Ensemble Forecasting

David Richardson

Head of Evaluation, Forecast Department

david.richardson@ecmwf.int

ECMWF

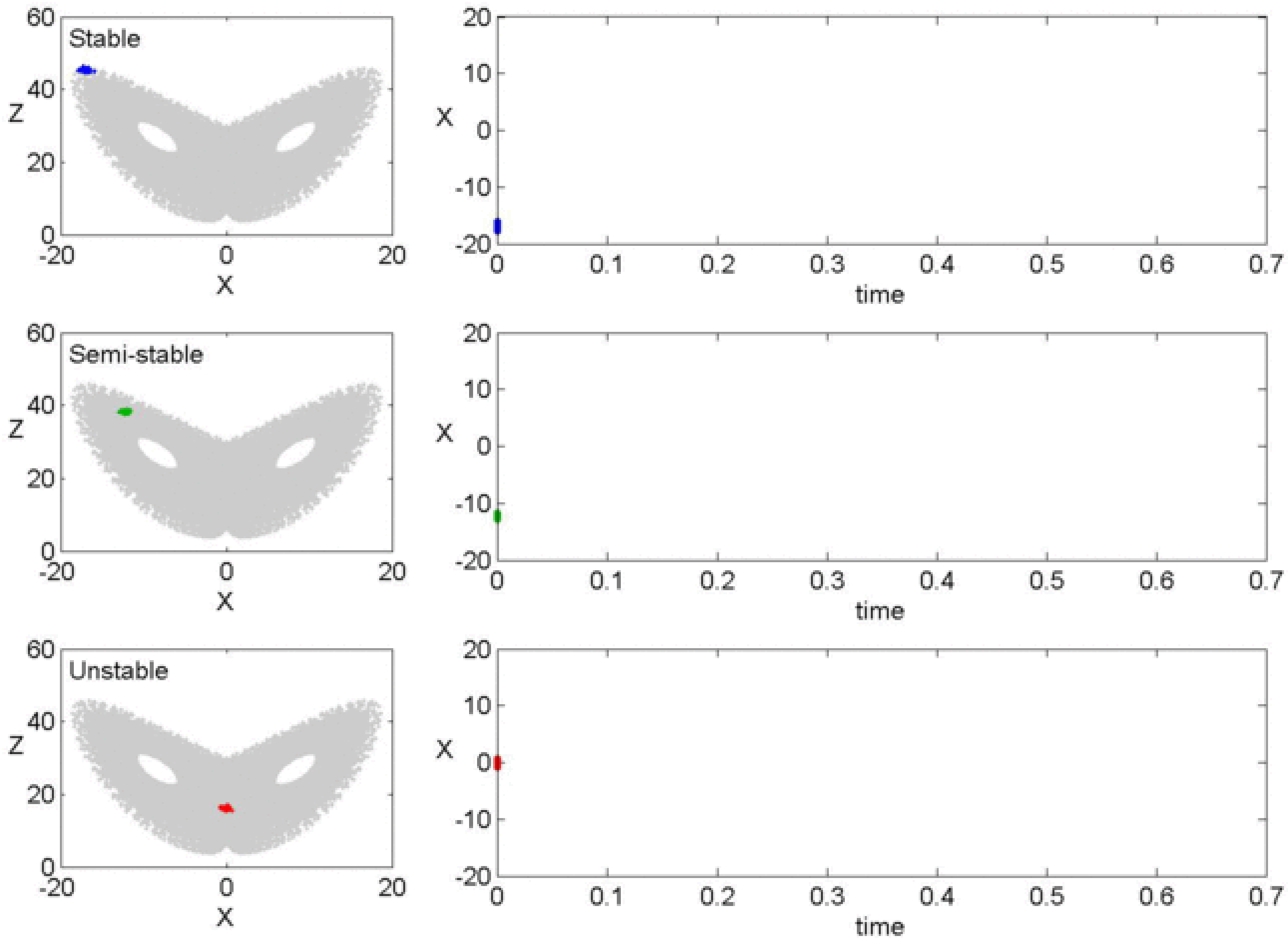
Outline

- Introduction
  - Why do forecast go wrong?
  - Observations, model, “chaos”
- The ECMWF ensemble
  - How does the ENS represent uncertainties?
  - Configuration of the ENS
- ENS products
  - Very short overview – much more in rest of course
- Use of ENS
  - Probabilities and decision support

Why are forecasts sometimes wrong?

- **Initial condition uncertainties**
  - Lack of observations
  - Observation error
  - Errors in the data assimilation
- **Model uncertainties**
  - Limited resolution
  - Parameterisation of physical processes
- **The atmosphere is chaotic**
  - small uncertainties grow to large errors (unstable flow)
  - small scale errors will affect the large scale (non-linear dynamics)
  - error-growth is flow dependant
- **Even very good analyses and forecast models are prone to errors**

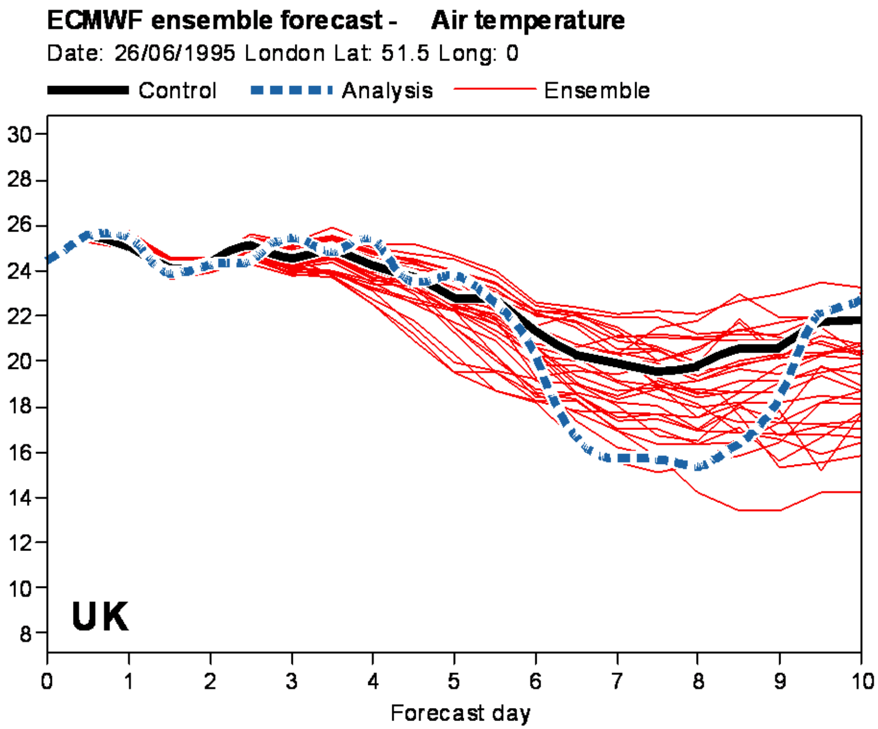
Chaos - the Lorenz attractor



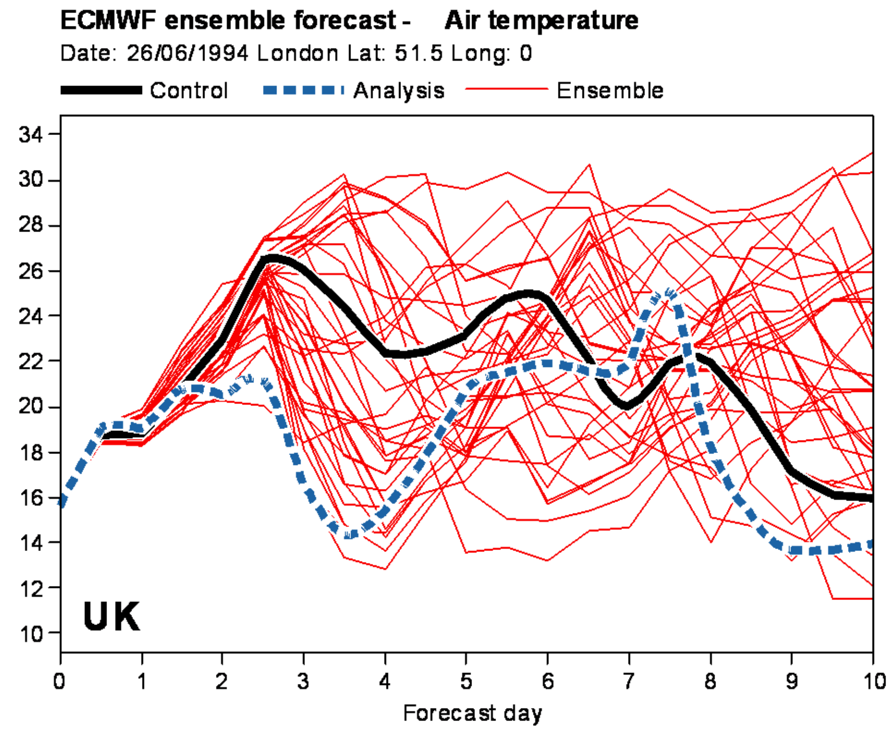
Tim Palmer, Oxford University

Flow dependence of forecast errors

26th June 1995



26th June 1994

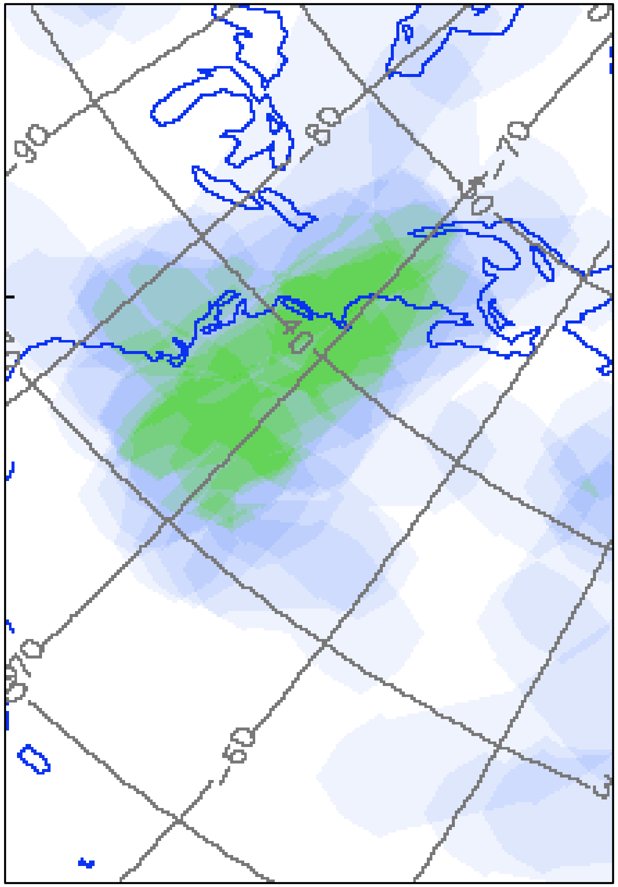


If the forecasts are coherent (small spread) the atmosphere is in a more predictable state than if the forecasts diverge (large spread)

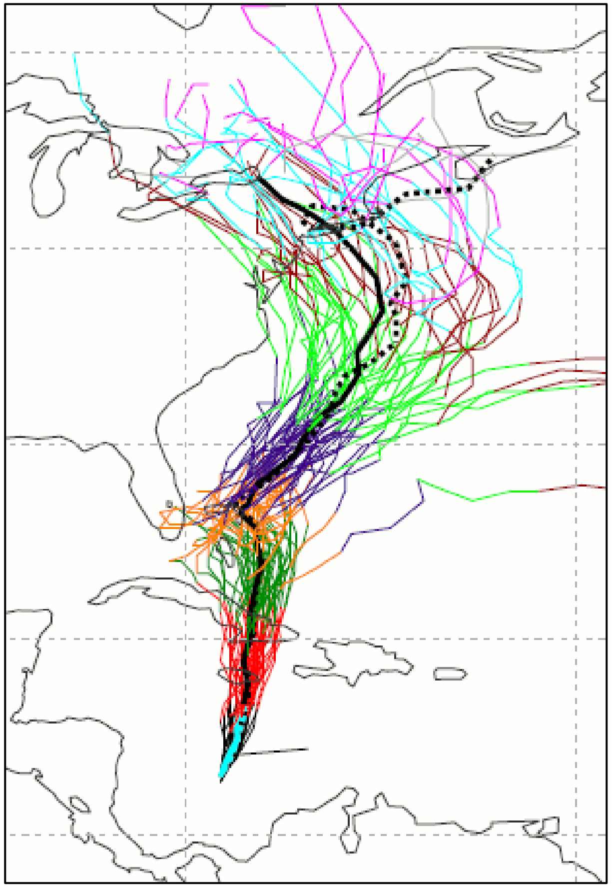
ECMWF

Superstorm Sandy

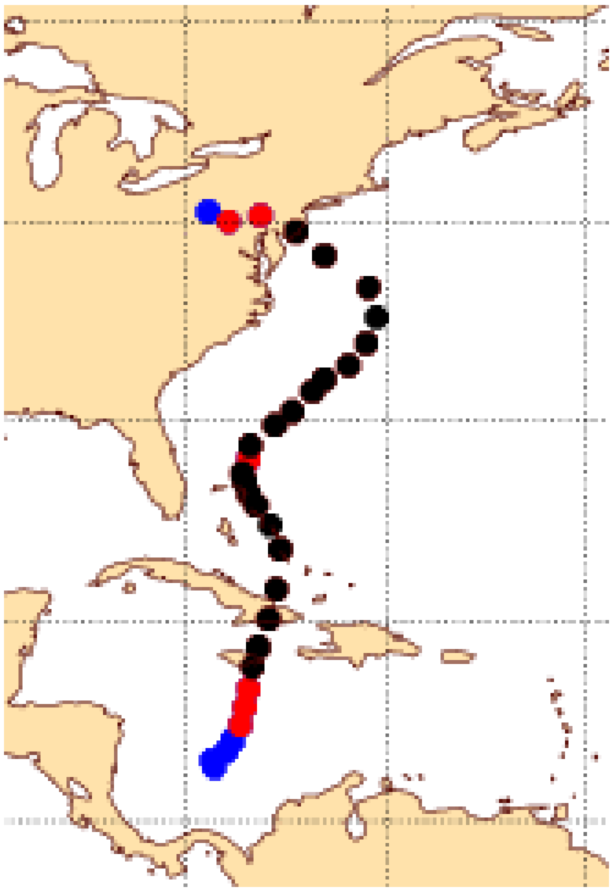
First indications 9.5 days before landfall



Track forecasts 6.5 days before landfall



Observed track of Sandy



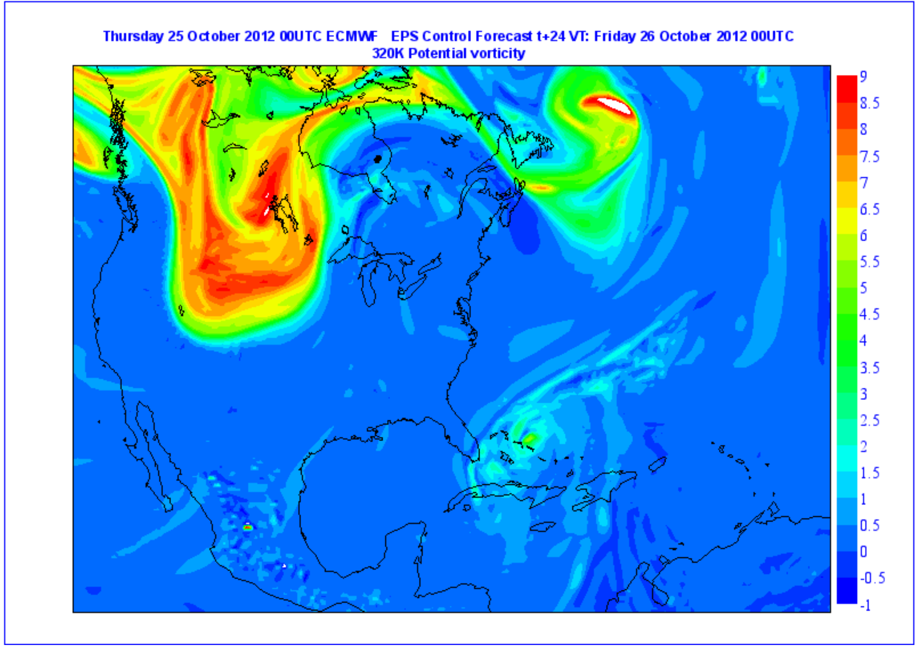
2 days before Sandy formed (9.5 days before landfall in New Jersey) there was already a significant probability (25%) of a severe wind storm affecting NE USA

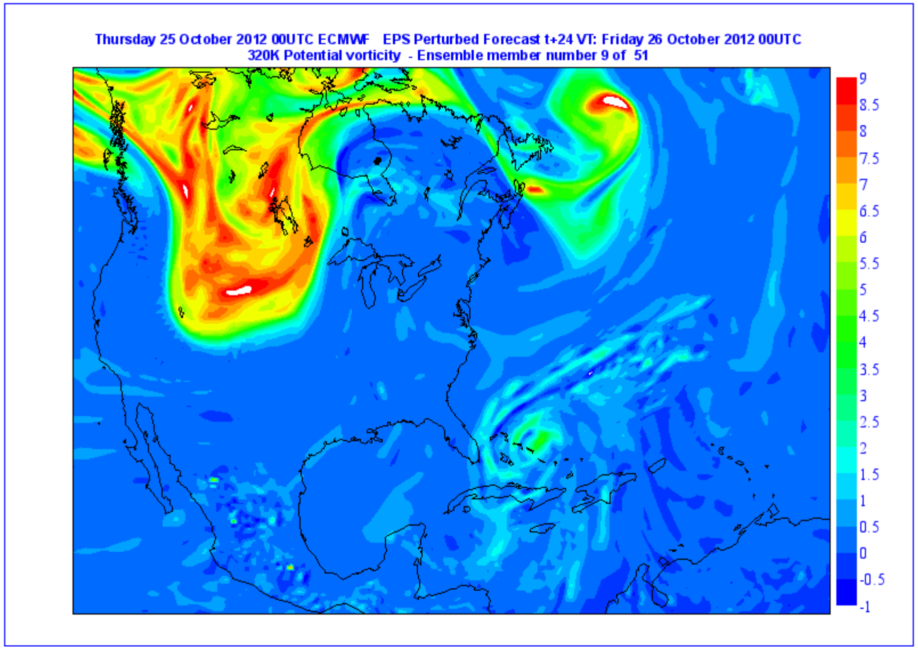
Sandy: ENS PV evolution

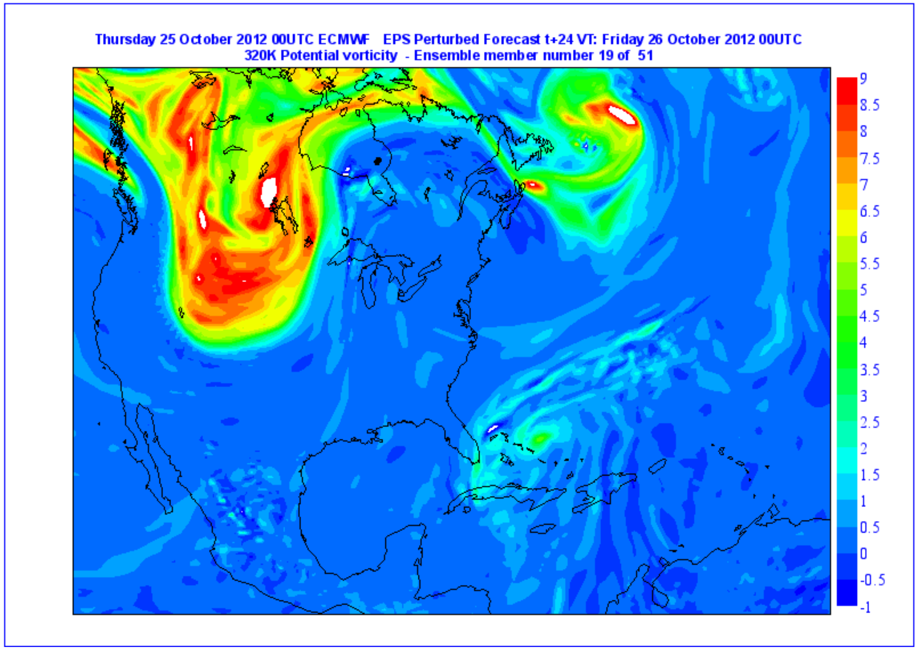
Forecast from 0 UTC on 25 October

three ensemble members:
- control (top)
- M09 (bottom L) “caught” too late
- M19 (bottom R) “escaped”

PV on 320K (6h steps)







What is an ensemble?

- A set of forecasts run from slightly different initial conditions to account for initial uncertainties
  - At ECMWF perturbations are generated using singular vectors and an ensemble of data assimilations
- The forecast model also contains approximations that can affect the forecast evolution
  - Model uncertainties are represented using “stochastic physics”
- The ensemble of forecasts provides a range of future scenarios consistent with our knowledge of the initial state and model capability
  - Provides explicit indication of uncertainty in today’s forecast

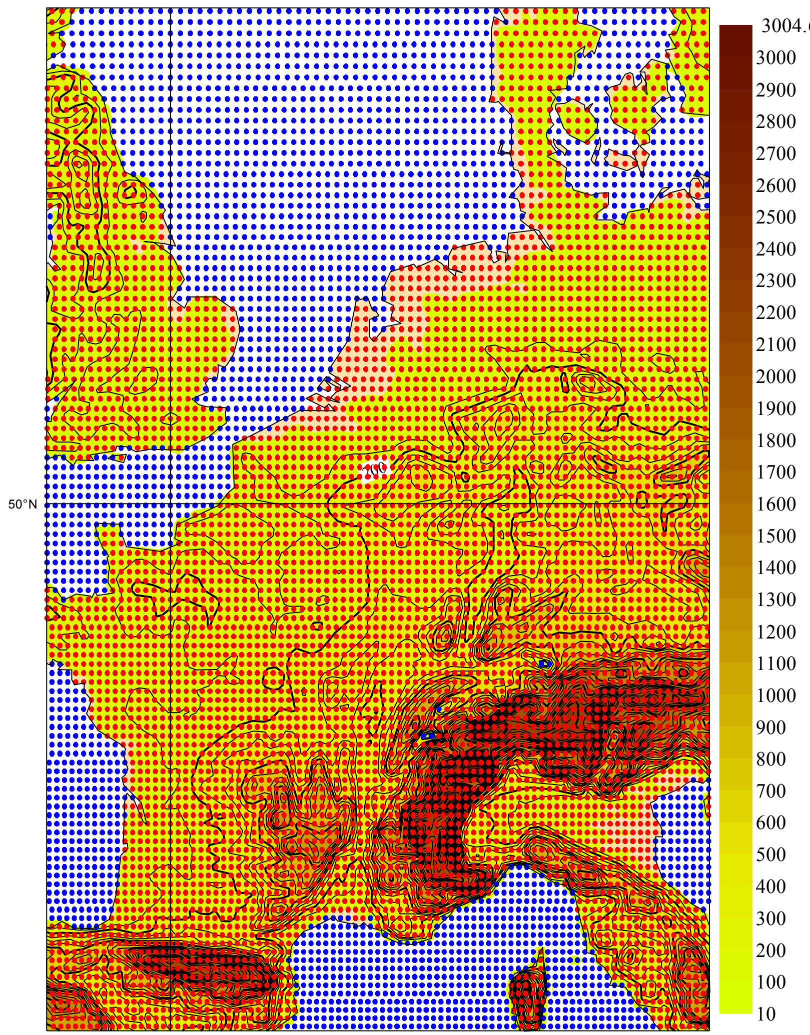
ECMWF medium-range forecasts

- High-resolution forecast (16 km grid, 137 levels) runs twice every day to 10 days
- Ensemble: same model but run at lower resolution (32 km, 91 levels; 64 km after day 10)
  - ensemble control (run from high-resolution analysis, no perturbation)
  - 50 perturbed members (account for initial and model uncertainties)
  - Ensemble coupled to ocean model from start of forecast

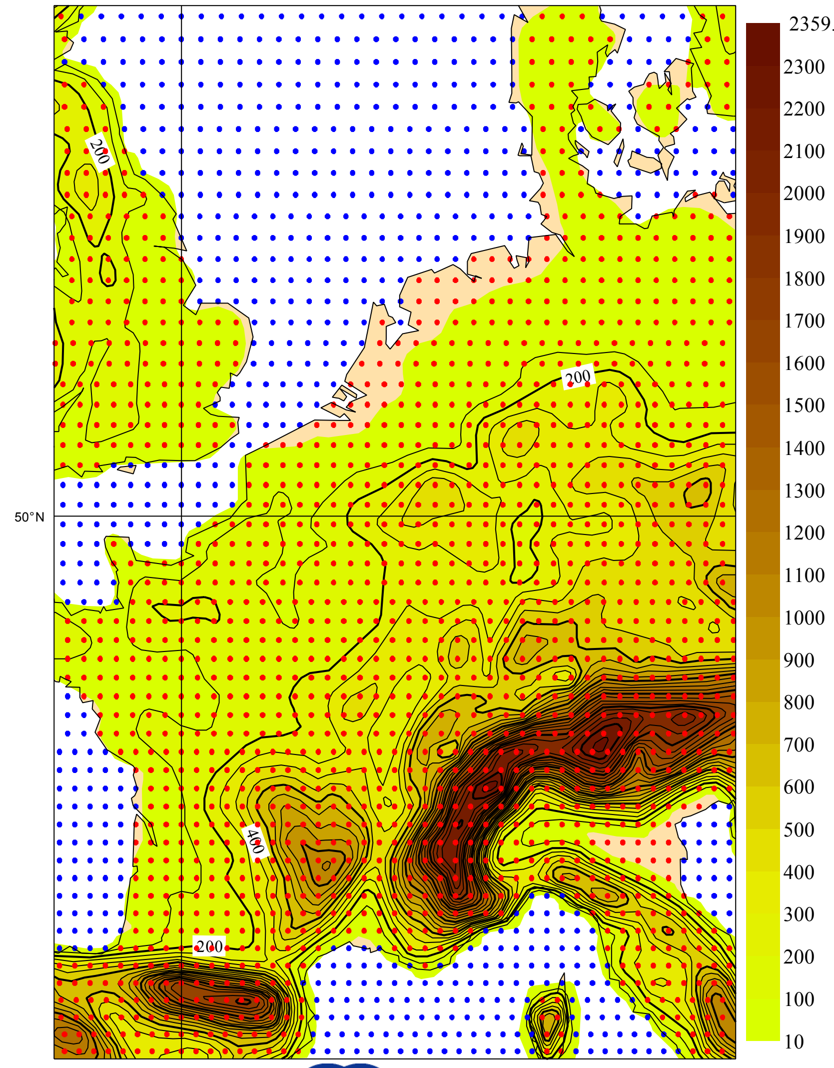
Model grids:

HRES (16km, T1279) ENS (32 km, T639)

OROGRAPHY, GRID POINTS AND LAND SEA MASK IN TL 1279 (OP 2010) ECMWF MODEL
orography shaded (height in m), land grid points (red), sea grid points (blue)

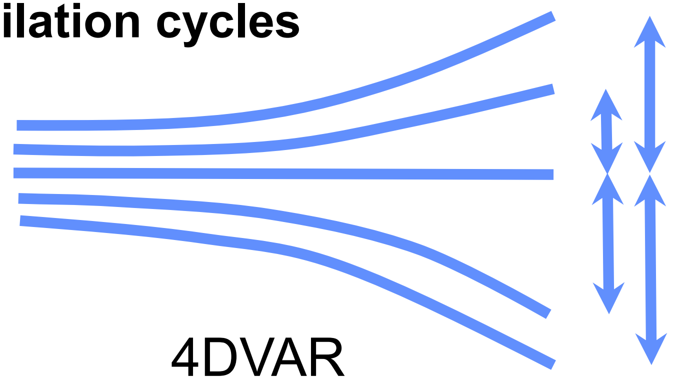


OROGRAPHY, GRID POINTS AND LAND SEA MASK IN TL 639 (EPS 2010) ECMWF MODEL
orography shaded (height in m), land grid points (red), sea grid points (blue)



Initial uncertainties

- Combination of 2 types of perturbations
- Ensemble of data assimilations (EDA)
  - Randomly perturbed observations and SST fields
  - Run 25 independent data assimilation cycles



- Singular vectors: perturbations that grow quickly over the first 48 hours of the forecast

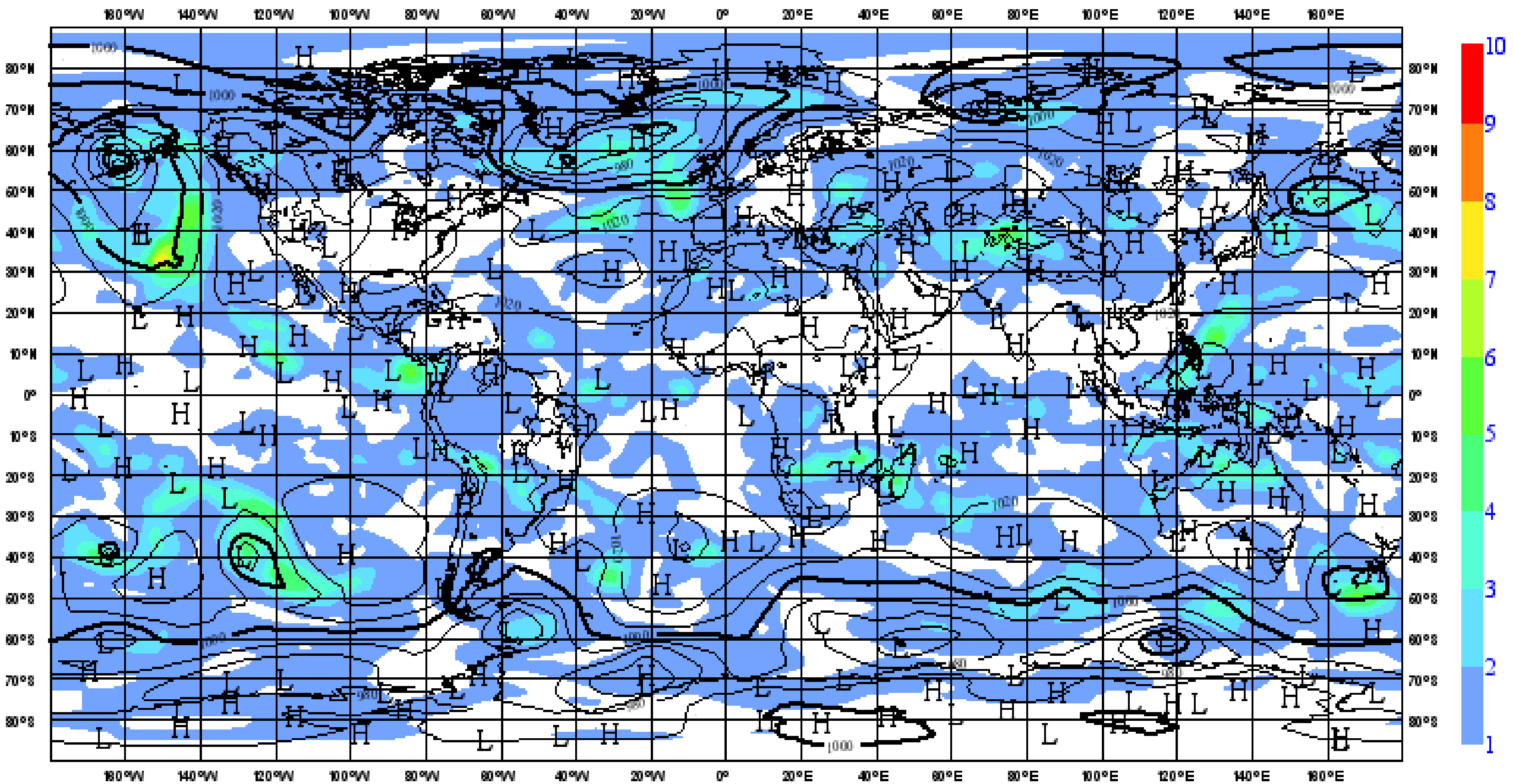
- Best approach given limited available computer resources

ENS initial perturbations

- SV- and EDA-based perturbations have different characteristics:
  - EDA-based perturbations are less localized than SV-based perturbations and have a smaller scale. They have a larger amplitude over the tropics. EDA-perturbations are more barotropic than SV-based perturbations, and grow less rapidly.
  - At initial time, SV-based perturbations have a larger amplitude in potential than kinetic energy, while EDA-based perturbations have a similar amplitude in potential and kinetic energy
- Since June 2010 SV- and EDA-based perturbations are used together to construct the initial perturbations for the EPS
- The perturbations are constructed so that all perturbed members are equally likely
- All perturbations are flow-dependent: they are different from day to day

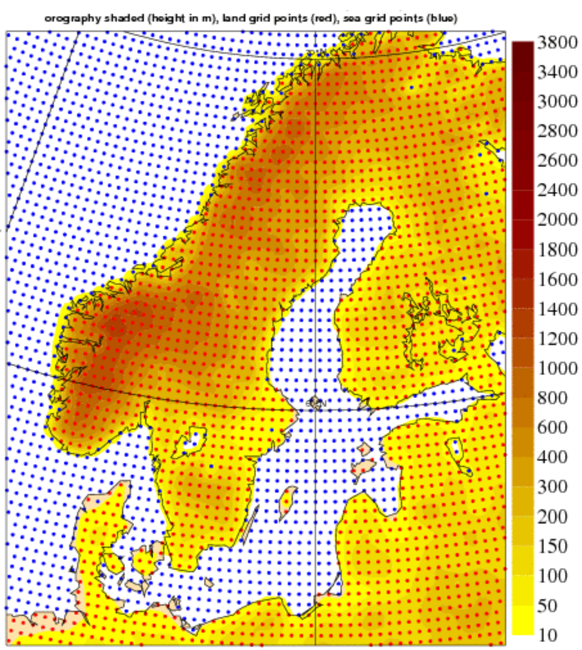
Ensembles of Data Assimilation (EDA)

The ensemble spread is flow-dependent but noisy. A filter is applied to remove it. This plot shows the EDA std in terms of vorticity at 500 hPa, +9h after filtering.

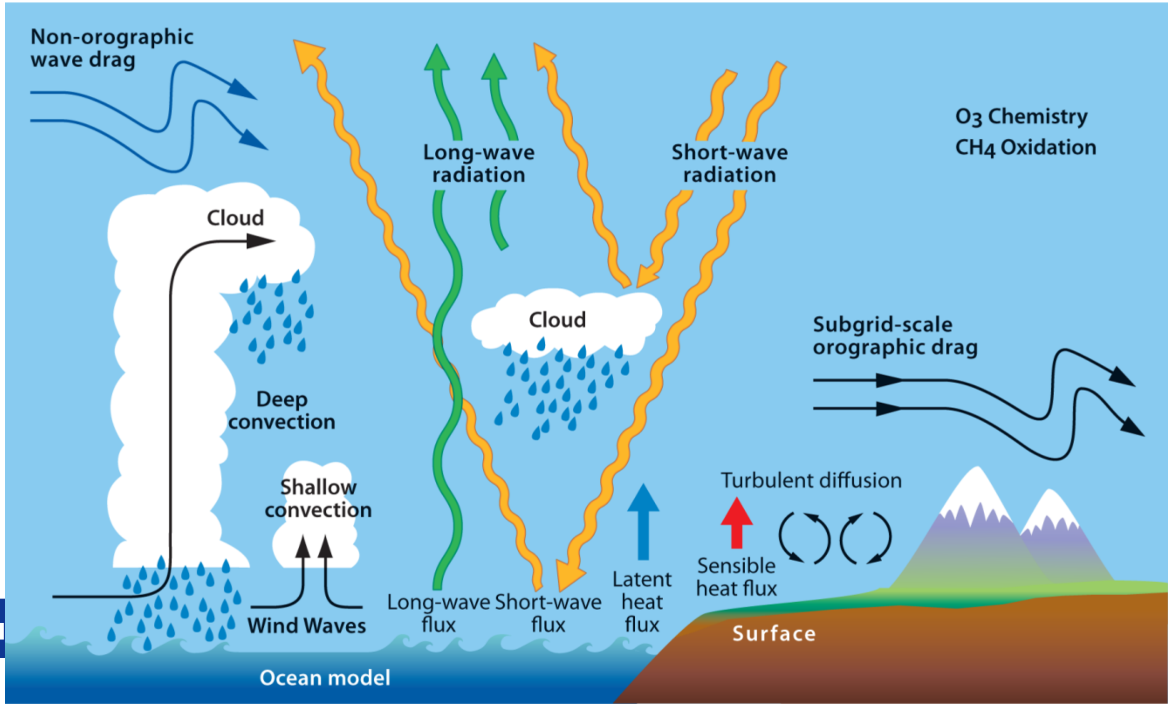


Model uncertainties – stochastic physics

- Parametrization – represent effects of unresolved (or partly resolved) processes on the resolved model state
- Statistical ensemble of sub-grid scale processes within a grid box; in equilibrium with grid-box mean flow
- Stochastic physics represents statistical uncertainty
  - allows for energy transfer from sub-grid scale to resolved flow, non-local effects

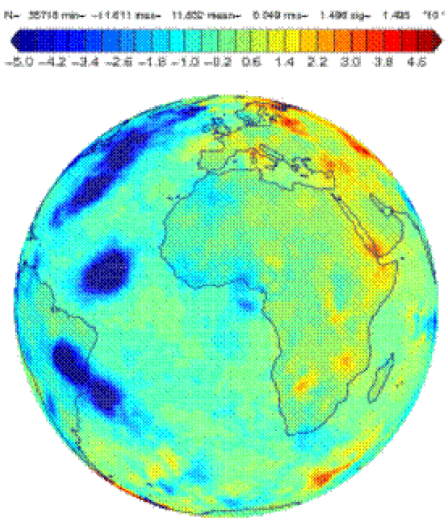


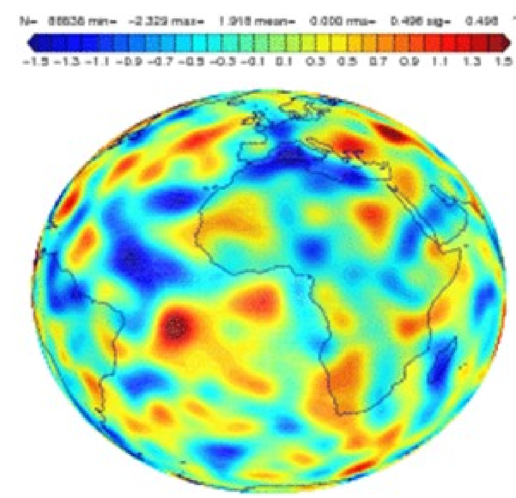
WF Produ



Model uncertainties – stochastic physics

- **2 components**
- **Stochastically Perturbed Parametrization Tendencies (SPPT)**
  - **Random pattern of perturbation to model fields**
  - **Initial scheme introduced 1999, revised 2009 (cycle 35r3)**
- **Spectral stochastic backscatter scheme (SPBS)**
  - **A fraction of the dissipated energy is backscattered upscale and acts as streamfunction forcing for the resolved-scale flow**
  - **Introduced in addition to SPPT in November 2010 (cycle 36r4)**

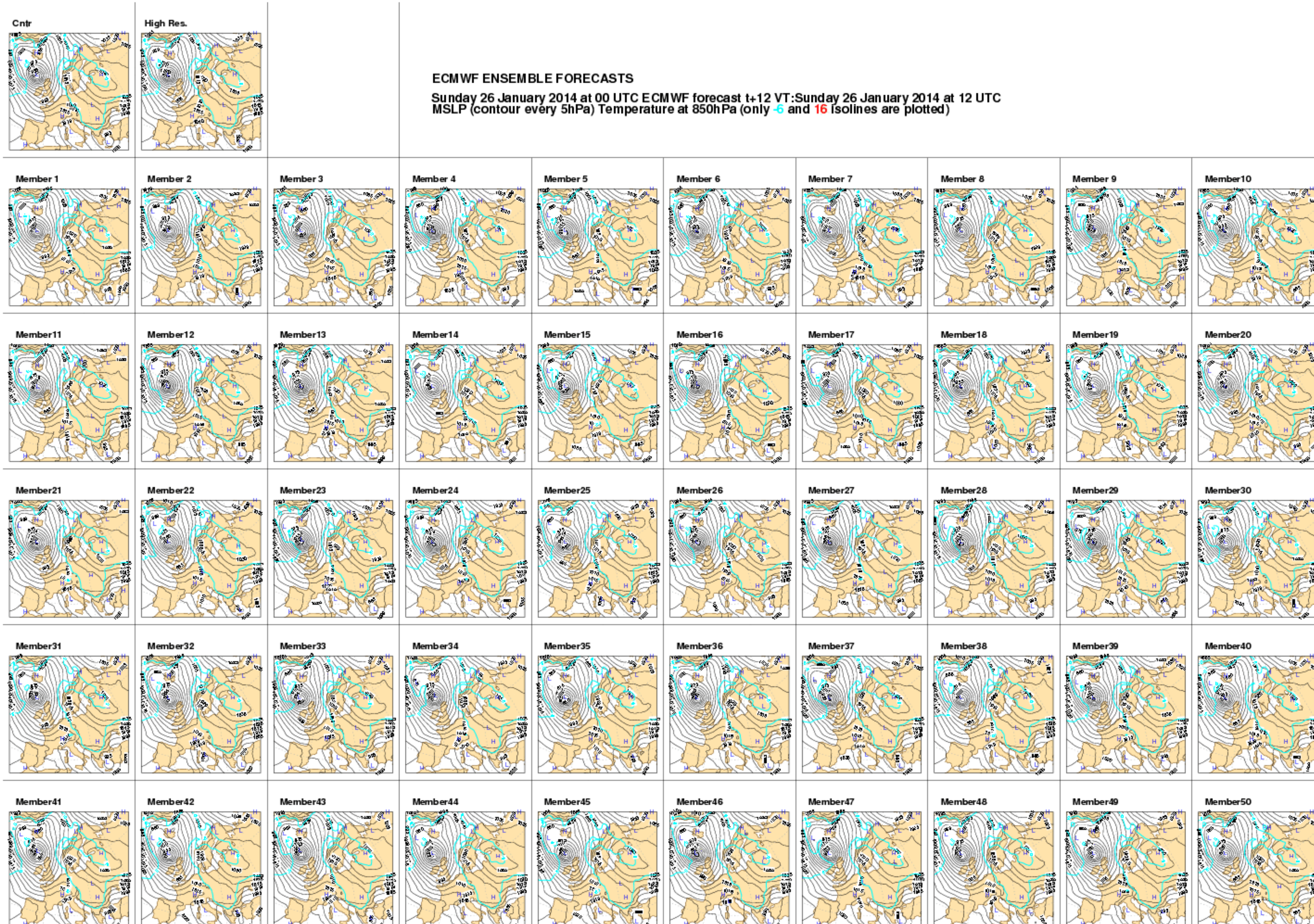




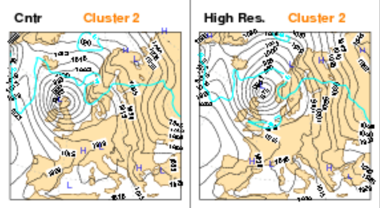
The ECMWF ensemble

- 91 levels, 32km (T639) to day 10, then 65km (T319) to day 15
- 1 control + 50 perturbed members
- Runs twice per day (00 and 12)
- Coupled to ocean model from start of forecast
- Extended to 32 days twice per week for monthly forecast (00 Thursday, Monday)

ENS

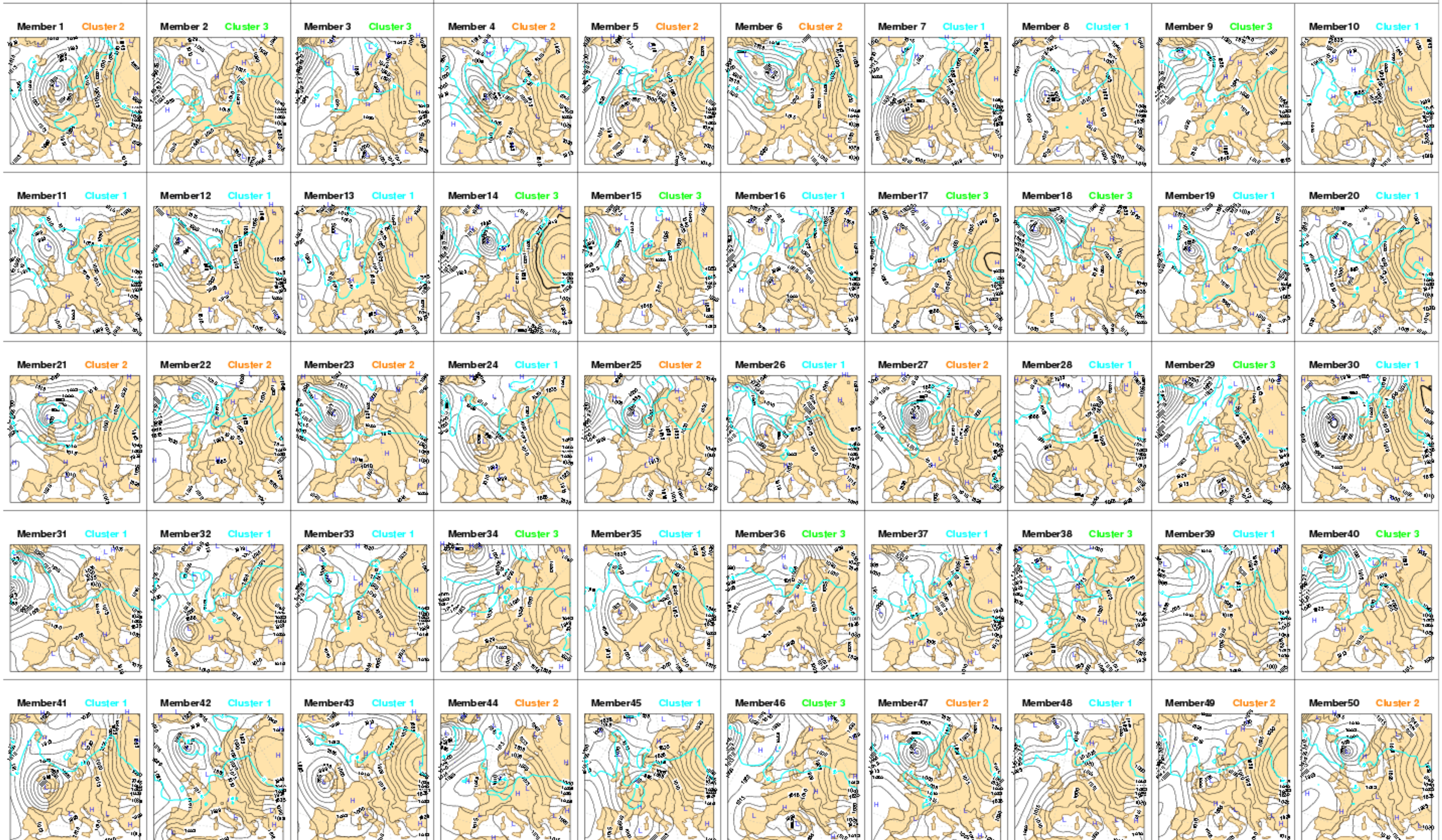

Cntr
High Res.
ECMWF ENSEMBLE FORECASTS
Sunday 26 January 2014 at 00 UTC ECMWF forecast t+12 VT:Sunday 26 January 2014 at 12 UTC
MSLP (contour every 5hPa) Temperature at 850hPa (only -6 and 16 isolines are plotted)
Member 1
Member 2
Member 3
Member 4
Member 5
Member 6
Member 7
Member 8
Member 9
Member10
Member11
Member12
Member13
Member14
Member15
Member16
Member17
Member18
Member19
Member20
Member21
Member22
Member23
Member24
Member25
Member26
Member27
Member28
Member29
Member30
Member31
Member32
Member33
Member34
Member35
Member36
Member37
Member38
Member39
Member40
Member41
Member42
Member43
Member44
Member45
Member46
Member47
Member48
Member49
Member50

ENS

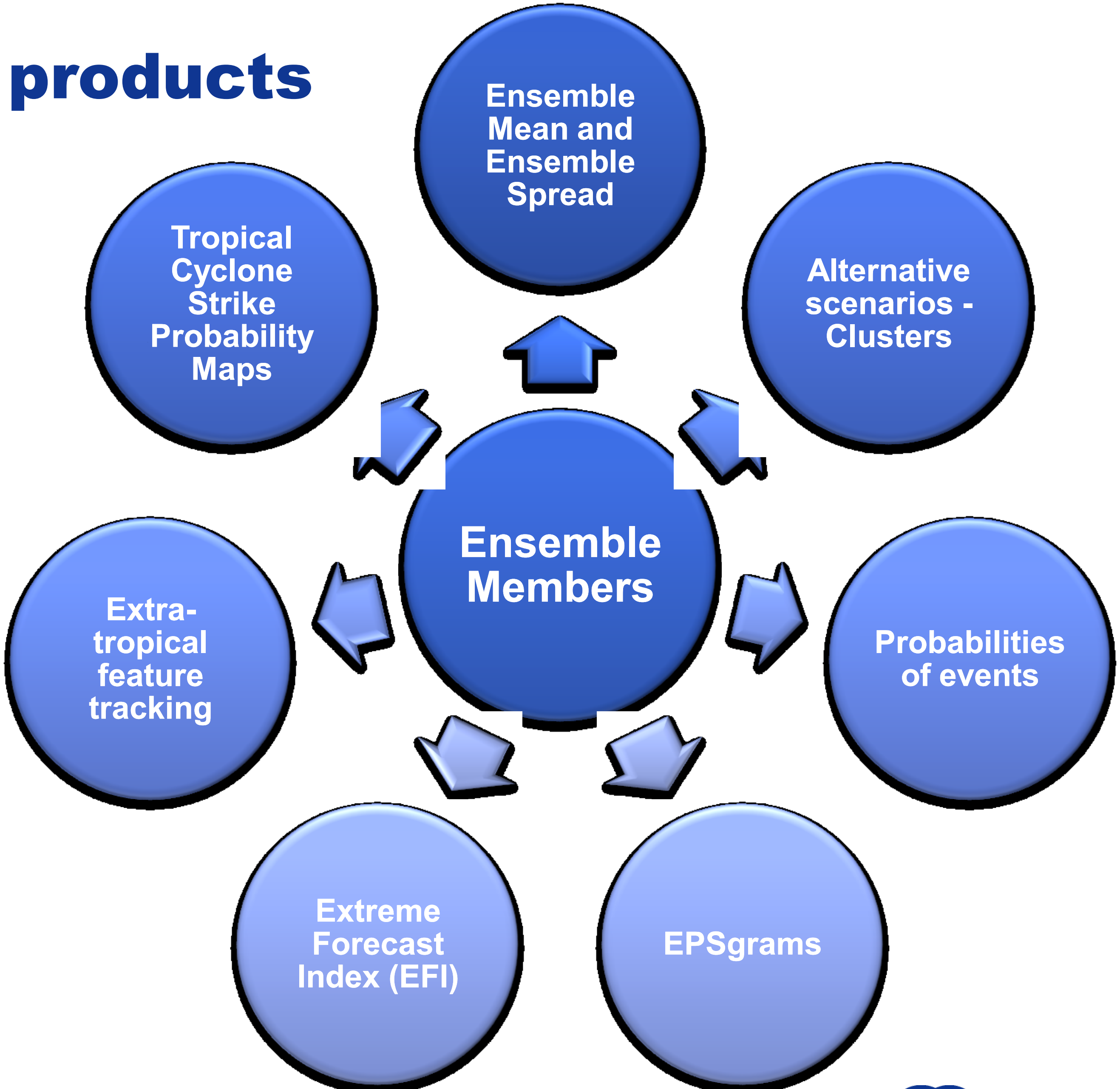


ECMWF ENSEMBLE FORECASTS

Sunday 26 January 2014 at 00 UTC ECMWF forecast t+168 VT:Sunday 02 February 2014 at 00 UTC
MSLP (contour every 5hPa) Temperature at 850hPa (only -6 and 16 isolines are plotted)

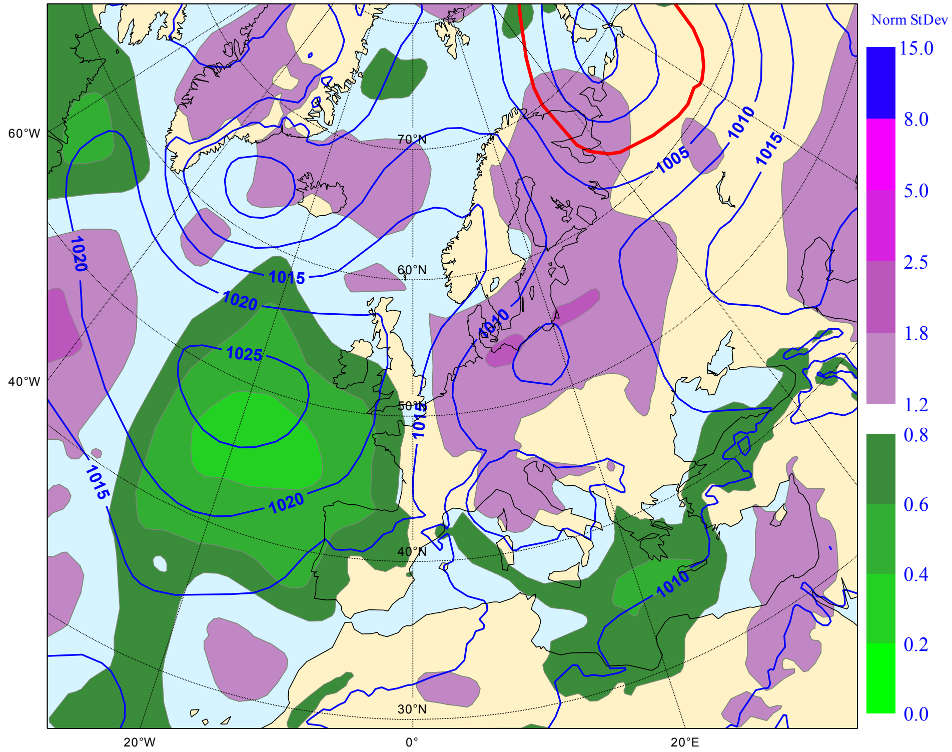


ENS products

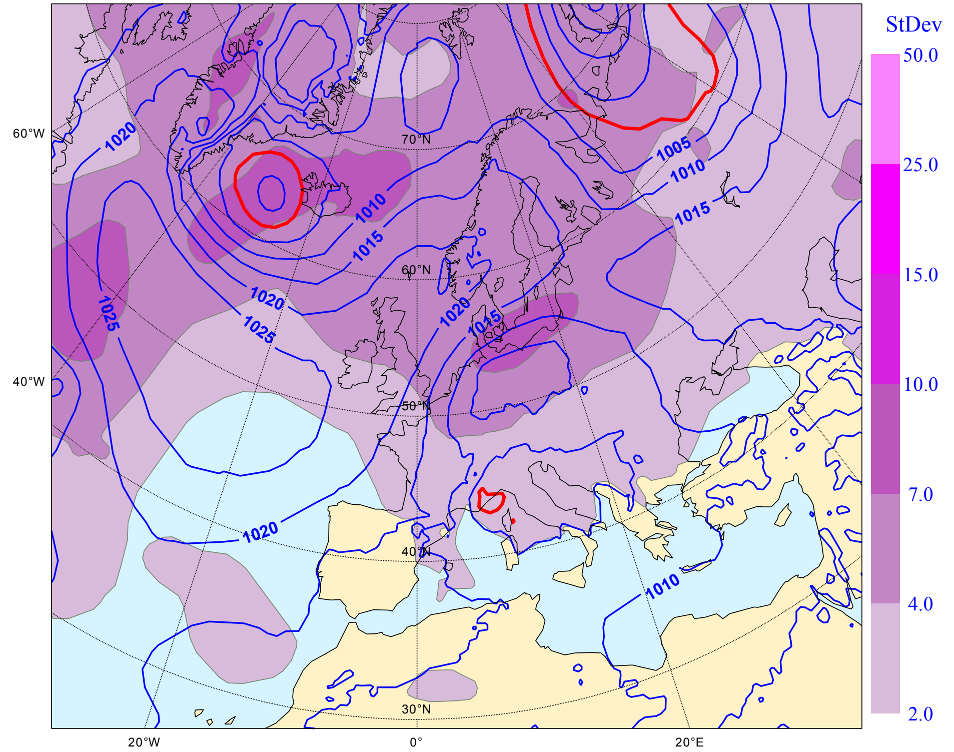


Ensemble mean and spread

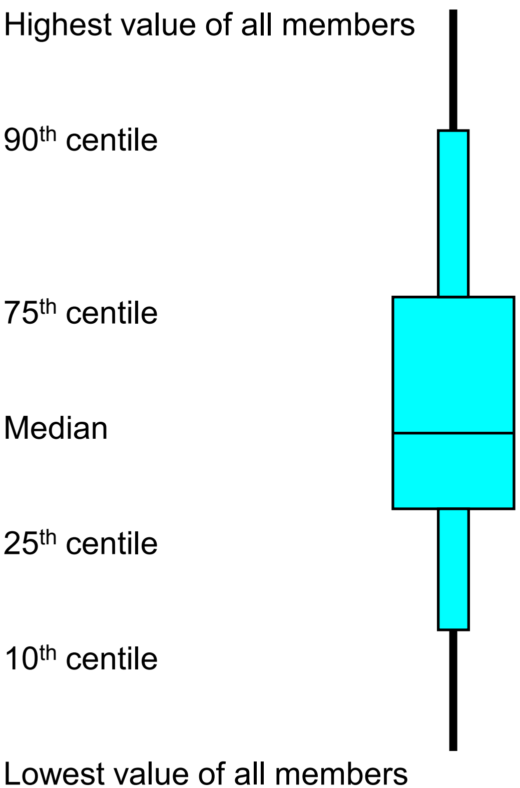
Monday 11 October 2010 12UTC ECMWF Forecast t+120 VT: Saturday 16 October 2010 12UTC
Mean sea level pressure (MSLP) Ensemble Mean and Normalised Standard Deviation (shaded)



Monday 11 October 2010 12UTC ECMWF Forecast t+120 VT: Saturday 16 October 2010 12UTC
Mean sea level pressure (MSLP) Deterministic Forecast and Standard Deviation (shaded)

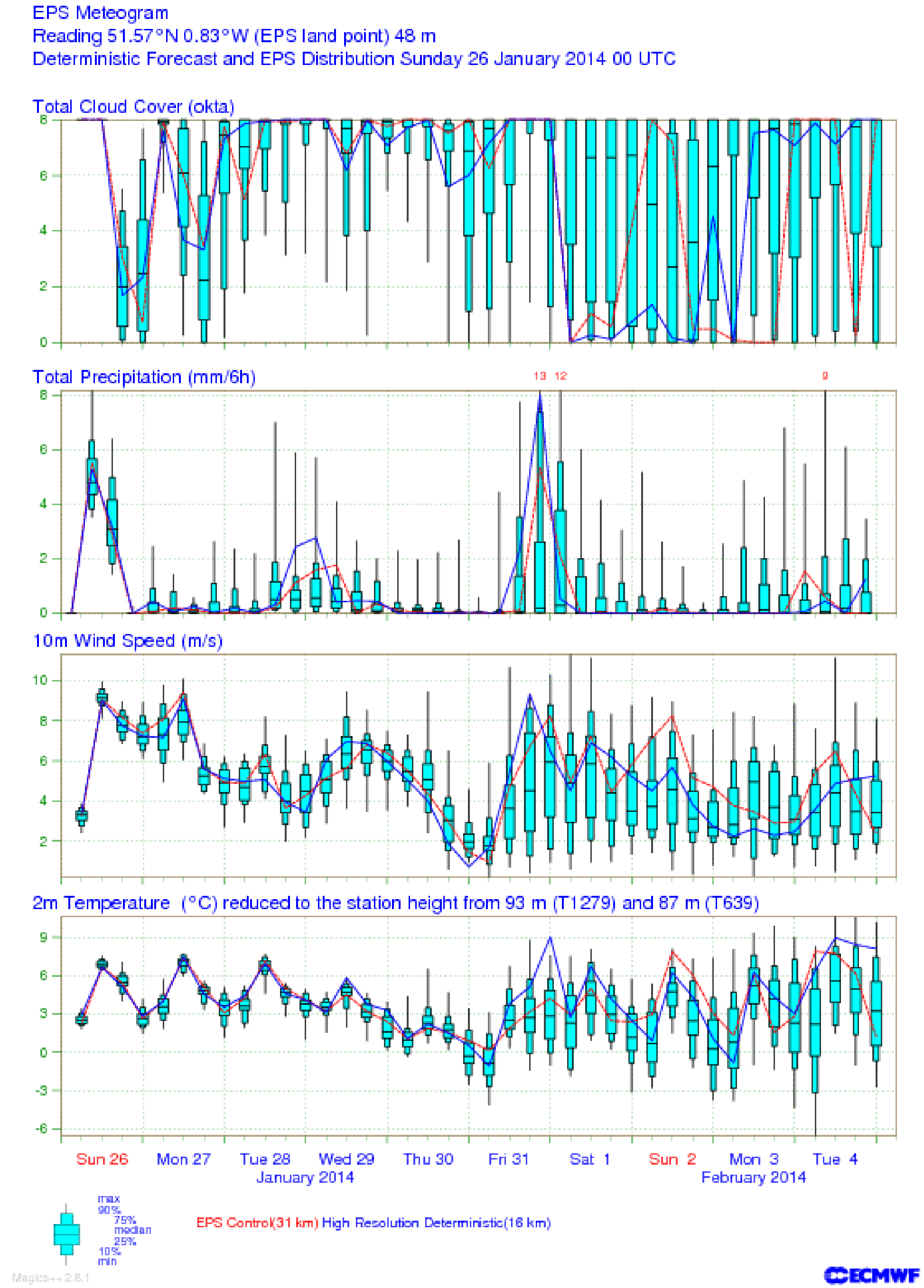


ENS forecasts: timeseries (EPSgram)



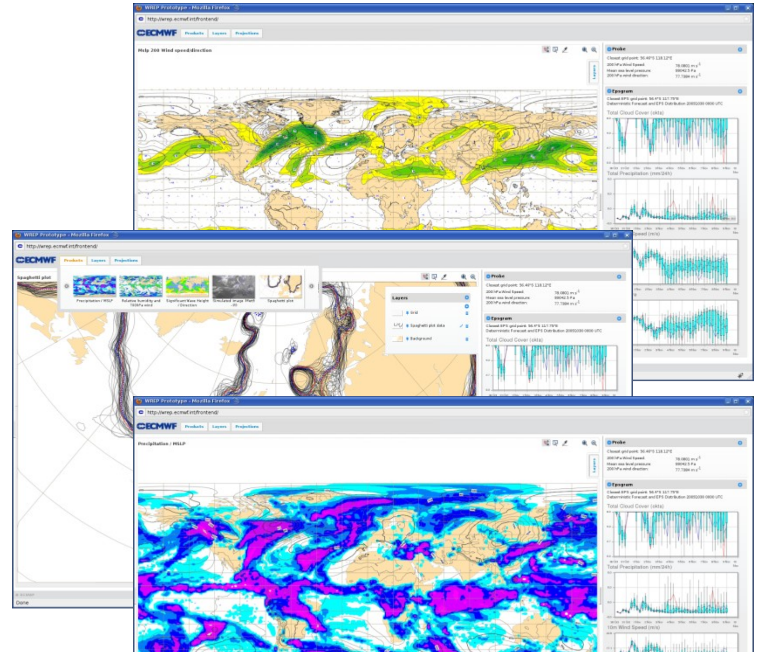
EPSgram for Reading

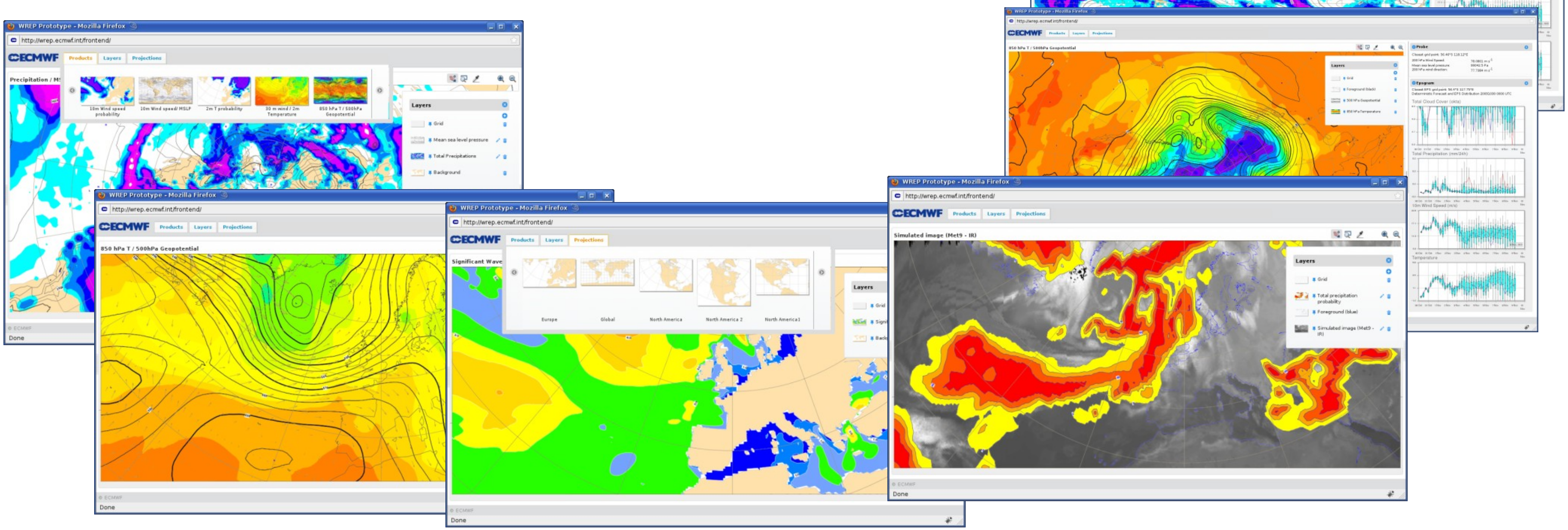
Start Sun 26/01/14 00 UTC



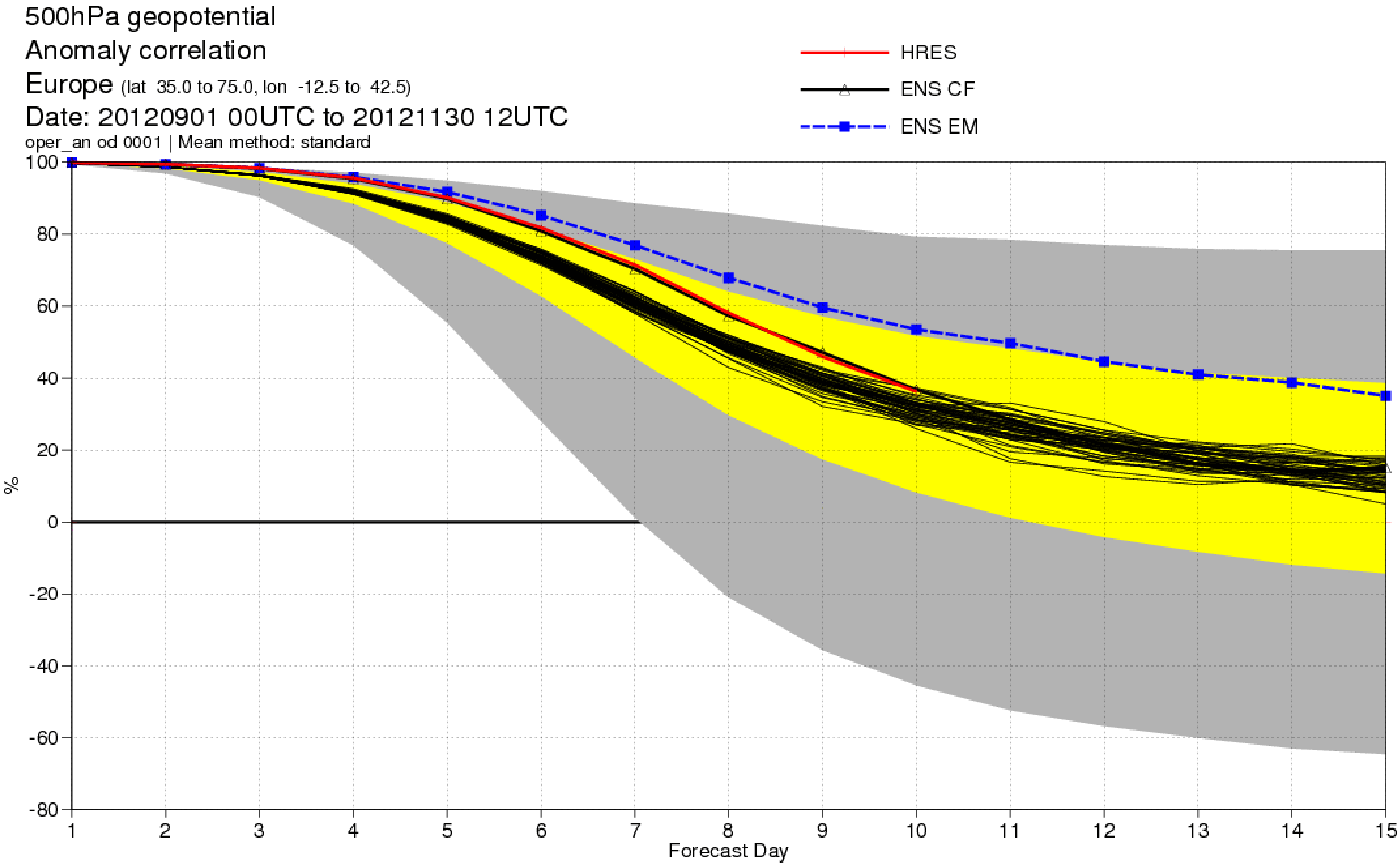
ecCharts

- Interactivity: zooming, panning, …
- Customisation:
  - Probabilities threshold, …
  - Show/hide, add/remove layers
- Related products: Meteograms



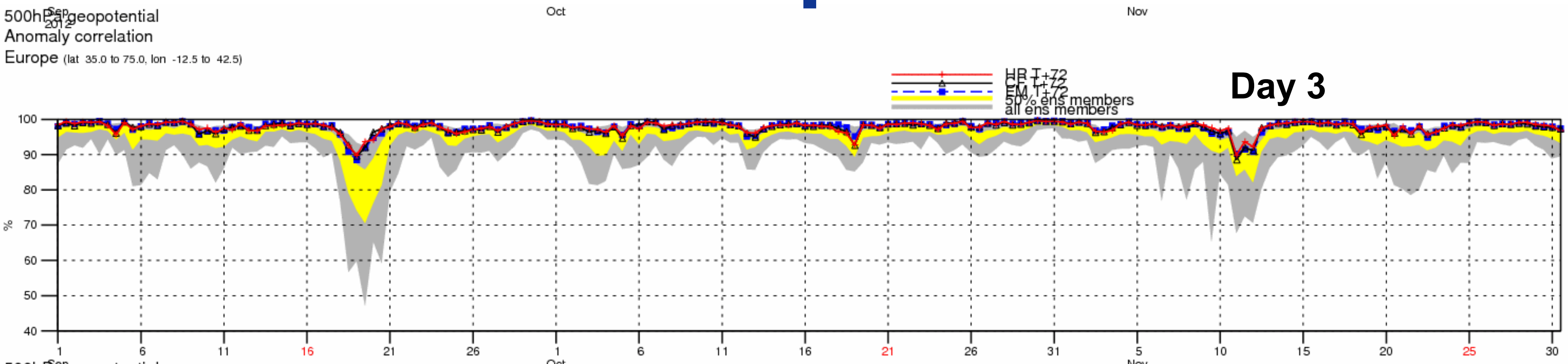


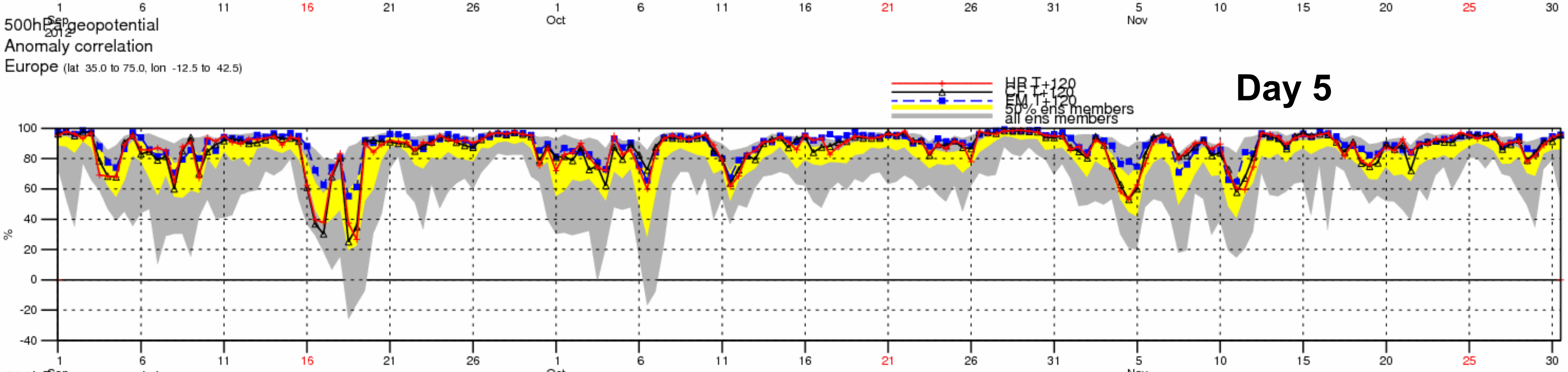
Ensemble skill Z500 Europe

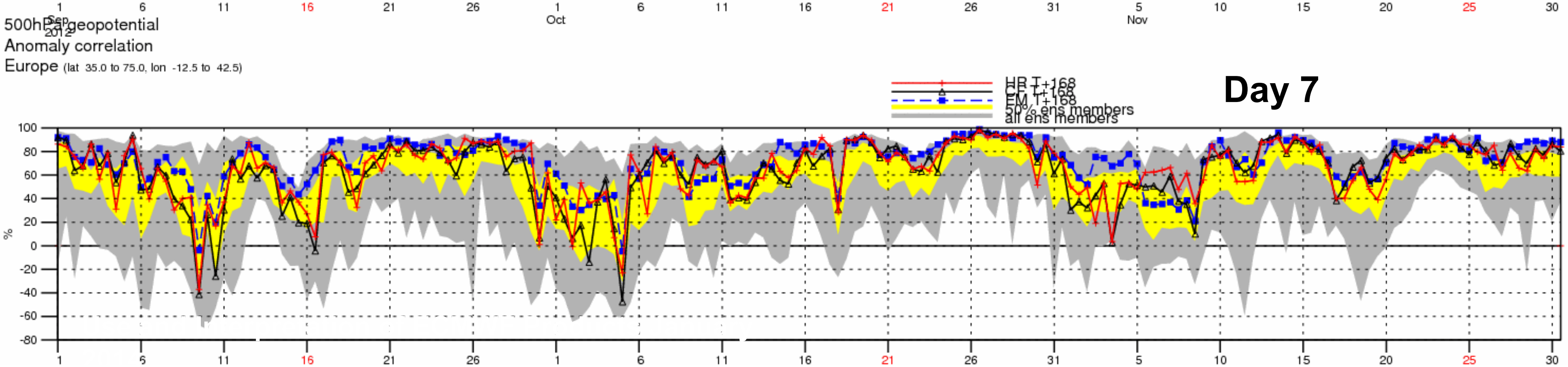


ECMWF

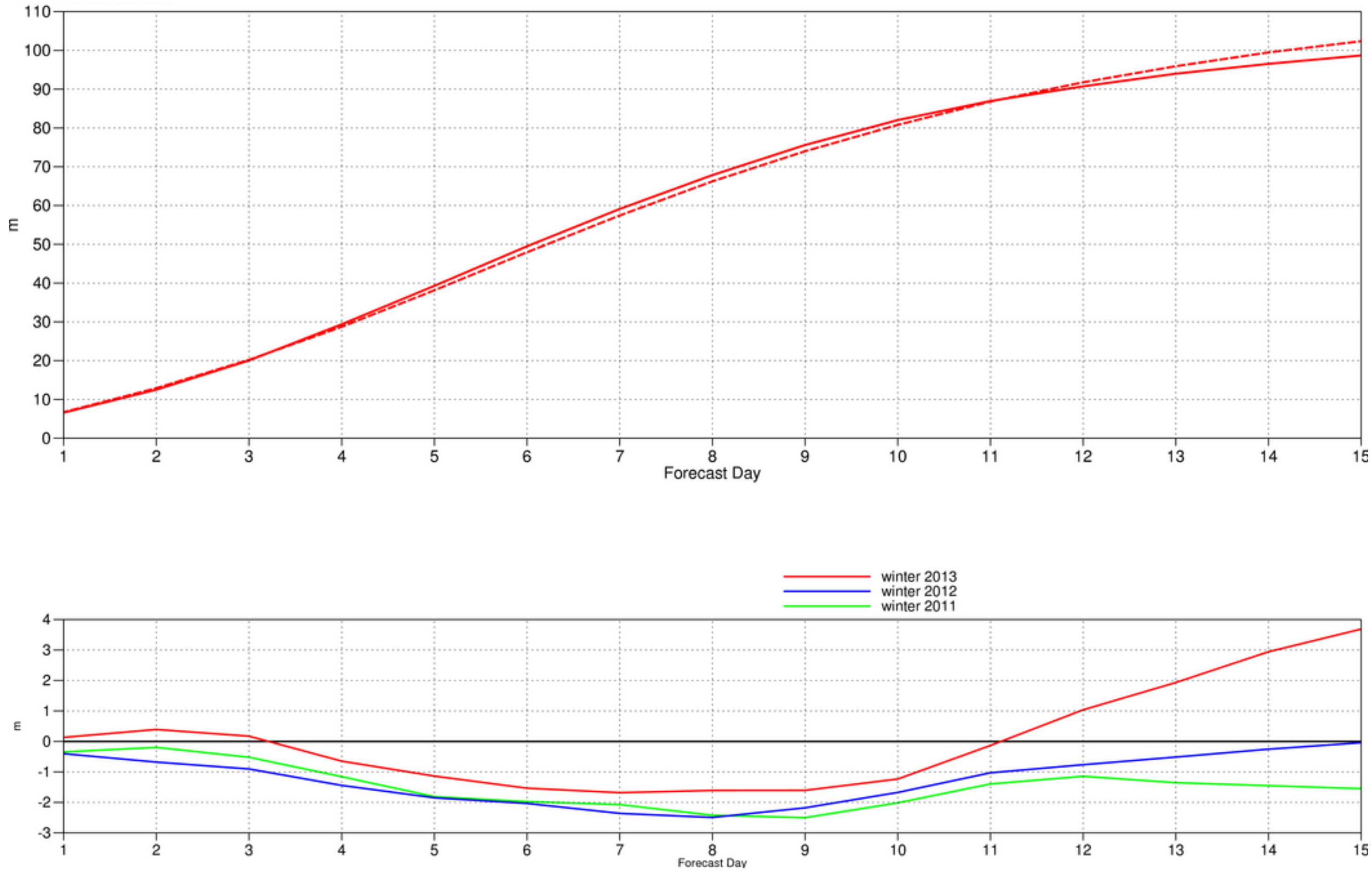
Ensemble: Z500 Europe







ENS spread and error, Z500, N.Hem

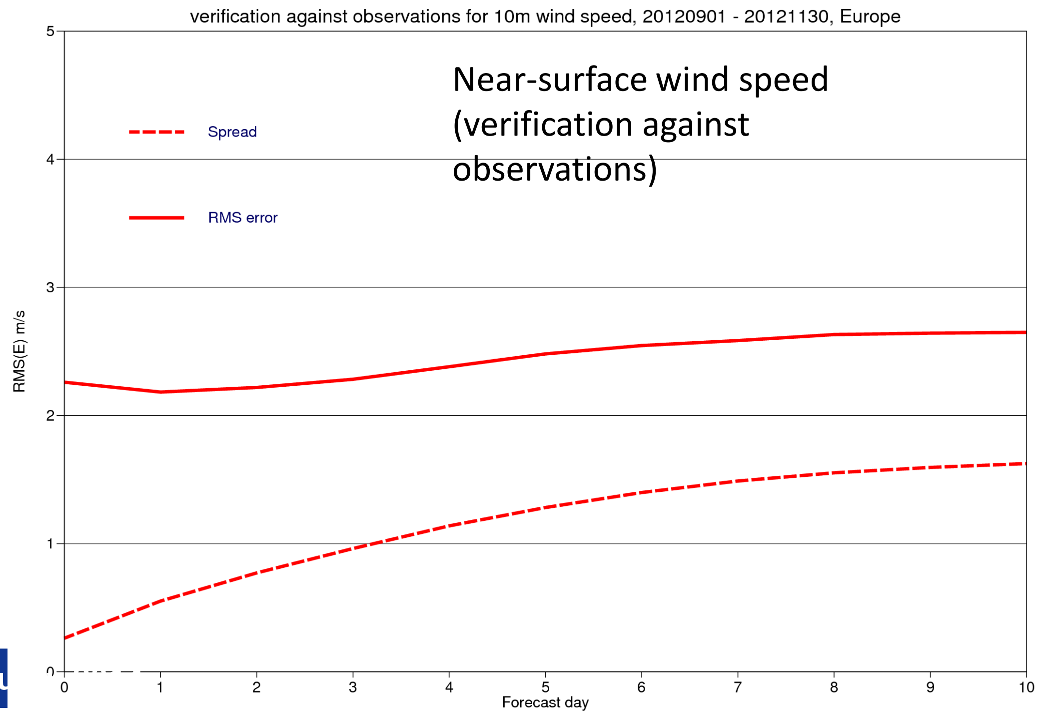


EPS spread (dashed), RMS error of ensemble-mean (full lines), and their difference (below) for Z500 hPa in winter 2010-11 (green), 2011-12 (blue) and 2012-13 (red).

ECMWF

Surface perturbations

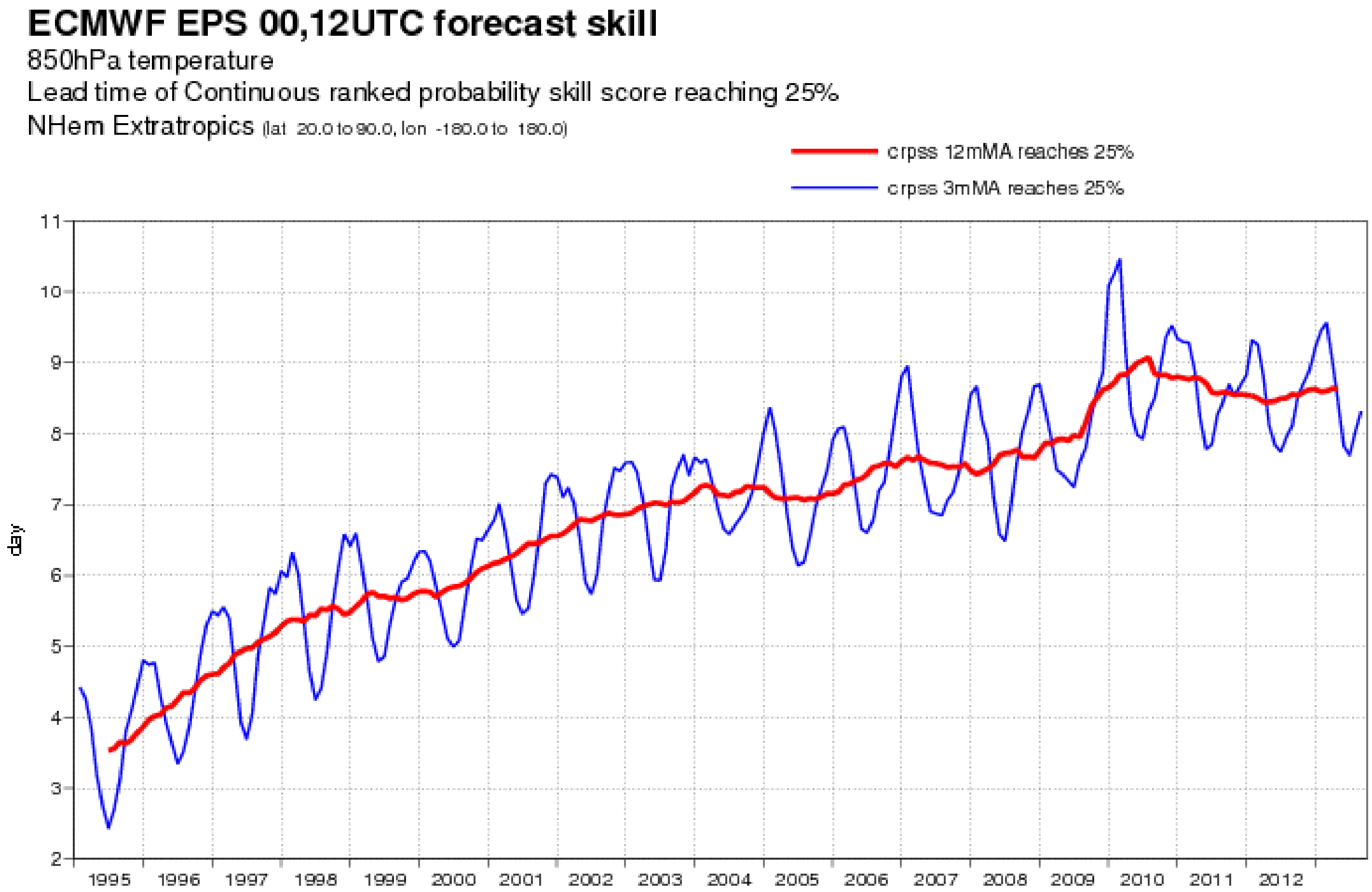
- ENS had too little spread for near surface weather parameters (e.g. 10-m wind)
  - representativeness (an individual observation is not equivalent to a model grid box average) and errors in the observations
  - ENS resolution: difficult to represent small-scale phenomena such as sting jets
  - Additional sources of uncertainty?
- Land-surface perturbations
  - Added November 2013



Ensemble spread (dashed) and root-mean-square error of ensemble-mean (solid) autumn (September-November) 2012 over Europe

ENS Probabilistic Score

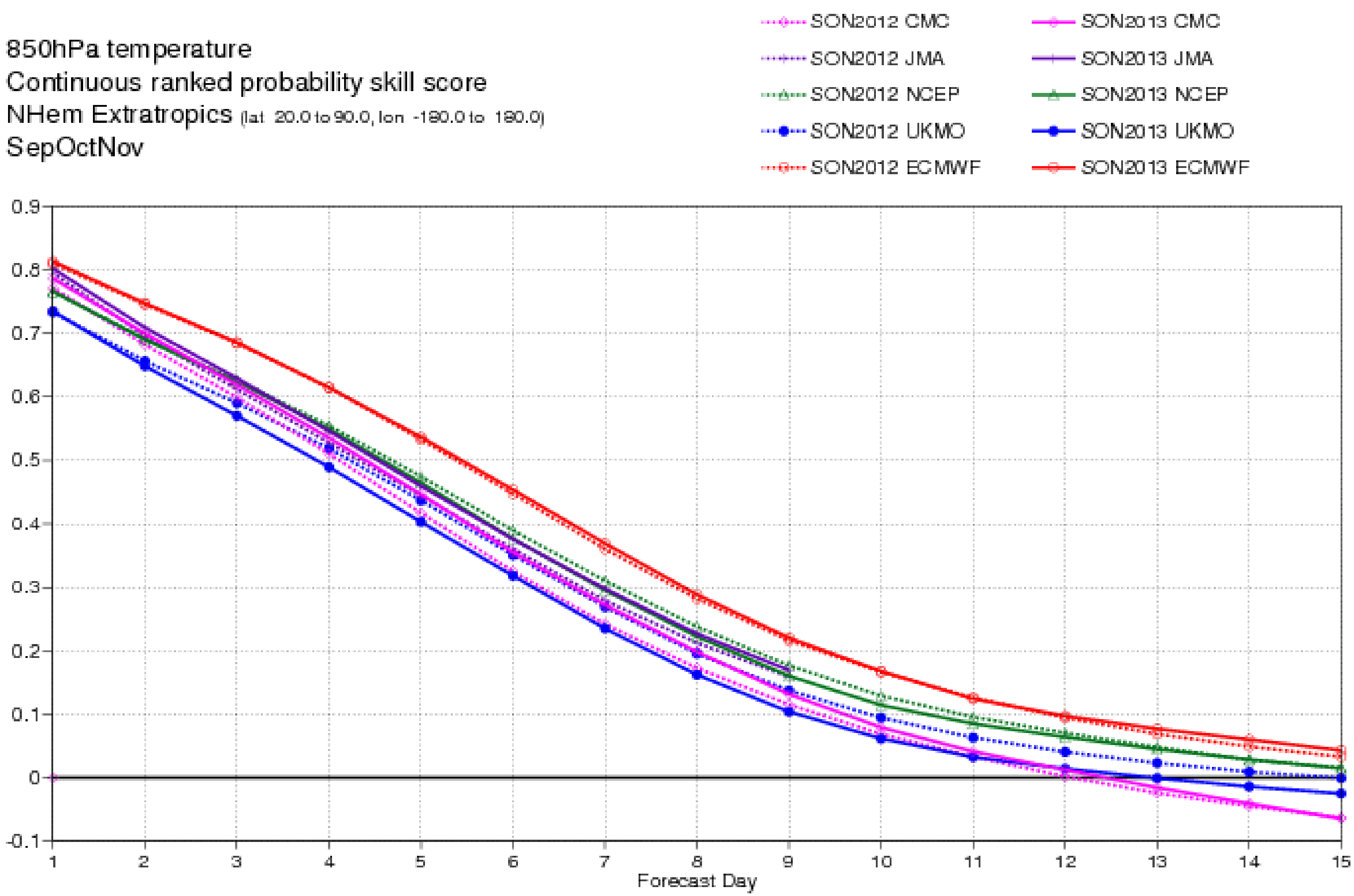
CRPSS, Temperature at 850 hPa N hemisphere



Monthly score (blue), and 12-month running mean (red) of Continuous Ranked Probability Skill Score. Day at which score reaches 25%.

ECMWF

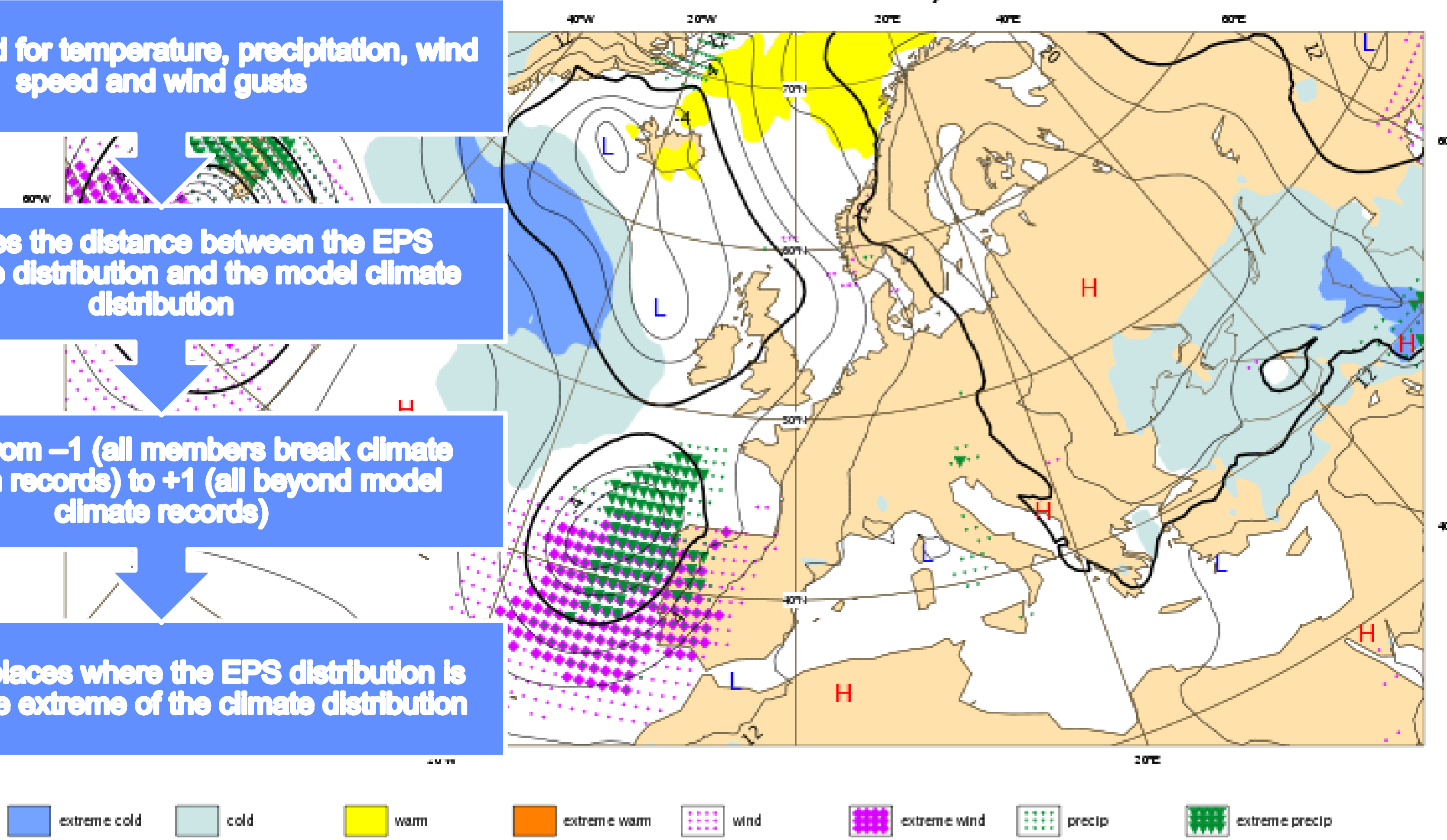
ENS Probabilistic Score
CRPSS, Temperature at 850 hPa N hemisphere



ECMWF

Extreme forecast index (EFI)

Anomalous weather predicted by EPS: Tuesday 25 October 2011 at 00 UTC
1000 hPa Z ensemble mean (Wednesday 26 October 2011 at 12 UTC)
and EFI values for Total precipitation,maximum 10m wind gust and mean 2m temperature (all 24h)
valid for 24hours from Wednesday 26 October 2011 at 00 UTC to Thursday 27 October 2011 at 00 UTC



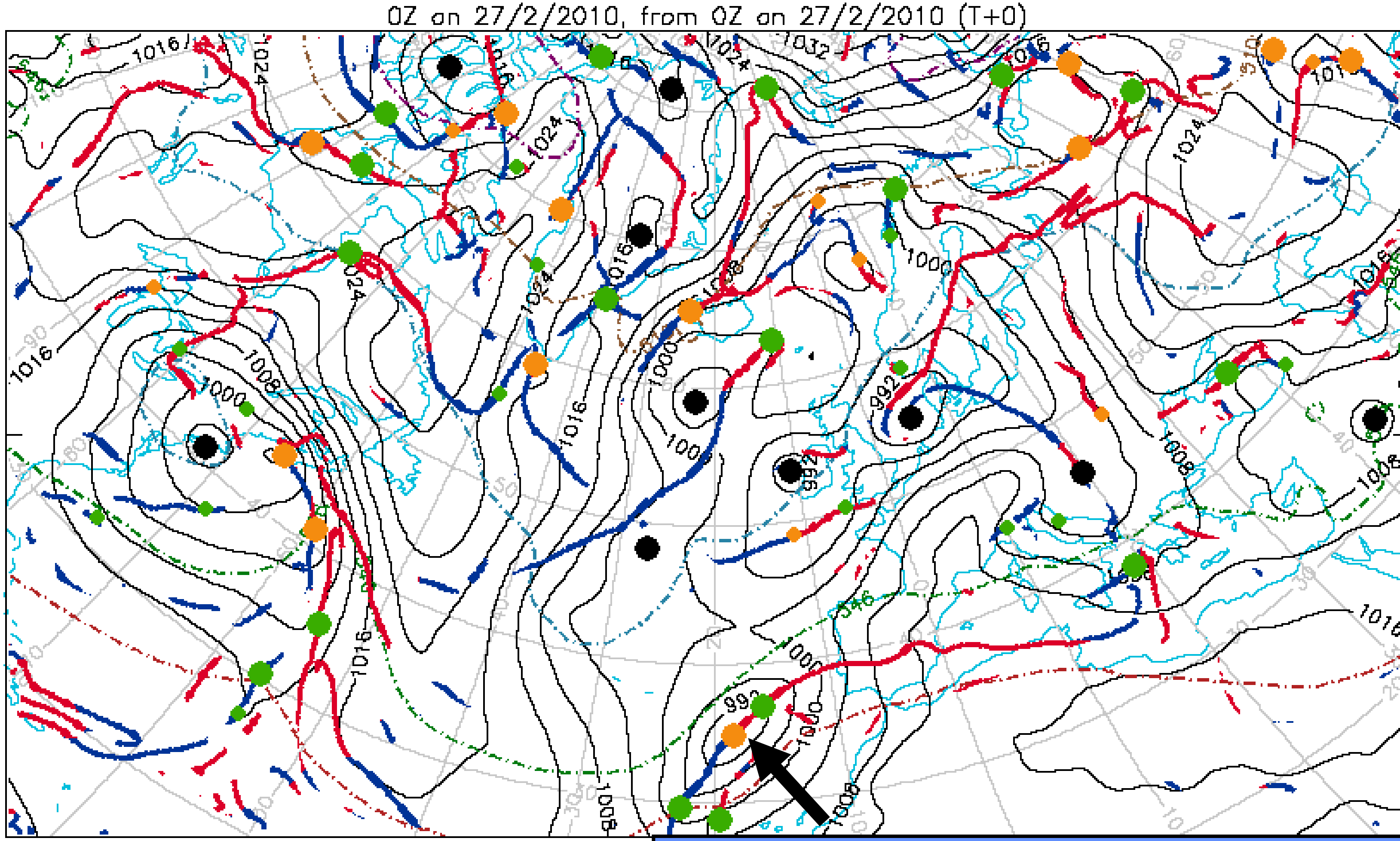
Is computed for temperature, precipitation, wind speed and wind gusts

Measures the distance between the EPS cumulative distribution and the model climate distribution

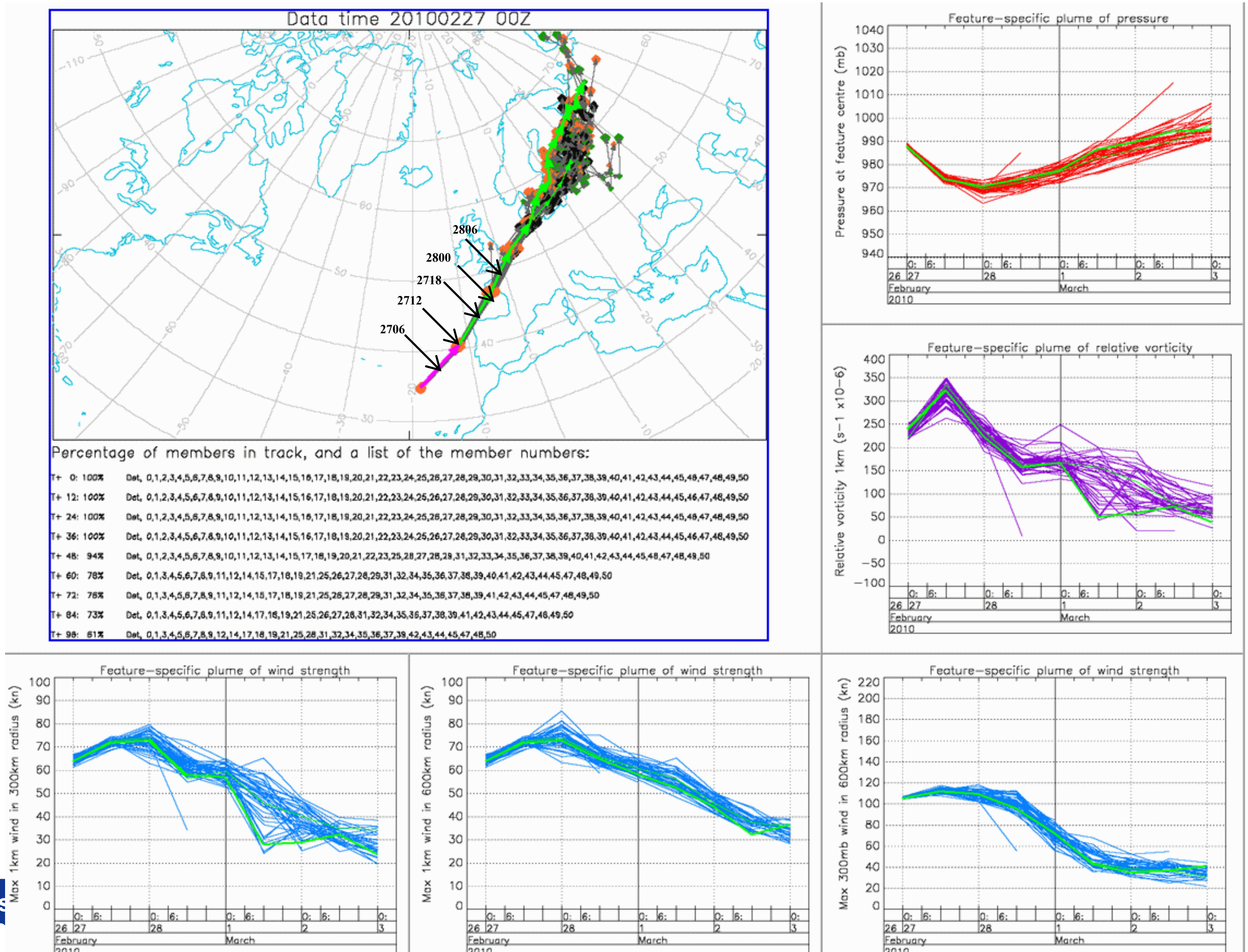
Ranges from –1 (all members break climate minimum records) to +1 (all beyond model climate records)

Indicates places where the EPS distribution is towards the extreme of the climate distribution

Extra-tropical feature tracking: Xynthia



User can click on any spot (= cyclonic feature) to see how that feature evolves in the EPS

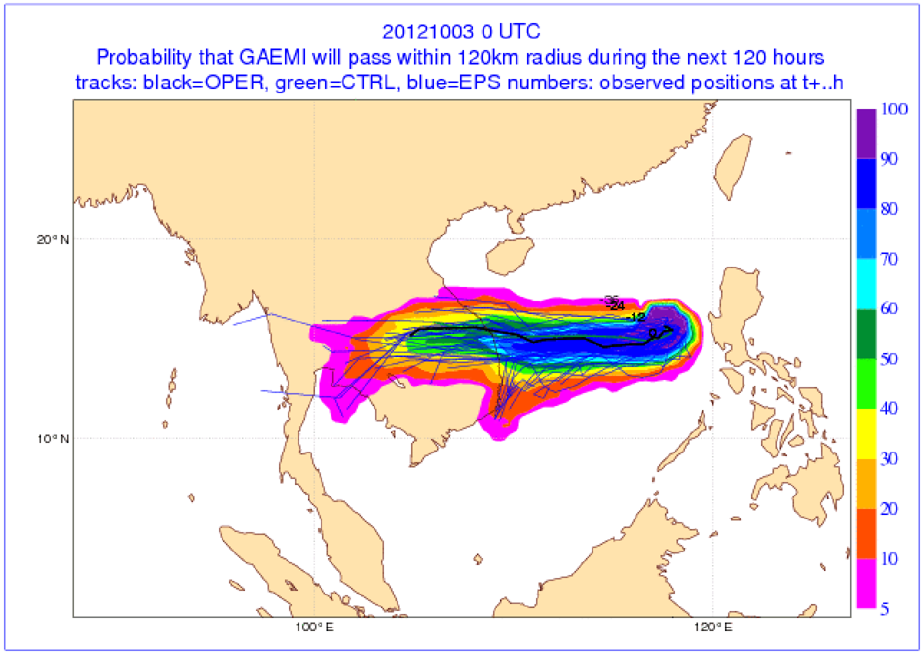


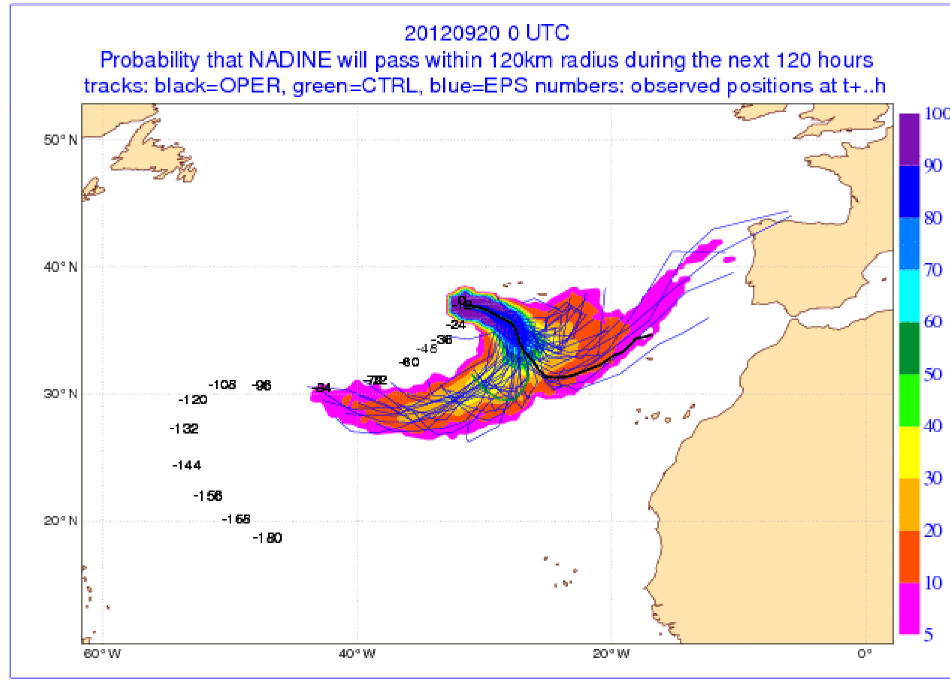
Data time 20100227 00Z
2806
2800
2718
2712
2706
Percentage of members in track, and a list of the member numbers:
T+ 0: 100% Det, 0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50
T+ 12: 100% Det, 0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50
T+ 24: 100% Det, 0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50
T+ 36: 100% Det, 0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50
T+ 48: 94% Det, 0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,17,18,19,20,21,22,23,25,26,27,28,29,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50
T+ 60: 76% Det, 0,1,3,4,5,6,7,8,9,11,12,14,15,17,18,19,21,25,26,27,28,29,31,32,34,35,36,37,38,39,40,41,42,43,44,45,47,48,49,50
T+ 72: 76% Det, 0,1,3,4,5,6,7,8,9,11,12,14,15,17,18,19,21,25,26,27,28,29,31,32,34,35,36,37,38,39,41,42,43,44,45,47,48,49,50
T+ 84: 73% Det, 0,1,3,4,5,6,7,8,9,11,12,14,17,18,19,21,25,26,27,28,31,32,34,35,36,37,38,39,41,42,43,44,45,47,48,49,50
T+ 96: 61% Det, 0,1,3,4,5,6,7,8,9,12,14,17,18,19,21,25,28,31,32,34,35,36,37,39,42,43,44,45,47,48,50
Feature-specific plume of pressure
Pressure at feature centre (mb)
Feature-specific plume of relative vorticity
Relative vorticity 1km (s-1 x10-6)
Feature-specific plume of wind strength
Max 1km wind in 300km radius (kn)
Feature-specific plume of wind strength
Max 1km wind in 600km radius (kn)
Feature-specific plume of wind strength
Max 300mb wind in 600km radius (kn)
26 27 28 1 2 3
February March
2010

Tropical cyclone tracks

Gamei

Nadine





strike probability

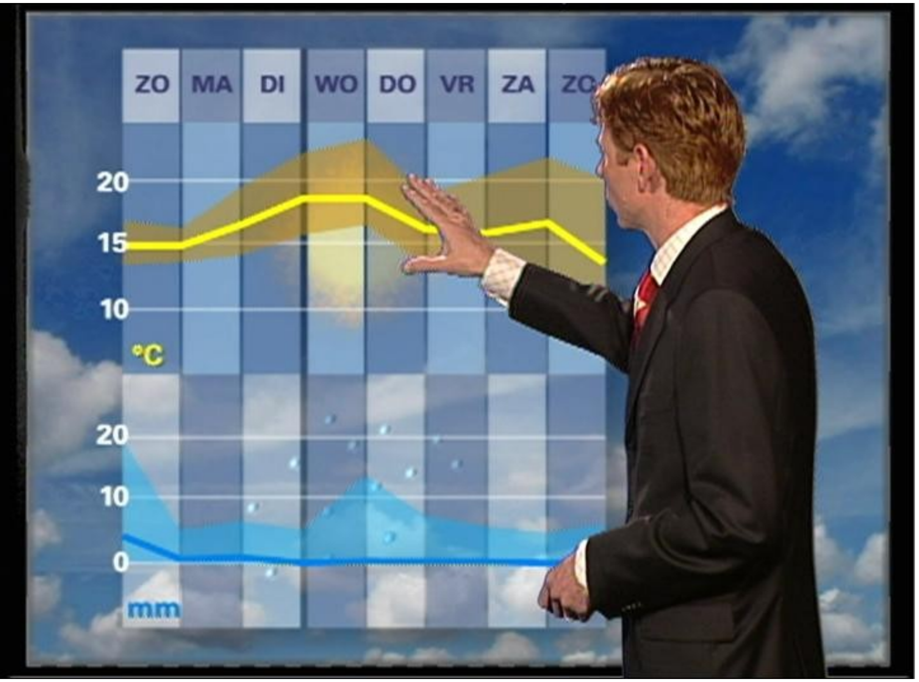
ECMWF

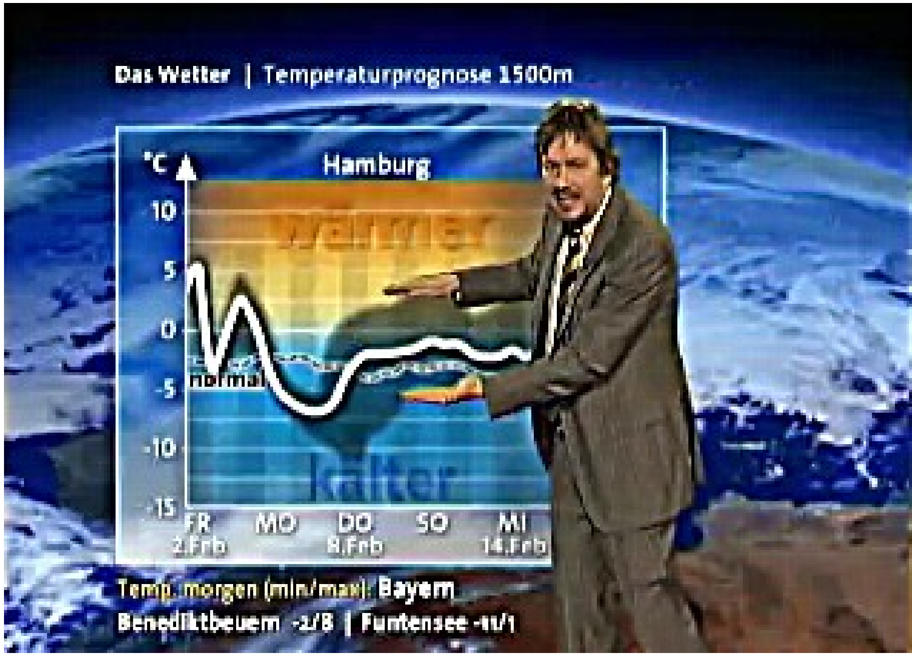
Great! But how can that help users who must make yes/no decisions?

ENS – communicating uncertainty

- All forecasts have errors
- It can be important for the user to know about the uncertainty in a forecast
  - what else could happen? what is the worst possibility?
- This is not a new idea
  - Forecasters are used to adjusting their forecast with their experience of model errors (flow dependence, forecast range dependency)
  - Inconsistency of the forecasts (in time, from one model to the other) were used as indication of the (un-)predictability of scenarios
- Ensembles give more information – they provide an explicit, detailed representation of model uncertainties, and potential of unusual events

Uncertainty information to public



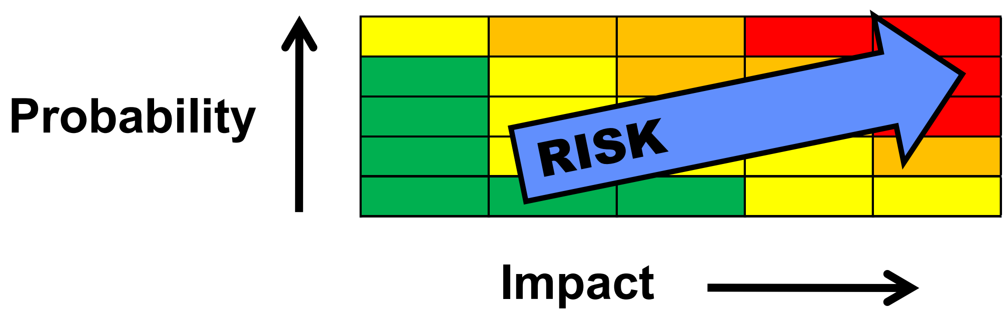


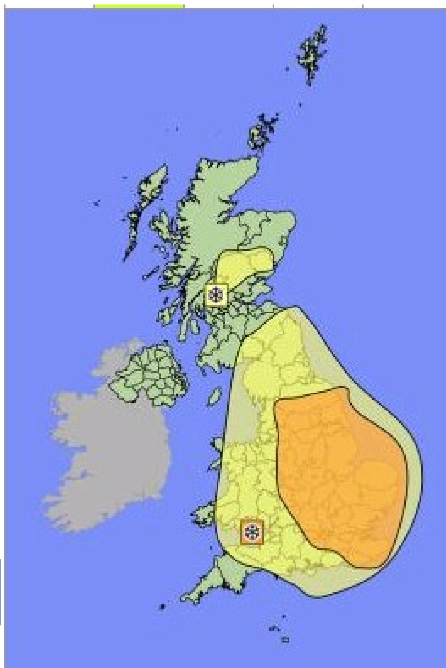
ECMWF

Uncertainty information to public

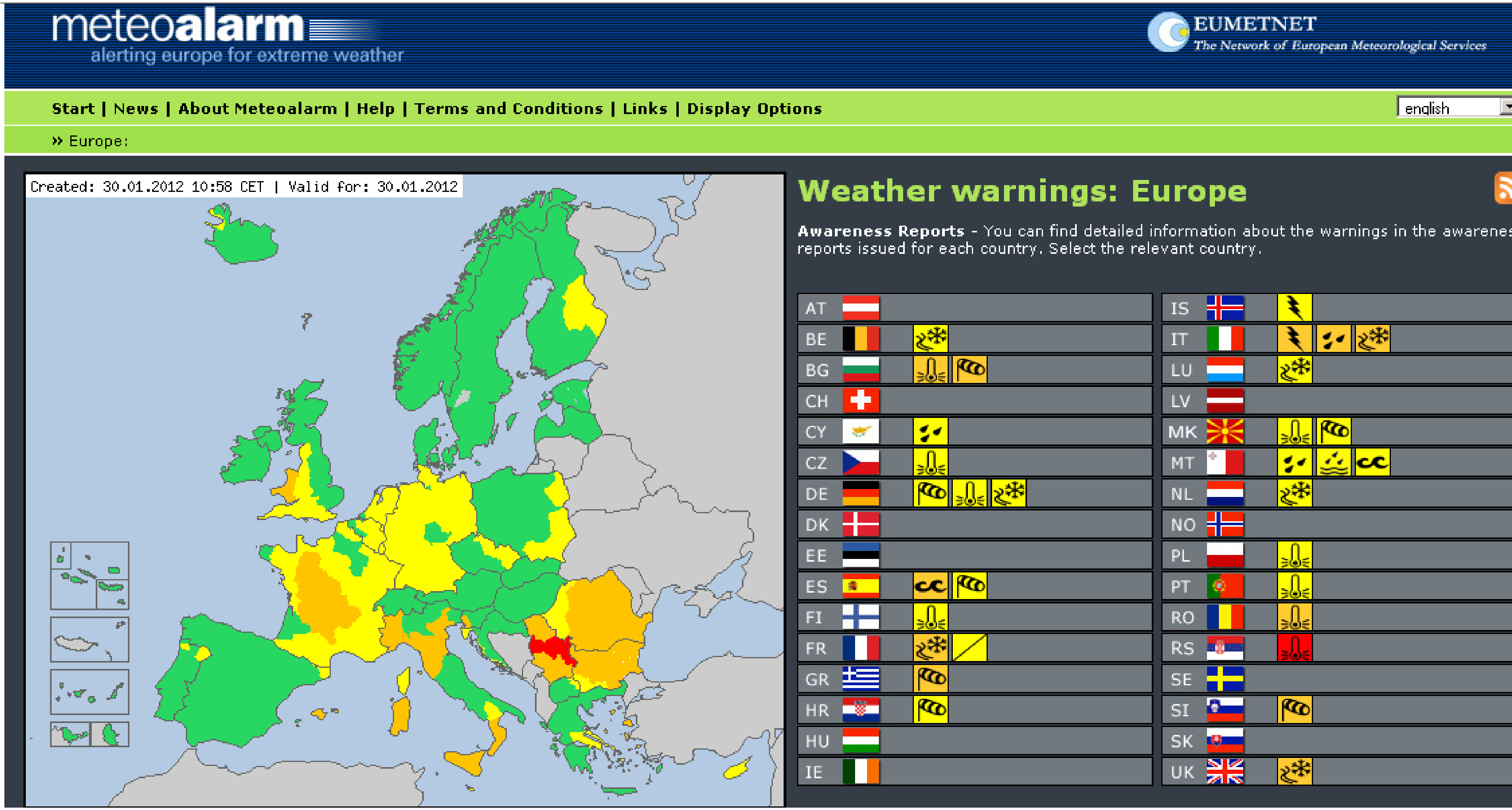
Value: the economic or societal worth of forecasts

- Forecasts only have value if people use them
  - make a decision or take an action which would not otherwise have been made
- Decisions can be based on deterministic forecasts, but …
- Decisions involve assessment of risk
- Risk = probability x impact
- To make a good decision need to know the probability and the impact (consequence to the individual user)





MeteoAlarm


meteoalarm
alerting europe for extreme weather
EUMETNET
The Network of European Meteorological Services
Start | News | About Meteoalarm | Help | Terms and Conditions | Links | Display Options
english
» Europe:
Created: 30.01.2012 10:58 CET | Valid for: 30.01.2012
Weather warnings: Europe
Awareness Reports - You can find detailed information about the warnings in the awareness reports issued for each country. Select the relevant country.
AT
BE
BG
CH
CY
CZ
DE
DK
EE
ES
FI
FR
GR
HR
HU
IE
IS
IT
LU
LV
MK
MT
NL
NO
PL
PT
RO
RS
SE
SI
SK
UK

Summary - why do we run an ensemble?

- The best method we have to produce flow-dependent probabilistic weather forecasts
- The ensemble of forecasts provides a range of future scenarios consistent with our knowledge of the initial state and model capability
  - Provides explicit indication of uncertainty in today's forecast
  - Range of ensemble based products for different users
- Ensembles provide the required input for a range of application models (hydrology, ship routing, energy demand), explicitly propagating the atmospheric uncertainty
- Read more in the ECMWF products User Guide
  - http://www.ecmwf.int/products/forecasts/guide/

ECMWF

Ensemble references

- Berner, J., G. J. Shutts, M. Leutbecher, and T. N. Palmer (2009), A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system, J. Atmos. Sci., 66, 603–626.
- Buizza, R., Leutbecher, M., & Isaksen, L., 2008: Potential use of an ensemble of analyses in the ECMWF Ensemble Prediction System. Q. J. R. Meteorol. Soc., 134, 2051-2066.
- Buizza, R., Bidlot, J.-R., Wedi, N., Fuentes, M., Hamrud, M., Holt, G., & Vitart, F., 2007: The new ECMWF VAREPS. *Q. J. Roy. Meteorol. Soc.*, 133, 681-695 (also EC TM 499).
- Buizza, R., 2008: Comparison of a 51-member low-resolution (TL399L62) ensemble with a 6-member high-resolution (TL799L91) lagged-forecast ensemble. *Mon. Wea. Rev.*, 136, 3343-3362 (also EC TM 559).
- Leutbecher, M. 2005: On ensemble prediction using singular vectors started from forecasts. ECMWF TM 462, pp 11.
- Leutbecher, M. & T.N. Palmer, 2008: Ensemble forecasting. *J. Comp. Phys.*, 227, 3515-3539 (also EC TM 514).
- Molteni, F., Buizza, R., Palmer, T. N., & Petroliagis, T., 1996: The new ECMWF ensemble prediction system: methodology and validation. *Q. J. R. Meteorol. Soc.*, 122, 73-119.
- Palmer, T N, Buizza, R., Leutbecher, M., Hagedorn, R., Jung, T., Rodwell, M, Virat, F., Berner, J., Hagel, E., Lawrence, A., Pappenberger, F., Park, Y.-Y., van Bremen, L., Gilmour, I., & Smith, L., 2007: The ECMWF Ensemble Prediction System: recent and on-going developments. A paper presented at the 36th Session of the ECMWF Scientific Advisory Committee (also EC TM 540).
- Palmer, T. N., Buizza, R., Doblas-Reyes, F., Jung, T., Leutbecher, M., Shutts, G. J., Steinheimer M., & Weisheimer, A., 2009: Stochastic parametrization and model uncertainty. ECMWF RD TM 598, Shinfield Park, Reading RG2-9AX, UK, pp. 42.
- Richardson, D. S., 2000. Skill and relative economic value of the ECMWF Ensemble Prediction System. *Q. J. R. Meteorol. Soc.,* 126, 649-668.
- Richardson, D.S., 2003. Economic value and skill. In *Forecast verification: a practitioner's guide in atmospheric science*, Jolliffe, I. T. and Stephenson, D. B., Eds., Wiley, 240pp.
- Vitart, F., Buizza, R., Alonso Balmaseda, M., Balsamo, G., Bidlot, J. R., Bonet, A., Fuentes, M., Hofstadler, A., Molteni, F., & Palmer, T. N., 2008: The new VAREPS-monthly forecasting system: a first step towards seamless prediction. *Q. J. Roy. Meteorol. Soc.*, 134, 1789-1799.
- Zsoter, E., Buizza, R., & Richardson, D., 2009: 'Jumpiness' of the ECMWF and UK Met Office EPS control and ensemble-mean forecasts'. *Mon. Wea. Rev.*, 137, 3823-3836.

ECMWF

Decision analysis - the cost-loss model

- Simplest possible case - but shows many important features
- There are only two important weather types: weather is either “good” or “bad”
- A particular user or decision maker will be affected by bad weather - they have a choice of two actions
  - If they do nothing and bad weather occurs they suffer a loss L
  - However, they can decide to take some protective action to prevent this possible loss, but it will cost C

Why is the probability forecast better?

- If the cost of protection is high wait until event is more certain
  - False alarms are more important
- If the loss is greater then protect even at low probability
  - Missed events are more important
- Changing the probability threshold at which to take action gives different hit rates and false alarm rates
- The optimal probability threshold depends on the user: $p_t = C/L$
- Using the probabilities allows decision makers to take decisive action according to their own risks – these are different for each user
- Even if the user does not have an explicit cost/loss they are still aware of the relative importance of false alarms and missed events

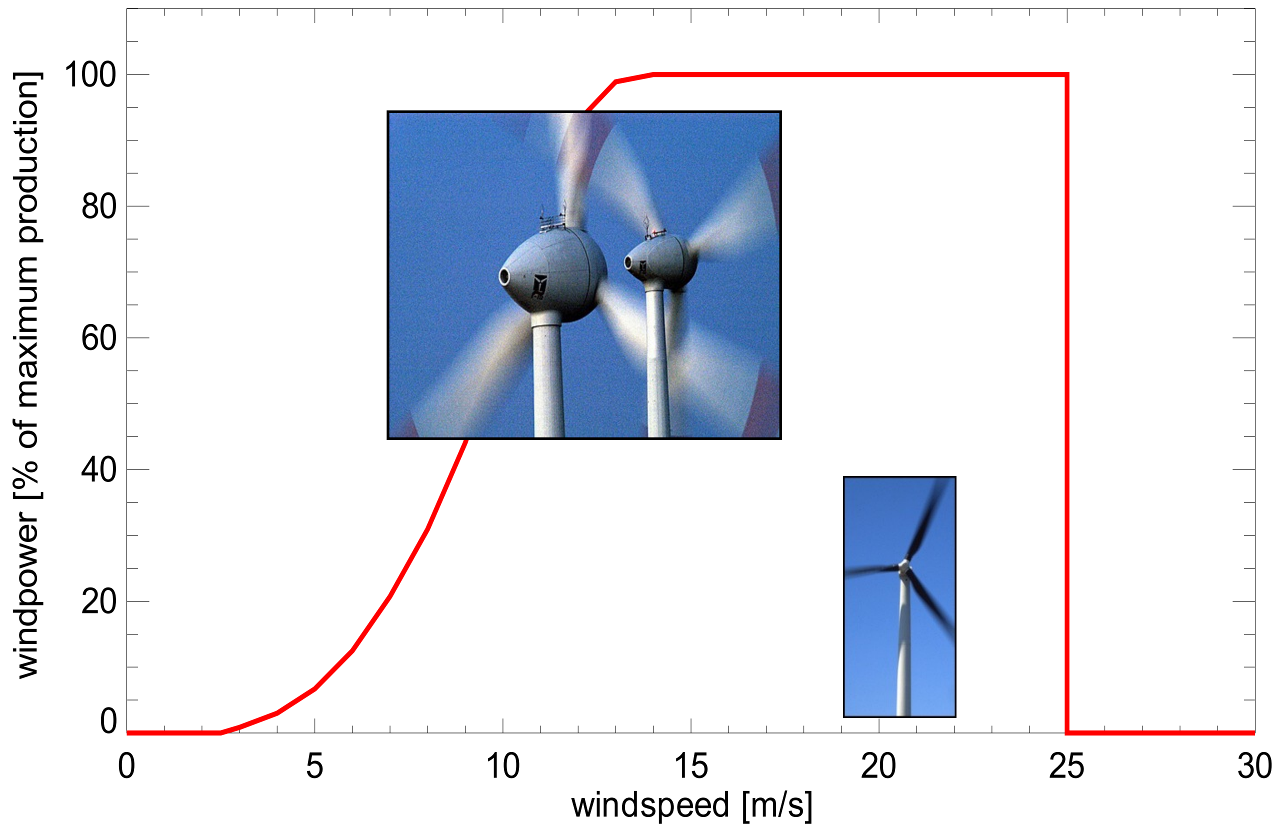
Wind farm example

turbines must be stopped in high winds

Must continue to supply electricity even if not generating

So may need to buy extra energy

Cheaper to buy in advance



Decision to make:

Should I buy extra energy to protect against ff>25 m/s, yes or no?

Financial costs to wind farm manager

	event occurs i.e. ff ≥ 25 m/s	event does NOT occur i.e. ff < 25 m/s
Protection: YES	**200 €**	**200 €**
Protection: NO	**1000 €**	**0 €**

Value of deterministic forecasts

- If no forecast just use climatological information
  - Always protect (if often occurs)
  - Never protect (if rarely occurs)
- Using forecast: protect when event is forecast
  - Can save money compared to using climate
- Value

$$V = \frac{\text{saving from using forecast}}{\text{saving from perfect forecast}}$$

- V = 0 forecast is no better than climate
- V = 1 forecast is perfect (no misses, no false alarms)

Value of deterministic forecast

Protect when event is forecast

Value of using forecast = saving compared to not using forecast

	event occurs i.e. ff ≥ 25 m/s	event does NOT occur i.e. ff < 25 m/s
Forecast: YES Protect: YES	Hit Cost = 200 €	False alarm Cost = 200 €
Forecast: NO Protect: NO	Miss Loss = 1000 €	Correct reject 0 €

Value, forecast quality and the user

Value can be written in terms of hit rate (H), false alarm rate (F) and the "cost-loss ratio" of the user (C/L):

$$V = (1-F) - \left(\frac{1-C/L}{C/L}\right)\left(\frac{\bar{o}}{1-\bar{o}}\right)(1-H) \quad \text{if} \quad C/L < \bar{o}$$

$$V = H - \left(\frac{C/L}{1-C/L}\right)\left(\frac{1-\bar{o}}{\bar{o}}\right)F \quad \text{if} \quad C/L > \bar{o}$$

- **Value depends on forecast quality: H and F**
- **but value also depends on the user (C/L)**
- **and on the weather event (ō)**

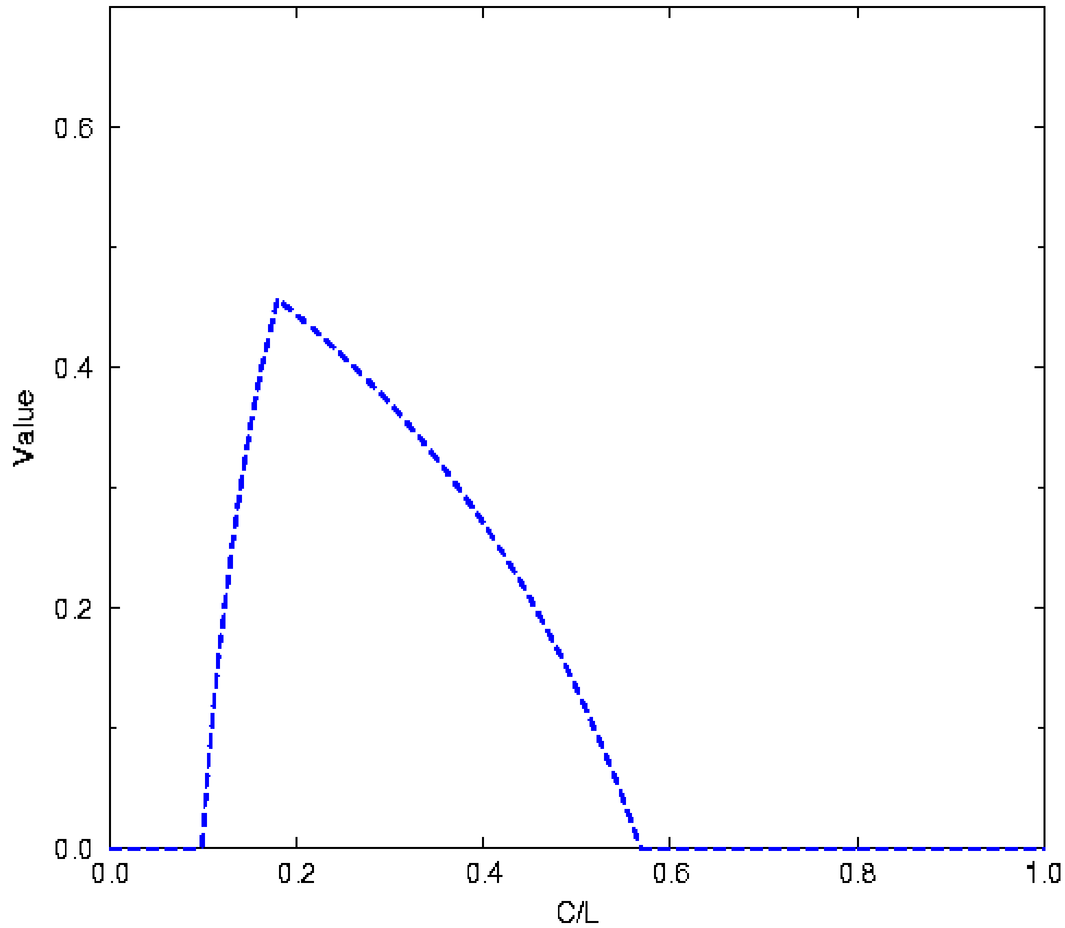
Cost-loss wind farm manager

Cost-loss ratio = 200/1000

= 0.2

	event occurs i.e. ff ≥ 25 m/s	event does NOT occur i.e. ff < 25 m/s
Protection: YES	200 €	200 €
Protection: NO	1000 €	0 €

Value for different users



High loss from missed event (hit rate important)

High cost to protect (false alarm rate important)

Value of probability forecasts

- Using a deterministic forecast is straightforward: take action if bad weather is forecast, otherwise do nothing
- What if the forecast is given as a probability of bad weather?
- To make a decision the probability forecast must be converted to a yes/no action
- Choose a probability threshold p_t
  - if $p>p_t$ then take action
  - if $p<p_t$ then do nothing
- Which probability threshold to choose?

Financial costs to wind farm manager

Probability is 30%

30

70

	event occurs i.e. ff ≥ 25 m/s	event does NOT occur i.e. ff < 25 m/s
Protection: YES	200 €	200 €
Protection: NO	1000 € **30,000 €**	0 €

Financial costs to wind farm manager

Probability is 30%

30 

70 

	event occurs i.e. ff ≥ 25 m/s	event does NOT occur i.e. ff < 25 m/s
Protection: YES	200 € **6,000 €**	200 € **14,000 €**
Protection: NO	1000 €	0 €

Financial costs to wind farm manager

Probability is 30%

30 

70 

Better to protect (costs €20000) than not protect (costs €30000)

	event occurs i.e. ff ≥ 25 m/s	event does NOT occur i.e. ff < 25 m/s
Protection: YES	200 € **6,000 €**	200 € **14,000 €**
Protection: NO	1000 € **30,000 €**	0 €

ECMWF

Financial costs to wind farm manager

Probability is 10%

10

90

	event occurs i.e. ff ≥ 25 m/s	event does NOT occur i.e. ff < 25 m/s
Protection: YES	200 €	200 €
Protection: NO	1000 € 10,000 €	0 €

ECMWF

Financial costs to wind farm manager

Probability is 10%

10 

90 

	event occurs i.e. ff ≥ 25 m/s	event does NOT occur i.e. ff < 25 m/s
Protection: YES	200 € **2,000 €**	200 € **18,000 €**
Protection: NO	1000 €	0 €

Financial costs to wind farm manager

Probability is 10%

10

90

Better to NOT protect (costs €10000) than protect (costs €20000)

	event occurs i.e. ff ≥ 25 m/s	event does NOT occur i.e. ff < 25 m/s
Protection: YES	200 € **2,000 €**	200 € **18,000 €**
Protection: NO	1000 € **10,000 €**	0 €

Financial costs to wind farm manager

Probability is 20%

20 

80 

	event occurs i.e. ff ≥ 25 m/s	event does NOT occur i.e. ff < 25 m/s
Protection: YES	200 €	200 €
Protection: NO	1000 € **20,000 €**	0 €

Financial costs to wind farm manager

Probability is 20%

20 

80 

	event occurs i.e. ff ≥ 25 m/s	event does NOT occur i.e. ff < 25 m/s
Protection: YES	200 € **4,000 €**	200 € **16,000 €**
Protection: NO	1000 €	0 €

Financial costs to wind farm manager

Probability is 20%

20

80

Same to protect as not protect (€20000)

	event occurs i.e. ff ≥ 25 m/s	event does NOT occur i.e. ff < 25 m/s
Protection: YES	200 € **2,000 €**	200 € **18,000 €**
Protection: NO	1000 € **20,000 €**	0 €

Probability threshold depends on user

- If the cost of protection is expensive wait until event is more certain (higher probability)
  - False alarms are more important
- If the loss is greater then protect even at low probability
  - Missed events are more important
- The threshold depends on the user: $p_t=C/L$

Value of probability and deterministic forecasts compared