of probabilistic forecasts. Such reforecasts provide a better dataset for training of statistical post-processing methods compared to using recent forecasts, as they provide a better sampling of different weather regimes and types. This can be particularly useful for optimizing the calibration of forecasts for rare or extreme events. However the running of reforecasts adds substantially to the computing cost of running an EPS, and depends also on the availability of a suitable reanalysis dataset to provide the initial conditions. As a result very few EPSs currently have reforecast datasets available, but their use is recommended where possible. Where a full reforecast dataset is not available, an alternative may be to use a recent archive of EPS forecasts from the same system, although is likely to provide a less reliable sampling of the full model climate.

9.4.1 Extreme Forecast Index (EFI)

One application of reforecasts is the computation of an Extreme Forecast Index.

NWP models and EPS systems do not represent accurately the climate of the real atmosphere, and identification of extreme events may be best done in relation to model climatology. The Extreme Forecast Index developed by ECMWF allows identification of forecasts which are extreme relative to the model climate, providing an alert to a risk of severe weather. The EFI does not provide explicit probabilities of severe events.



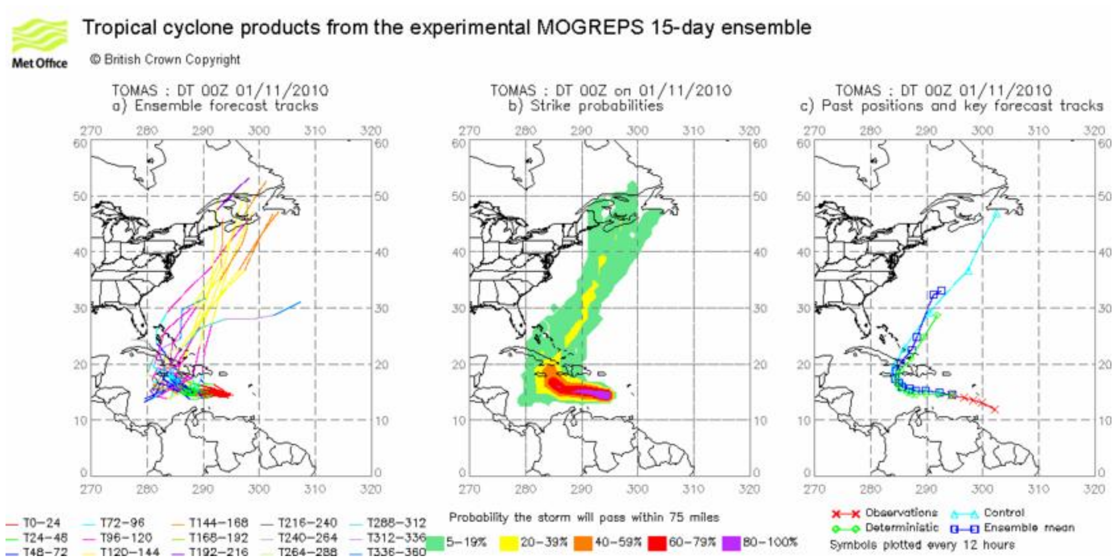
Reforecasts can also be used to assess forecast severity in relation to climatological return periods, which can be a useful way to communicate the severity of an event.

9.4.2 Quantile-Quantile Matching

Another approach to forecast calibration which can be used where an estimate of the model climate is available is Quantile Matching. For example the value corresponding to the 90th percentile of the model climate may be interpreted to represent the 90th percentile of the real observed climate distribution for a particular location. In general this method requires the use of a reforecast dataset to provide the model climate.

9.5 Feature Tracking

A useful technique for lower resolution EPS such as global EPS is to track meteorological features in each member of the ensemble. A good example is Tropical Cyclones (TC) which are not well-resolved in the model, but the global models can nevertheless predict the movement of the storms quite well. A global EPS could not be expected to predict the intensity of strong winds or heavy rain in a TC but could track its position. The forecaster can interpret the probabilities of severe weather by knowing the characteristics of tropical cyclones, combined with the ensemble information on where it is likely to go. The example below shows tracks of Hurricane Tomas in the members of the ensemble, probabilities that the storm will pass close to locations on the map, and summary tracks such as the ensemble mean track. These types of charts are often made available to the Tropical Cyclone RSMCs.



10. Use of EPS in Prediction of Severe Weather and issue of Warnings

Severe or high-impact weather events occur on a wide range of scales in space and time, from Tropical cyclone, extra-tropical cyclone, monsoon, winter storms and other large scale systems, to smaller scale systems such as local severe storms, orographic precipitation, thunderstorms and tornados. Forecasters must take account of the different predictabilities of different types of events (e.g. do not try to predict a thunderstorm 3 days in advance).

A well structured NMHS severe weather warning system should have appropriate thresholds, lead-times and level of service agreed with users. Thresholds should normally reflect the level of impact the weather is expected to have on society, including danger to life and property, and disruption to everyday life. Features which should be considered in a warning system include:

- Types of warnings; regions; thresholds (severity/impact and probability)
    - Risk = Probability x Impact
- A good warning system is one that will be easily understood by users, with standard thresholds adhered to by forecasters.

  - Many countries now use a 4-colour *traffic light* system (Green, Yellow, Amber and Red) indicating different levels of risk and corresponding levels of action which users should take.
- A good warning system will require feedback from users to NMHSs. The NMHSs in turn should give feedback to producers enabling them to design appropriate products.

EPS are a powerful tool in predicting severe weather events. For impact-based warnings systems the EPS may be used to help estimate the probability of weather hazards for use in the estimate of Risk = Probability x Impact. However, EPS can only predict severe weather which the model(s) can resolve:

- Numerical Weather Prediction has limitations in explicitly resolving smaller scale phenomena, which leads to under-estimation of extreme events likelihood within EPS.
- Sometimes can identify pre-cursor conditions for severe developments or favorable large scale environment such as convective indices
- Lower resolution EPS (Global) is less likely to be able to resolve details of an extreme event
- Regional EPS, which usually has higher resolution, should provide more detailed uncertainty estimates at the smaller scales.

Hazard thresholds used in the EPS may need to be calibrated to take account of the above limitations.

Early indications of some *extreme* events will be predicted in the tail of the ensemble distribution.

- Therefore forecasters and users should not ignore low probability events, especially when those events are very rare.
  - For example, ignoring probabilities below 20% or even 10% could result in missing the most important events signaled by the EPS.
  - To be able to use low probabilities, forecasters need verification information
  - “false alarms” are actually correct features of low probabilities. However low probabilities may be required in potential high-impact situations
  - It is expected that the probability will increase closer to the event – usually but not always

An extreme event may also be forecast essentially correctly, but with errors or uncertainties in location or timing.

Synoptic interpretation (e.g. weather feature tracking, use of analogues) or statistical downscaling tools are ways to add skill to the basic EPS.

- Note that some statistical methods require large data samples for training, and may not be well-suited to rare or extreme events.
- Cyclone tracking products (for both tropical and extra-tropical cyclones) can provide a useful summary of the development of high-impact storms.
- There is potential for development of more feature-based diagnostics for poorly resolved severe weather systems.

The Extreme Forecast Index (EFI) can be a useful tool in alerting forecasters to a potential severe event.

- EFI does not provide explicit probabilities of specific events, and should be interpreted in conjunction with other tools.
- Currently only a small number of systems can provide an EFI due to the need for a model climatology.

Consideration of input from multiple forecasting systems (EPS and deterministic) may give additional information on the probability of extreme events

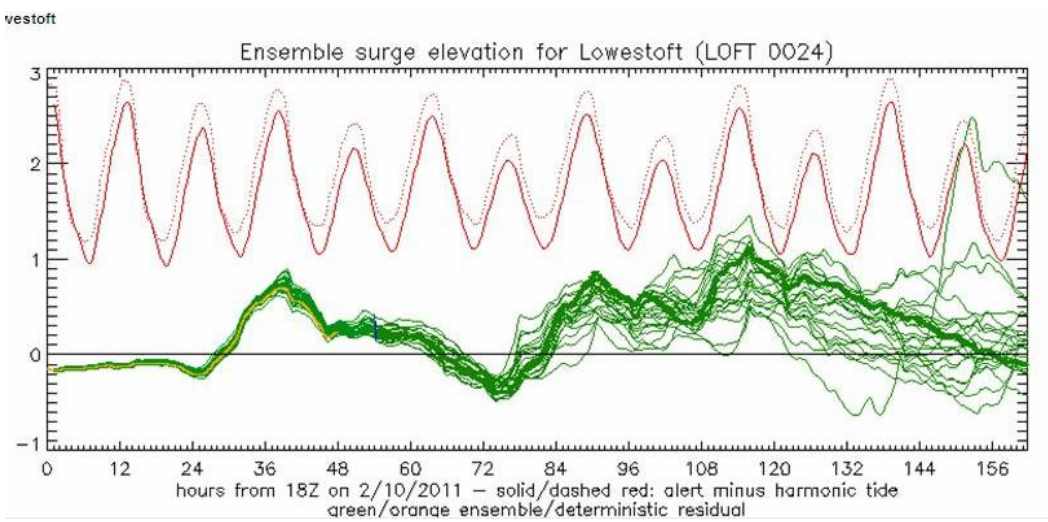
- Production of verification highlighting the skill and limitations of EPS is important.
  - Users of EPS should be aware of those limitations and strengths.

- However, due to the rarity of most extreme events it is often impossible to provide reliable (or statistically valid) verification of probabilistic performance. It may be possible to gain some estimate of skill for extreme events by extrapolating from the verification of less-severe events.

- Given the diminishing of the EPS skill with increasing lead time, latest available products are generally given higher credibility. However, previous runs of an EPS may still provide useful information about a rare event because of the lack of spread (limitation in the sample size).

11. Severe Weather Impact Modeling

The uncertainty in the weather forecast can be propagated through to uncertainty in impact by coupling ensemble members to impact models and generating a distribution of impact predictions. Examples include hydrological models for probabilistic flood forecasting, coastal storm surge models, heat health models etc. This is an advanced application which is being increasingly applied in the more advanced centres. The example below shows an ensemble forecast of storm surge at a coastal port, where the weather forecasting EPS has been used to force an ensemble with a storm surge model. The red lines at the top of the graph show the flood danger level oscillating up and down with the tide, and a flood risk is indicated where the ensemble forecast surge lines cross above the red lines. This is an interesting example as one member of the ensemble produces an extreme surge at day 7, indicating a low probability of severe coastal flooding. In this situation the user needs to be able to take some early preparedness action but without over-reacting because the probability of the flooding occurring is low:



12. Verification

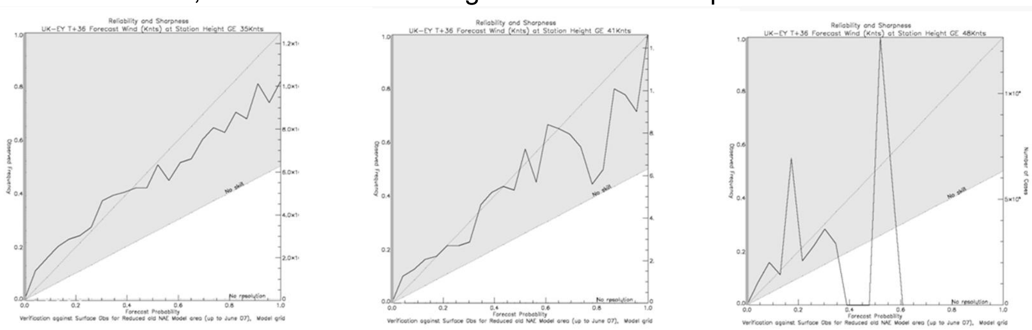
Verification is a very important part of everything we do in forecasting. If we do not verify our forecasts – measure how good they are by looking back afterwards and seeing how well the forecast matched what actually happened – then we have no way of learning and improving our forecasts in the future. This is just as true with probabilistic forecasts. You will often find people say that a probability forecast can never be wrong (unless we say 0% or 100%). Some people will also say that it is just a way for the forecaster to avoid making a decision. The way to challenge these views is to demonstrate that we do verify the forecasts, and that they have useful skill.

We do not provide here a detailed guide on verification of forecasts, but we describe a few important points:

- A *single* probability forecast cannot be right or wrong.
  - If we predict something with a high probability and it happens, it is often tempting to say "Look, we got it right!" We should avoid doing this, because when we forecast

something with a low probability and it happens we will want to say to the user "We did say it was a possibility even though it was a low probability".

- If we say there is a 30% probability that we will get more than 10mm of rain, and the observation shows that we get only 1mm, the forecast is not right or wrong. We have to measure the actual observed amount for many occasions when we make such a forecast – out of every 100 times that we say this, we should get over 10mm on 30 occasions. This is what the forecast means. Out of 100 times that we predict 80% probability, we should get it 80 times.
- The simplest way to present verification is using a reliability diagram, which plots a graph of the observed frequency against the forecast frequencies – so it plots exactly the test described above. Below are three examples of reliability diagrams for probabilities of wind speeds exceeding Beaufort Force 8, 9 and 10. The ideal is that the line should lie up the main diagonal, from (0,0) to (1,1). The first diagram on the left for Force 8 is quite good and shows that forecasts of high probability do mean the event is much more likely – the slope of the graph is slightly less than ideal, but good. The second is similar but not quite so good for the highest probabilities at the top right of the graph. The third on the right, for Force 10, shows useful skill for probabilities up to 30%, but at probabilities above that there is no useful information. In fact this is a rare event and there are not enough samples in the dataset to measure whether there is useful skill – this is a common problem with verifying extreme events, we do not have enough data to measure probabilistic skill.



- There are many other measure of probabilistic forecast. We list some other common ones here – much more information is easily available from an internet search for these terms, or from standard guides to forecast verification:
    - Brier Score – a root mean square error for probability forecasts of a particular event threshold.
    - Brier Skill Score – compares the Brier Score of the forecasts with the Brier Score of some reference forecast system.
    - Reliability – measures how well forecast probabilities match observed frequencies.
    - Resolution – measures how good the system is at predicting probabilities which are different from "normal".
    - ROC (Relative Operating Characteristic) – measure how good the forecasts are for decision-making – similar to resolution.
    - CRPS and RPS – (Continuous) Ranked Probability Score – like a Brier Score for multiple thresholds of the weather variable.

WMO CBS has defined a standard set of verification scores for comparison of EPSs, and these are displayed for a number of global EPSs at the Lead Centre website at http://epsv.kishou.go.jp/EPSv/ .

13. Forecaster Training

In general, forecaster training should include components on predictability and ensemble forecasting:

- Motivation for probabilistic forecasts – chaos theory and its impact.
- Statistical background theory and approaches.

- Aims of initial condition and model perturbations.
- Standard ensemble verification tools and their meaning.
- Explanation of basic meaning of products (e.g. lines on chart).
- Methods of post-processing and their impacts.

Learning Through Doing – The training of forecasters in the use of EPS guidance should be a practical experience using tools which are as close as possible to those used in operations. The optimal benefit from practical training on EPS is only obtained when an NMHS has access to operational EPS data, the operational time to use it and the products and tools to make direct use of it.

  - Benefits of training which is not reinforced by operational practice are rapidly lost.
  - Provision of training in conjunction with a demonstration project such as the SWFDP can help to ensure that the training is reinforced and consolidated by the provision of relevant operational EPS data.
  - During training, case studies should be worked through showing the appropriate use of EPS guidance, both in routine and severe weather scenarios.
  - Web based tools can be valuable in training, as they can be used on any workstation system through a standard browser to ensure continued access afterwards.
- In the relatively new area of EPS, periodic training is expected to generate the best benefit. Forecasters require time to build experience in using this guidance followed by further training to reinforce key concepts. It would also be of benefit if various NMHSs could share their experience with EPS.
- Training resources
  - ECMWF Users' guide http://www.ecmwf.int/products/forecasts/guide/
  - COMET ensemble modules: http://deved.meted.ucar.edu/nwp/pcu1/ensemble/

ANNEX V

DRAFT TEXT FOR THE NEW *MANUAL* ON GLOBAL ENSEMBLE PREDICTION, COORDINATION OF EPS VERIFICATION AND LIMITED AREA ENSEMBLE PREDICTION

II.2.1.3 – Global Ensemble Prediction

SPECIFICATION

Centres participating in activity 2.1.3, global Ensemble Prediction, shall:

- Prepare global ensemble forecast fields of basic and derived atmospheric parameters
- Make available on the WIS a range of these products. The minimum list to be made available, including parameters, forecast range, time steps and frequency, is given in Appendix A.II.2.1.3-a
- Prepare verification statistics according to the standard defined in Appendix A.II.2.3.2, and make them available to the Lead Centre(s) for EPS Verification
- Make available on a web site up-to-date information on the characteristics of its global Ensemble Prediction System. The minimum information to be provided is given in Appendix A.II.2.1.3-b

RESPONSIBILITY AND *(if required)* COORDINATION ***(Changes to Activity Specification)***			
To be proposed by:	CBS/ET-EPS	CBS/ICT-DPFS	
To be approved by:	CBS		
To be decided by:	EC / Congress		
DESIGNATION			
To be approved by:	CBS		
To be decided by:	EC / Congress		
COMPLIANCE			
To be monitored by:	CBS/ET-EPS		
To be reported to:	CBS/ICT-DPFS	CBS	

Appendix A.II.2.1.3-a

MINIMUM LIST OF PRODUCTS TO BE MADE AVAILABLE ON THE WIS

Parameter	Level	Thresholds	Resolution (lat/lon grid)	Forecast range	Time steps	Frequency
Probability of Precipitation	Surface	1, 5, 10, 25, 50 mm and 100 mm/24 hours	1.5º x 1.5º	10d (or the maximum range if less)	Every 12h	Once a day
Probability of 10 m sustained wind and gusts	Surface	10, 15 and 25 m s^{-1}				
Probability of Temperature anomalies	850	± 1, ± 1.5, ± 2 standard deviations with respect to a reanalysis climatology specified by the producing Centre				
Ensemble mean + spread (standard deviation) of Geopotential height	500					
Ensemble mean + spread (standard deviation) of MSL pressure	Surface					
Ensemble mean + spread (standard deviation) of wind speed	850/250					

Additional recommended products

- Location-specific time series of temperature, precipitation, wind speed, depicting the most likely solution and an estimation of uncertainty (“EPSgrams”). The definition, method of calculation and the locations should be documented.
- Tropical storm tracks (lat/long locations from EPS members).

Appendix II.2.1.3-b

CHARACTERISTICS OF GLOBAL EPS

1. Ensemble System
Ensemble name (version)
Date of implementation
2. Configuration of the EPS
Horizontal resolution of the model, with indication of grid spacing in km
Number of model levels
Top of model
Forecast length and forecast step interval
Runs per day (Times in UTC)
Is there an unperturbed control forecast included? (Y/N)
Number of perturbed ensemble members (excluding control)
Is model coupled to an ocean, waves, sea ice models? Specify which models
Integration time step
Additional comments
3. Initial conditions and Perturbations
Initial perturbation strategy
Optimisation time in forecast (if applicable)
Horizontal resolution of perturbations (if different from model resolution)
Initial perturbed area
Data assimilation method for control analysis
Are perturbations to observations employed? (Y/N)
Perturbations added to control analysis or derived directly from ensemble analysis
Perturbations in +/- pairs? (Y/N)
Additional comments
4. Model Uncertainty Perturbations
Is model physics perturbed? If yes, briefly describe method(s).
Do all ensemble members use exactly the same model version, or are, for example, different parameterization schemes used? Please describe any differences.
Is model dynamics perturbed? If yes, briefly describe method(s).
Are the above model uncertainty perturbations applied to the control forecast?
Additional comments
5. Surface Boundary Perturbations
Perturbations to sea-surface temperature? If yes, briefly describe method(s).
Perturbations to soil moisture? If yes, briefly describe method(s).
Perturbations to surface wind stress or roughness? If yes, briefly describe method(s).
Any other surface perturbations? If yes, briefly describe method(s).
Are the above surface perturbations applied to the control forecast?
Additional comments
6. Other details of model
What kind of soil scheme is in use?
How are radiations parametrized?

What kind of Large scale dynamics is in use (e.g. gridpoint semi-Lagrangian)? Hydrostatic?

What kind of boundary layer parametrization is in use?

What kind of convection parametrization is in use?

What Cloud scheme is in use?

Other relevant details?

7. Further Information

Operational contact point

URLs for system documentation

URL for list of products

II.2.3.2 – Coordination of EPS verification

SPECIFICATION

The centre(s) participating in activity 2.3.2, coordination of EPS verification, shall be designated as Lead Centre(s) for EPS verification.
These centre(s) shall:

- Provide the facility for the GDPFS Centres producing global EPS to automatically deposit their standardized verification statistics as defined in appendix A.II.2.3.2, and access to these verification statistics
- Maintain an archive of the verification statistics to allow the generation and display of trends in performance
- Monitor the received verification statistics and consult with the relevant participating centre if data is missing or suspect
- Provide access to standard data sets needed to perform the standard verification, including climatology and lists of specified observation sites and keep this up to date according to CBS recommendation
- Provide on its (their) website(s):
    - consistent up-to-date graphical displays of the verification results from participating Centres through processing of the received statistics
    - relevant documentation including access to the standard procedures required to perform the verification, and links to the websites of GDPFS participating Centres
    - contact details to encourage feedback from NMHSs and other GDPFS Centres on the usefulness of the verification information

These centre(s) may also provide access to standardized software for calculating scoring information.

RESPONSIBILITY AND *(if required)* COORDINATION ***(Changes to Activity Specification)***			
To be proposed by:	CBS/ET-EPS	CBS/ICT-DPFS	
To be approved by:	CBS		
To be decided by:	EC / Congress		
DESIGNATION			
To be approved by:	CBS		
To be decided by:	EC / Congress		
COMPLIANCE			
To be monitored by:	CBS/ET-EPS		
To be reported to:	CBS/ICT-DPFS	CBS	

Appendix A.II.2.3.2

STANDARD VERIFICATION MEASURES OF EPS

EXCHANGE OF SCORES

Monthly exchanges:

Ensemble mean
For verification of ensemble mean, the specifications in Appendix A.II.2.3.1 for variables, levels, areas and verifications shall be used.

Spread
Standard deviation of the ensemble averaged over the same regions and variables as used for the ensemble mean.

Probabilities
Probabilistic scores (excluding the CRPS) are exchanged in the form of reliability tables. Details of the format of the exchange of verification data are provided on the website of the Lead Centre for EPS verification.

List of parameters

PMSL anomaly ± 1, ± 1.5, ± 2 standard deviation with respect to a centre-specified climatology
Verified for areas defined for verification against analysis

Z500 with thresholds as for PMSL. Verified for areas defined for verification against analysis

850 hPa wind speed with thresholds of 10, 15, 25 m s^{-1}. Verified for areas defined for verification against analysis

850 hPa u and v wind components with thresholds of 10th, 25th, 75th and 90th percentile points with respect to a centre-specified climatology. Verified for areas defined for verification against analysis

250 hPa u and v wind components with thresholds of 10th, 25th, 75th and 90th percentile points with respect to a centre-specified climatology. Verified for areas defined for verification against analysis

T850 anomalies with thresholds ± 1, ± 1.5, ± 2 standard deviation with respect to a centre-specified climatology. Verified for areas defined for verification against analysis

Precipitation with thresholds 1, 5, 10, and 25 mm/24 hours every 24 hours verified over areas defined for deterministic forecast verification against observations

Observations for EPS verification should be based on the GCOS list of surface network (GSN). Producing centres shall have the right to omit certain observation sites should they fail a quality control. Verification of precipitation may alternatively be against a proxy analysis i.e. short range forecast from the control or high-resolution deterministic forecast, e.g. 12-36h forecast to avoid spin-up problems.

NOTE: Where thresholds are defined with respect to climatology, the daily climate should be estimated.

Scores (computed by Lead Centre(s) based on reliability tables provided by participating centres)

Brier Skill Score (with respect to climatology) (see definition below*)

Relative Operating Characteristic (ROC)
Relative economic value (C/L) diagrams
Reliability diagrams with frequency distribution
Continuous Rank Probability Score (CRPS)

NOTES:

In the case of CRPS, centres are encouraged to submit this for both EPS and the deterministic (control and high-resolution) forecast as well - CRPS for deterministic forecast is equal to the mean absolute error.

---

* The Brier Score (BS) is most commonly used for assessing the accuracy of binary (two-category) probability forecasts. The Brier Score is defined as: *[Move to Part II, 1.4]*

$$BS = \frac{\sum_{ij}\left(F_{ij} - O_{ij}\right)^2}{N}$$

where the observations *Oij* are binary (0 or 1) and *N* is the verification sample size. The Brier Score has a range from 0 to 1 and is negatively-oriented. Lower scores represent higher accuracy.

The Brier Skill Score (BSS) is in the usual skill score format, and may be defined by:

$$BSS = \frac{BS_C - BS_F}{BS_C} \times 100 = \left[1 - \frac{\sum_{ij}\left(F_{ij} - O_{ij}\right)^2}{\sum_{ij}\left(C_{ij} - O_{ij}\right)^2}\right] \times 100$$

where *C* refers to climatology and *F* refers to the forecast.

II.2.1.4 – Limited Area Ensemble Prediction

SPECIFICATION

Centres participating in activity 2.1.4, Limited Area Ensemble Prediction, shall:

- Prepare limited area ensemble forecast fields of basic and derived atmospheric parameters
- Make available on the WIS a range of these products. The minimum list to be made available, including parameters, forecast range, time steps and frequency, is given in Appendix A.II.2.1.4-a
- Prepare verification statistics according to the standard defined in Appendix A.II.2.3.2, adapted for the region covered by the model at a resolution of 0.5°, and make available consistent up-to-date graphical displays of the verification results on a web site
- Make available on a web site up-to-date information on the characteristics of its limited area Ensemble Prediction System. The minimum information to be provided is given in Appendix A.II.2.1.4-b

RESPONSIBILITY AND *(if required)* COORDINATION ***(Changes to Activity Specification)***			
To be proposed by:	CBS/ET-EPS	CBS/ICT-DPFS	
To be approved by:	CBS		
To be decided by:	EC / Congress		
DESIGNATION			
To be approved by:	CBS		
To be decided by:	EC / Congress		
COMPLIANCE			
To be monitored by:	CBS/ET-EPS		
To be reported to:	CBS/ICT-DPFS	CBS	

Appendix A.II.2.1.4-a

MINIMUM LIST OF PRODUCTS TO BE MADE AVAILABLE ON THE WIS

Parameter	Level	Thresholds	Resolution (lat/lon grid)	Forecast range	Time steps	Frequency
Probability of Precipitation	Surface	1, 5, 10, 25, 50 mm and 100 mm/24 hours	0.5º x 0.5º	2d (or the maximum range if less)	Every 6h	Once a day
Probability of 10 m sustained wind and gusts	Surface	10, 15 and 25 m s^{-1}				
Probability of Temperature anomalies	850	± 1, ± 1.5, ± 2 standard deviations with respect to a reanalysis climatology specified by the producing Centre				
Ensemble mean + spread (standard deviation) of Geopotential height	500					
Ensemble mean + spread (standard deviation) of MSL pressure	Surface					
Ensemble mean + spread (standard deviation) of wind speed	850/250					

Additional recommended products

- Location-specific time series of temperature, precipitation, wind speed, depicting the most likely solution and an estimation of uncertainty ("EPSgrams"). The definition, method of calculation and the locations should be documented.
- Tropical storm tracks (lat/long locations from EPS members).

Appendix II.2.1.4-b

CHARACTERISTICS OF LIMITED AREA EPS

1. Ensemble System
Ensemble name (version
Date of implementation
2. Configuration of the EPS
Horizontal resolution of the model, with indication of grid spacing in km
Number of model levels
Top of model
Forecast length and forecast step interval
Runs per day (Times in UTC)
Is there an unperturbed control forecast included? (Y/N)
Number of perturbed ensemble members (excluding control)
Is model coupled to an ocean, waves, sea ice models? Specify which models
Integration time step
Additional comments
3. Initial conditions and Perturbations
Initial perturbation strategy
Optimisation time in forecast (if applicable)
Horizontal resolution of perturbations (if different from model resolution)
Initial perturbed area
Data assimilation method for control analysis
Are perturbations to observations employed? (Y/N)
Perturbations added to control analysis or derived directly from ensemble analysis
Perturbations in +/- pairs? (Y/N)
Additional comments
4. Model Uncertainty Perturbations
Is model physics perturbed? If yes, briefly describe method(s).
Do all ensemble members use exactly the same model version, or are, for example, different parameterization schemes used? Please describe any differences.
Is model dynamics perturbed? If yes, briefly describe method(s).
Are the above model uncertainty perturbations applied to the control forecast?
Additional comments
5. Surface Boundary Perturbations
Perturbations to sea-surface temperature? If yes, briefly describe method(s).
Perturbations to soil moisture? If yes, briefly describe method(s).
Perturbations to surface wind stress or roughness? If yes, briefly describe method(s).
Any other surface perturbations? If yes, briefly describe method(s).
Are the above surface perturbations applied to the control forecast?
Additional comments
6. Other details of model
What kind of soil scheme is in use?
How are radiations parametrized?

What kind of Large scale dynamics is in use (e.g. gridpoint semi-Lagrangian)? Hydrostatic?

What kind of boundary layer parametrization is in use?

What kind of convection parametrization is in use?

What Cloud scheme is in use?

Other relevant details?

7. Regional Ensemble specifics

Regional domain descriptor (lat/long of boundaries)

Normal source of boundary conditions

Are boundary conditions perturbed?

Specification of boundary conditions required.

Are boundary condition requirements compatible with any other global models or standards? If so, please describe

Additional comments

8. Further Information

Operational contact point

URLs for system documentation

URL for list of products

ANNEX VI

TERMS OF REFERENCE OF THE ET-EPS

Expert Team on Ensemble Prediction Systems

(a) Provide advice on EPS in relation to probabilistic forecasts in the context of short- and medium-range EPS products, focusing on applications concerned with all aspects of the EPS systems which forecast the weather on a daily basis;

(b) Propose guidance for the generation of EPS products for the communication of uncertainty in prediction of the weather and its impact;

(c) Liaise with the PWS programme to promote and support the use and communication of probabilistic information available from the GDPFS centres;

(d) Promote and support the education and training of forecasters, including rationale of concepts and strategies of EPS, and on the nature, interpretation and application of EPS products;

(e) In consultation with the Coordination Group on verification, review verification system for EPS products and provide guidance on the interpretation of verification;

(f) Support the further development of the Lead Centre on Verification of EPS by reporting on verification measures and determining the best way of presenting skill of ensemble forecasting systems;

(g) To review the Manual on the GDPFS (WMO-No. 485) and propose updates as necessary concerning EPS;

(h) Develop specifications for the introduction of probabilistic information into products from RSMCs and Regional Centres of the SWFDP;

(i) Liaise with THORPEX GIFS-TIGGE WG to trial new products and advise on their suitability for operational application.

ANNEX VII

LIST OF ABBREVIATIONS AND ACCRONYMS

ATM	Atmospheric Transport Modelling
BV	Breeding Vector
CAT	Clear Air Turbulence
CBS	Commission for Basic Systems (WMO)
Cg	World Meteorological Congress (WMO)
CGMS	Coordination Group of Meteorological Satellite Operators
CMA	China Meteorological Administration
CMC	Canadian Meteorological Centre (MSC)
COSMO	Consortium for Small-scale Modelling
ECMWF	European Centre for Medium-range Weather Forecasts
EDA	Ensemble Data Assimilation
EFAS	European Flood Forecasting System
EFI	Extreme Forecast Index
EnKF	Ensemble Kalman Filter
EPS	Ensemble Prediction Systems
ET	Expert Team
ET-EPS	Expert Team on Ensemble Prediction Systems (CBS)
ETKF	Ensemble Transform Kalman Filter
ETR	Ensemble Transform and Rescaling
EuroPP	European Post-Processing
GCI	GEOSS Common Infrastructure
GDPFS	Global Data-Processing and Forecasting System (WMO)
GEM	Global Environmental Multi-scale
GEO	Global Earth Observation
GEOSS	Global Earth Observation System of Systems
GEOWOS	GEOSS interoperability for Weather, Ocean and Water
GEPS	Global Ensemble Prediction Systems
GFS	Global Forecast System
GIFS	Global Interactive Forecast System (TIGGE)
GlobalPP	Global Post-Processing
GSM	Global Spectral Model (JMA)
IP	Initial Perturbation
JMA	Japan Meteorological Agency
KF	Kalman Filter
KMA	Korea Meteorological Administration
LAM	Limited Area Model
LDC	Least Developed Country
MAP	Mesoscale Alpine Programme
MRI	Meteorological Research Institute (JMA)
MSC	Meteorological Service of Canada
NAEFS	North American Ensemble Forecast System
NCEP	National Centres for Environmental Prediction (US)
NDGD	National Digital Guidance Database
NMHS	National Meteorological and Hydrological Services
NMSM	National Meteorological Service of Mexico
NOAA	National Oceanic and Atmospheric Administration
NOMADS	National Operational Model Archive and Distribution System (NOAA)
NSWWS	National Severe Weather Warning Service (Met Office UK)
NWP	Numerical Weather Prediction
NWS	National Weather Service (NOAA)
REPS	Regional Ensemble Prediction Systems
RP	Random Parameters
RSMC	Regional Specialized Meteorological Centre (WMO)

RTMA	Real Time Mesoscale Analysis
PDF	Probability Density Function
PEARP	Prévision d'Ensemble ARPège (Météo-France)
PPT	Perturbations of Physics Tendencies
PWS	Public Weather Services
RA	Regional Association (WMO)
SG-SWFDP	Steering Group of the SWFDP
SKEB	Stochastic Kinetic Energy Backscatter
SSPS	Site-Specific Processing System
STTP	Stochastic Total Tendency Perturbations
SV	Single Vector
SWFDP	Severe Weather Forecasting Demonstration Project
TC	Tropical Cyclone
TD	Technical Document (WMO)
THORPEX	The Observing System Research and Predictability Experiment
TIGGE	THORPEX Interactive Grand Global Ensemble
ToR	Terms of Reference
UK	United Kingdom
UKMO	Met Office UK
UM	Unified Model (Met Office UK)
UKPP	UK Post-Processing
US	United States of America
WDS	Weather and Disaster Risk Reduction Services Department (WMO)
WMO	World Meteorological Organization
WWRP	World Weather Research Programme (WMO)