Ensemble Forecasts
Initial Perturbation 1

Simon Lang

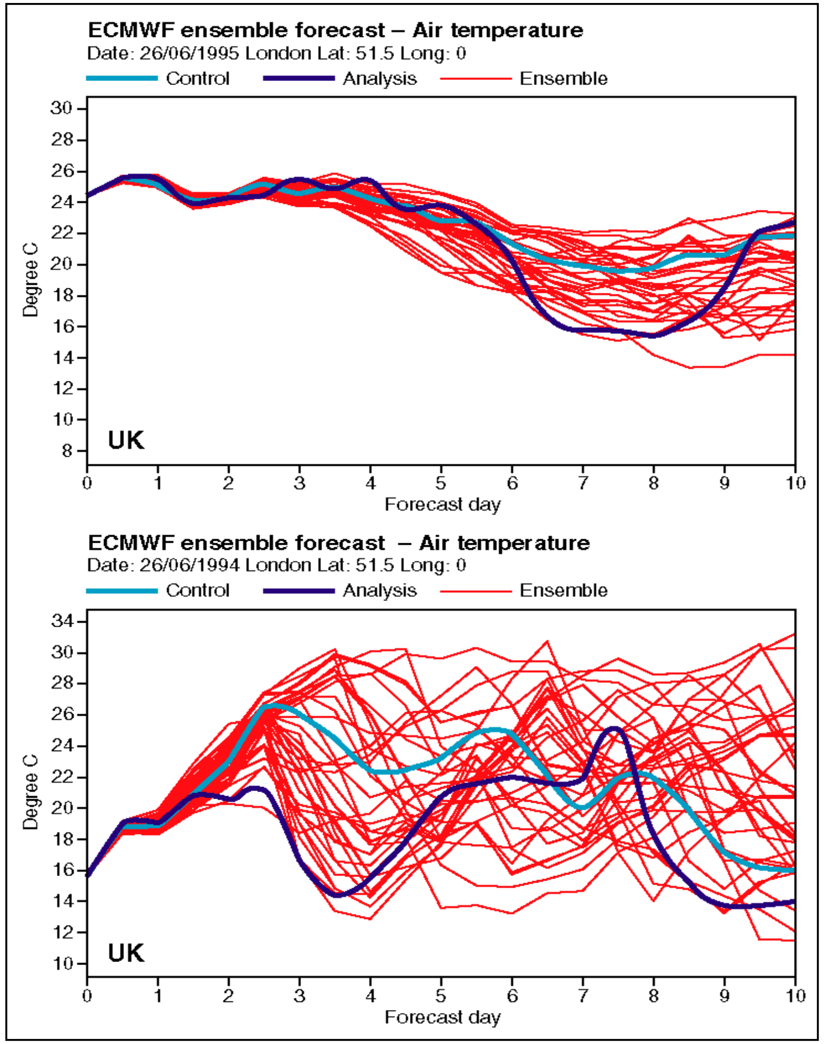
Chaos and weather prediction

The atmosphere is a chaotic system

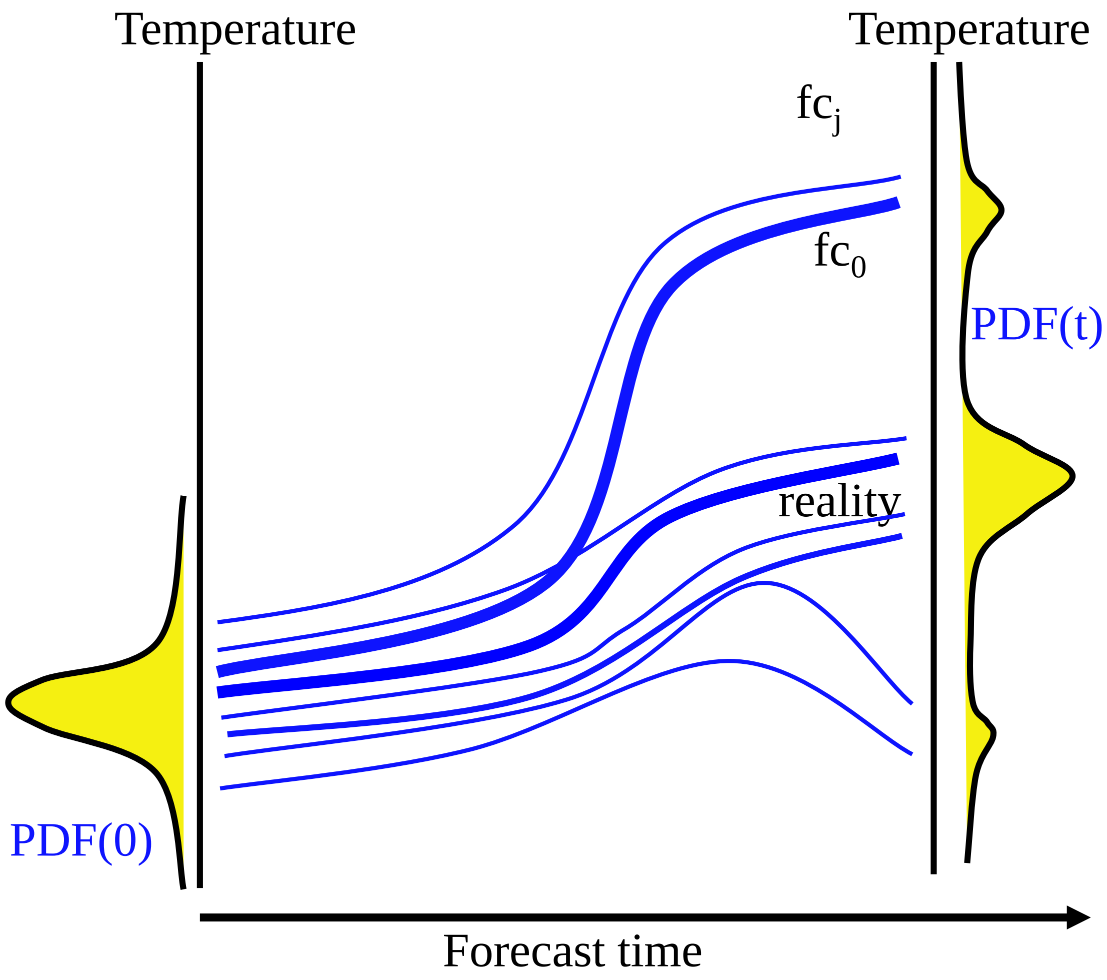
- Small errors can grow to have major impact
- We can never perfectly measure the current state of the whole atmosphere

Ensemble Forecasts

- Parallel set of forecasts from very slightly different initial conditions and model formulation
- Assess uncertainty of today's forecast



- 51 Members (50 perturbed + control member without perturbations), TCo639 (~ 18 km) to day 15, then TCo319 (~ 36 km)
- 91 vertical levels
- Coupled to NEMO ocean model (1/4 degree) and LIM2 ice model
- Initial perturbation via an ensemble of data assimilations and singular vectors, 5 member ocean data assimilation
- Model error representation - SPPT



from R. Buizza

Sources of Uncertainty:

- Initial Conditions
- Model Formulation

Perturbations to the initial conditions:

Methods that rely on the dynamics only, e.g.:

- bred vectors
- singular vectors

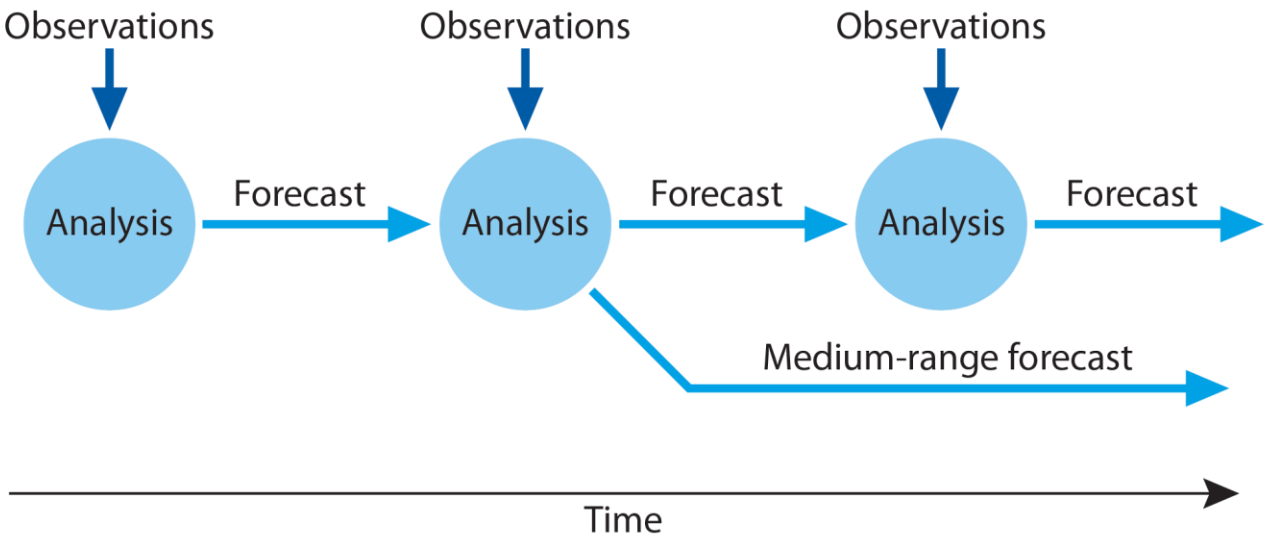
Ensemble data assimilation methods, e.g.:

- Ensemble of 4D-Var data assimilations (EDA)
- Ensemble Kalman Filter

ECMWF: combination of EDA and singular vectors

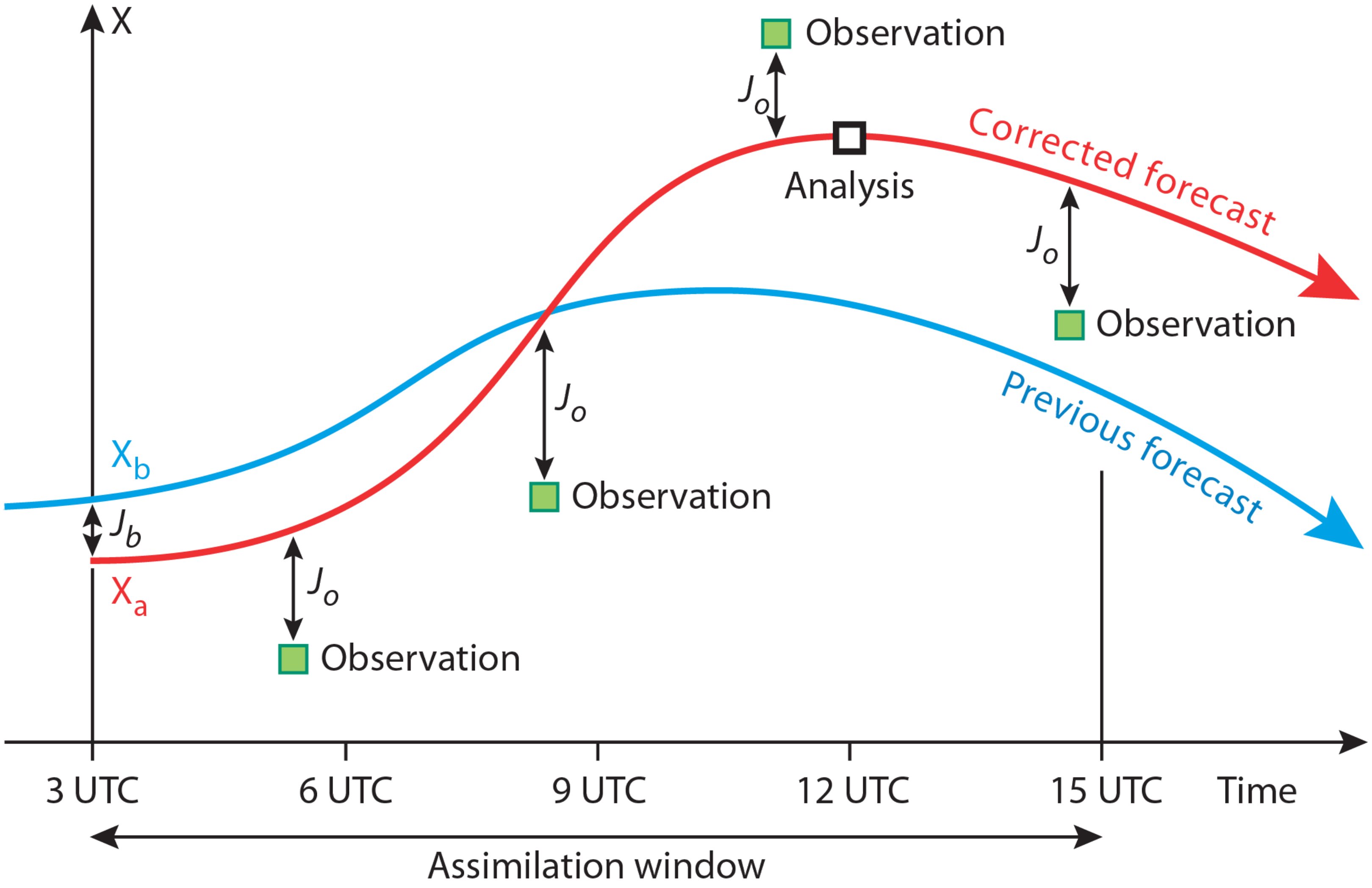
Starting the Medium-Range Forecast – the 'Analysis'

Analysis: 3 dimensional virtual image of the atmosphere at a given time.



- The short range forecast from the previous analysis is our 'first estimate' of the current state of the atmosphere.

4D-Var assimilation

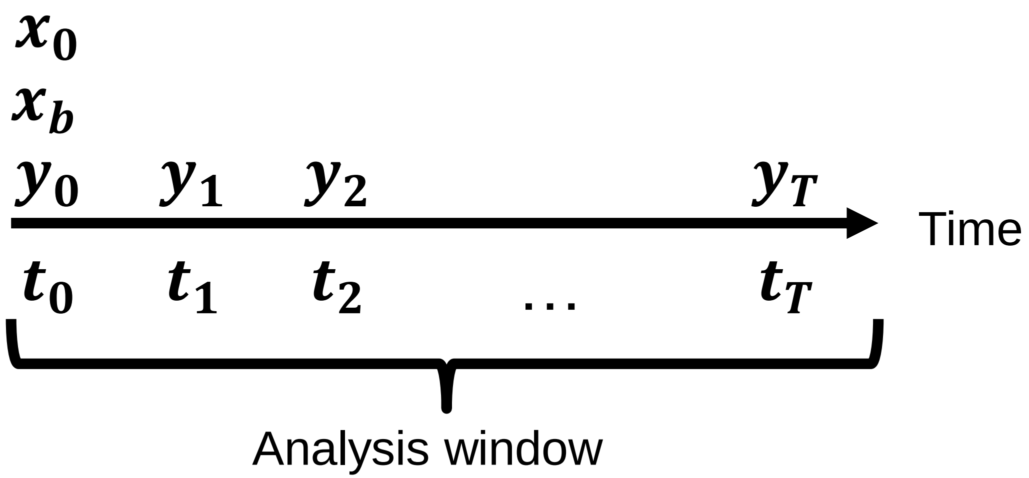


ECMWF

4D-Var assimilation

To find model trajectory that best fits the observations over an assimilation interval (t=0,1,…,T) - > finding the minimum of the 4DVar cost function:

$$J(\mathbf{x}_0)=(\mathbf{x}_b-\mathbf{x}_o)^T(\mathbf{P}^b)^{-1}(\mathbf{x}_b-\mathbf{x}_o)+\sum_{t=0}^{T}(\mathbf{y}_t-H_tM_{0\to t}(\mathbf{x}_0))^T\,\mathbf{R}_t^{-1}(\mathbf{y}_t-H_tM_{0\to t}(\mathbf{x}_0))$$



See lectures in DA Training

Ensemble of 4D-Var data assimilations (EDA)

- 25 perturbed ensemble members + 1 control:
  TCo639 outer loops (~ 18 km), 137 levels, TL191/TL191 inner loops.
  (HRES DA: TCo1279 outer loops (~ 9 km), TL255/TL319/TL399 inner loops).

- Observations randomly perturbed according to their estimated error covariances (R)

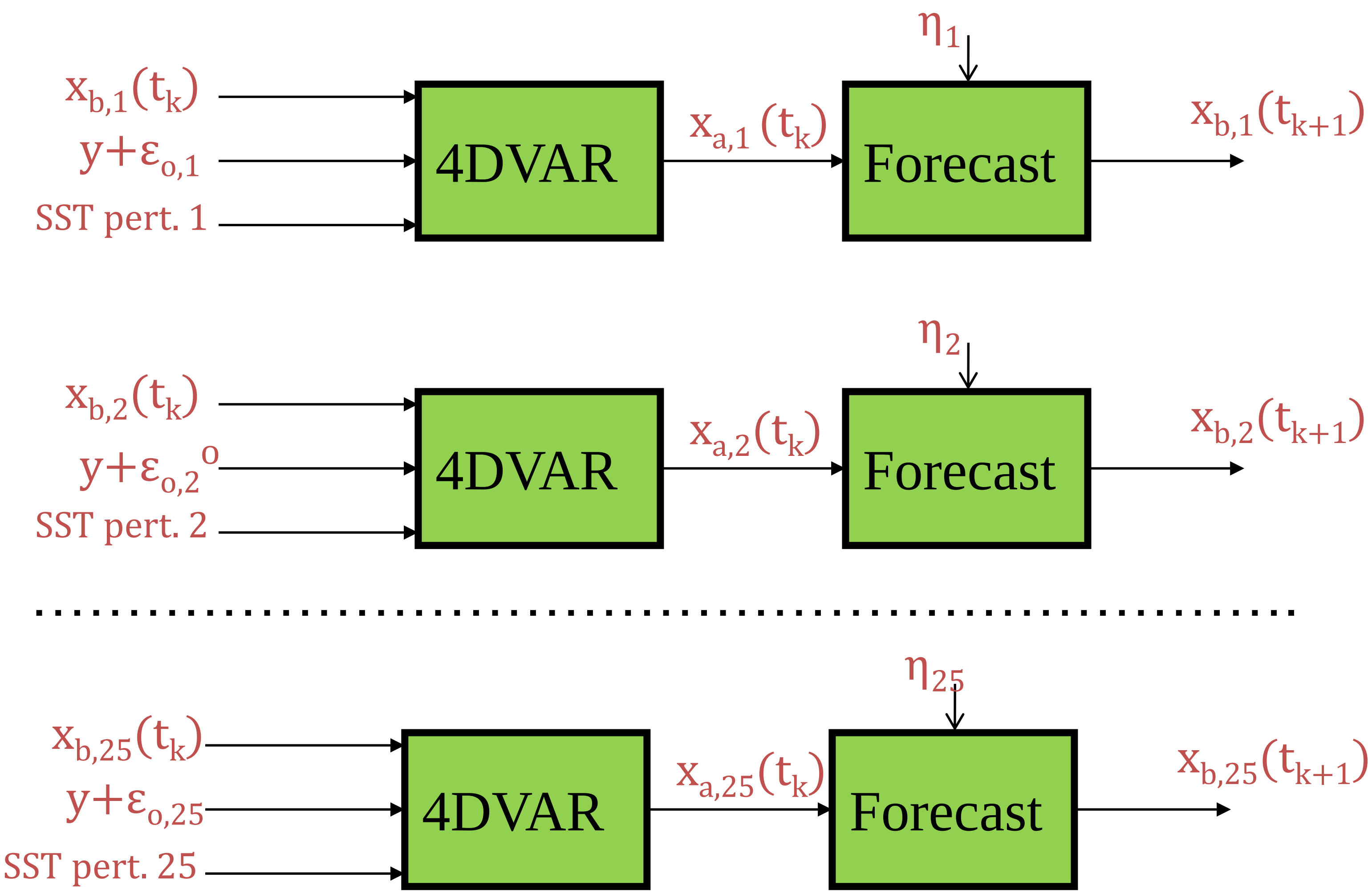
- SST perturbed with climatological error structures

- Model error representation via Stochastically Perturbed Parametrization Tendencies (SPPT)

<u>The EDA simulates the error evolution of the 4DVar analysis cycle:</u>

→ uncertainty estimates to initialize ensemble forecasts

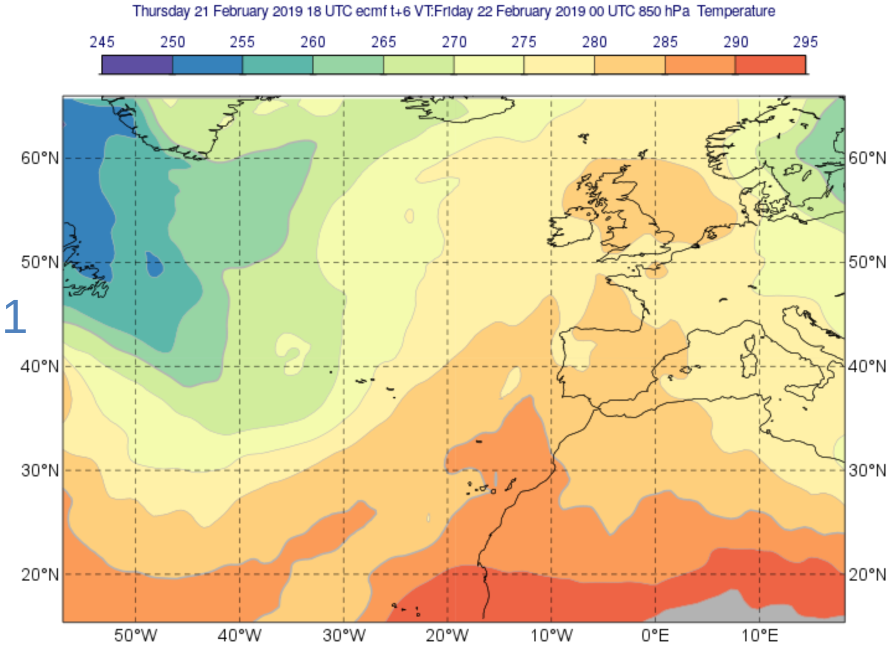
→ Flow dependent estimates of background error covariances for use in 4D-Var



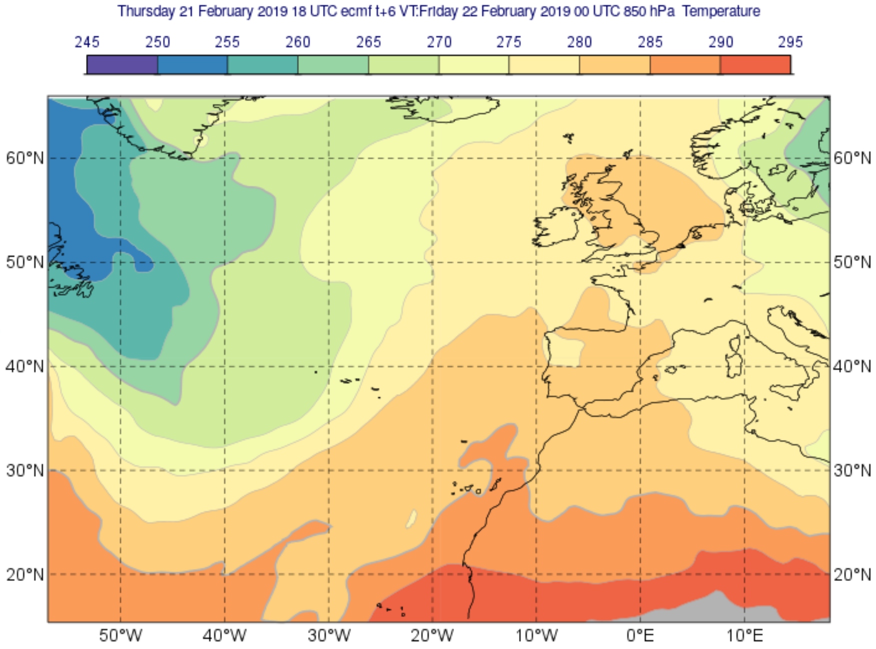
See also Massimo Bonavita's Talk in DA Training

T850hPa

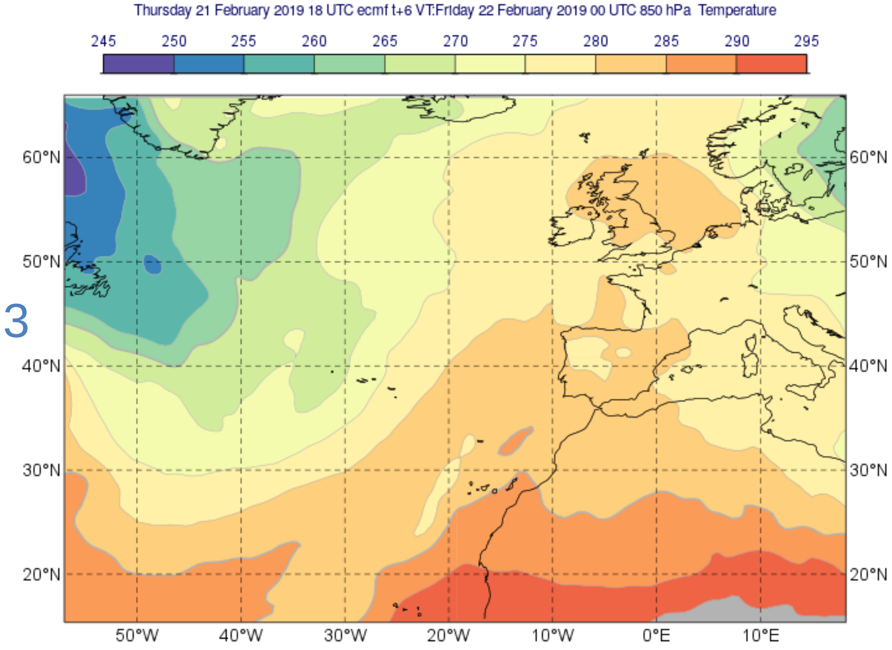
EDA Member 1



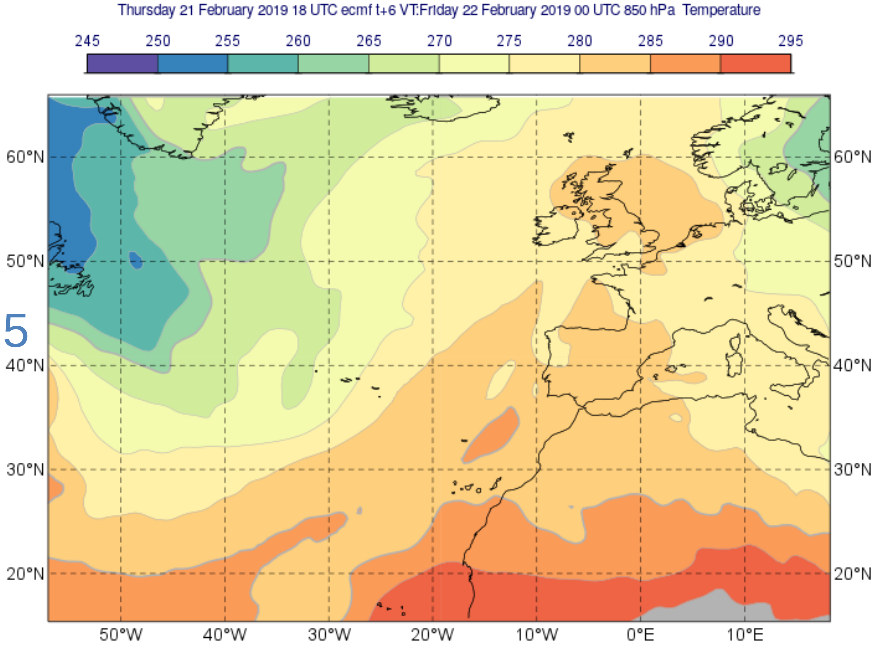
EDA Member 2



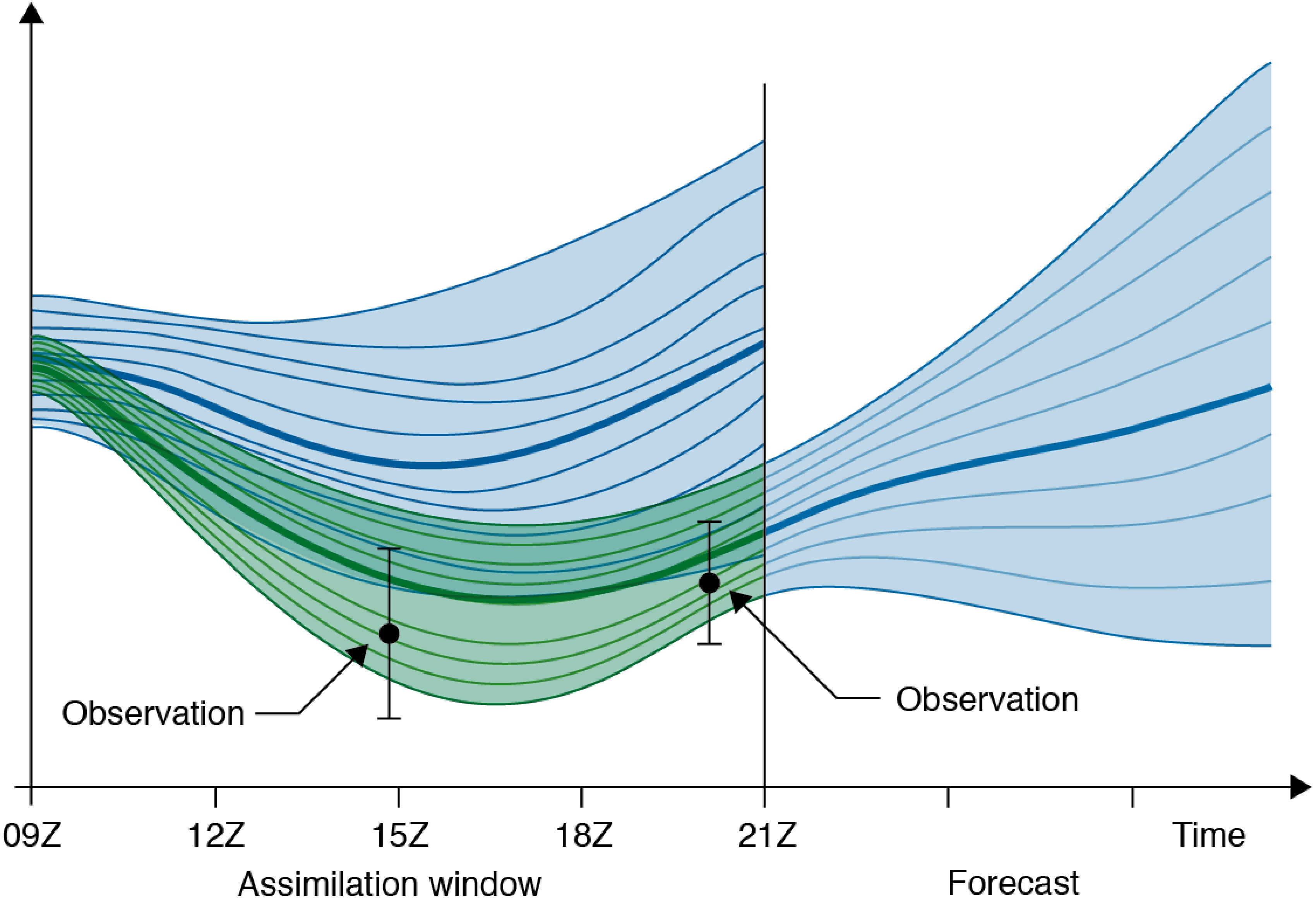
EDA Member 3



EDA Member 25

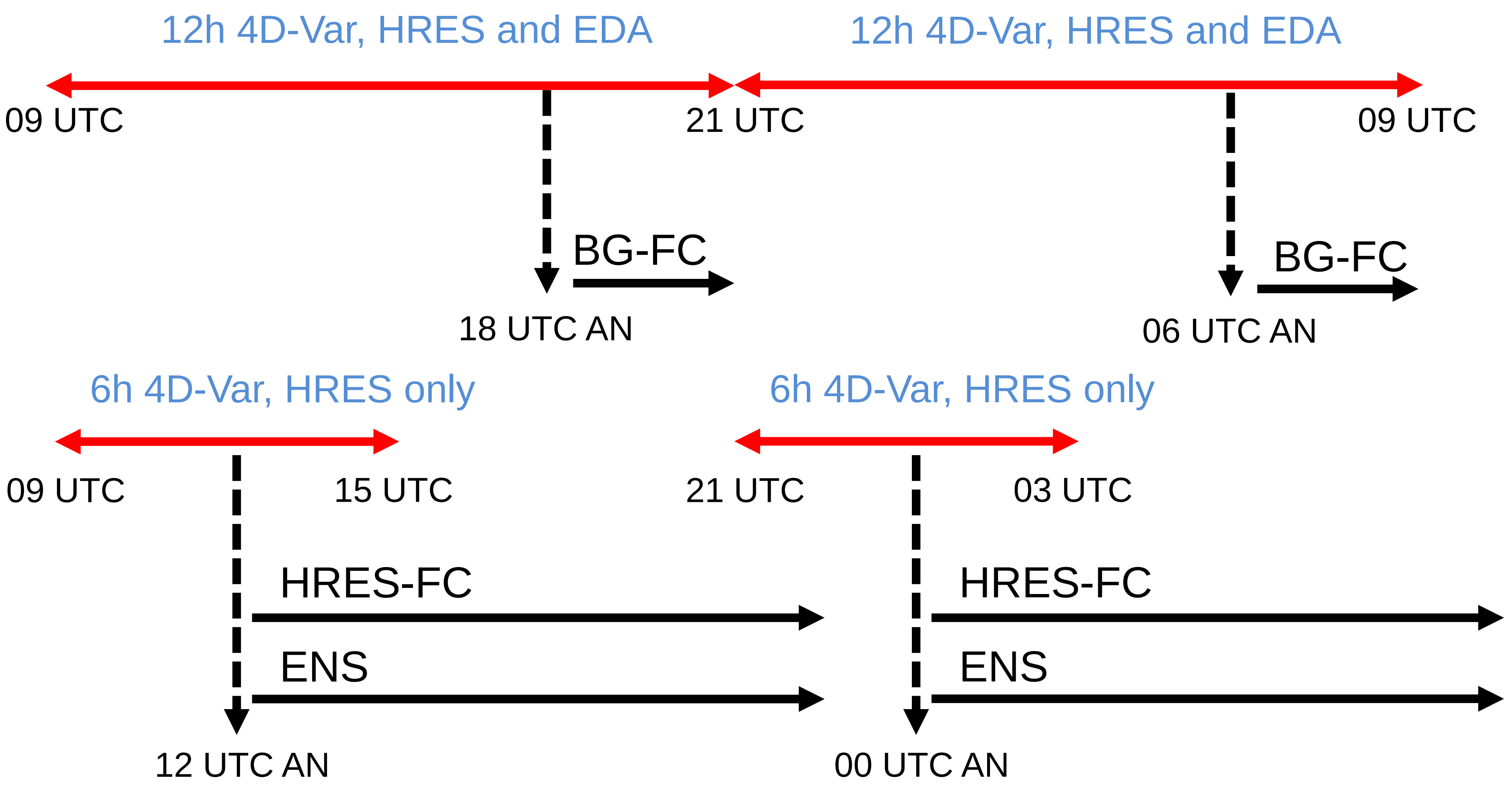


Ensemble Data Assimilation | Ensemble Forecast



ECMWF

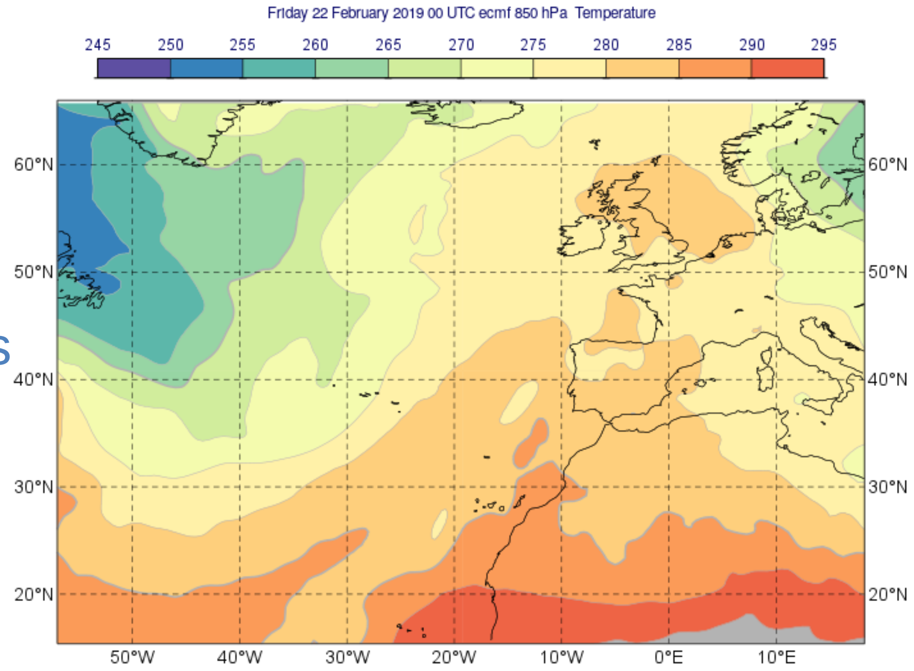
Early delivery Suite:



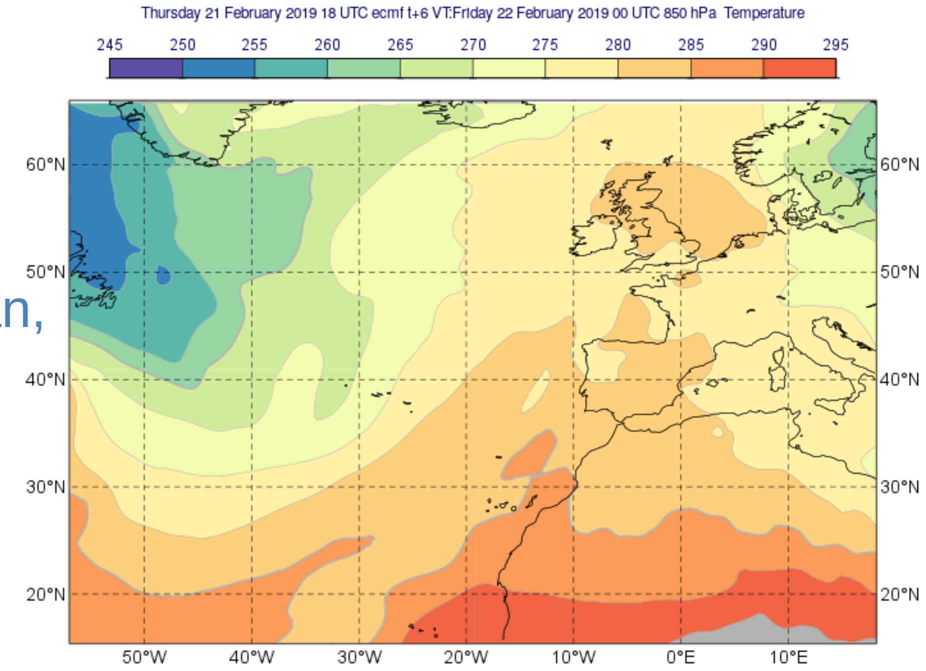
T850hPa

What is available when we start the 00 UTC ensemble forecasts?

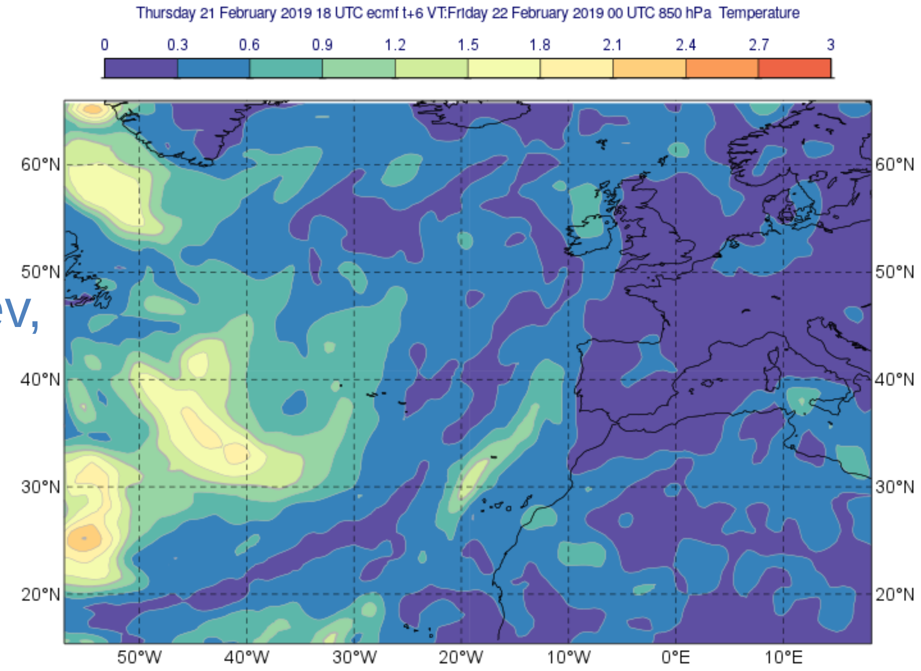
HRES Analysis 00 UTC



EDA Mean, 18 UTC, 6h Fcsts



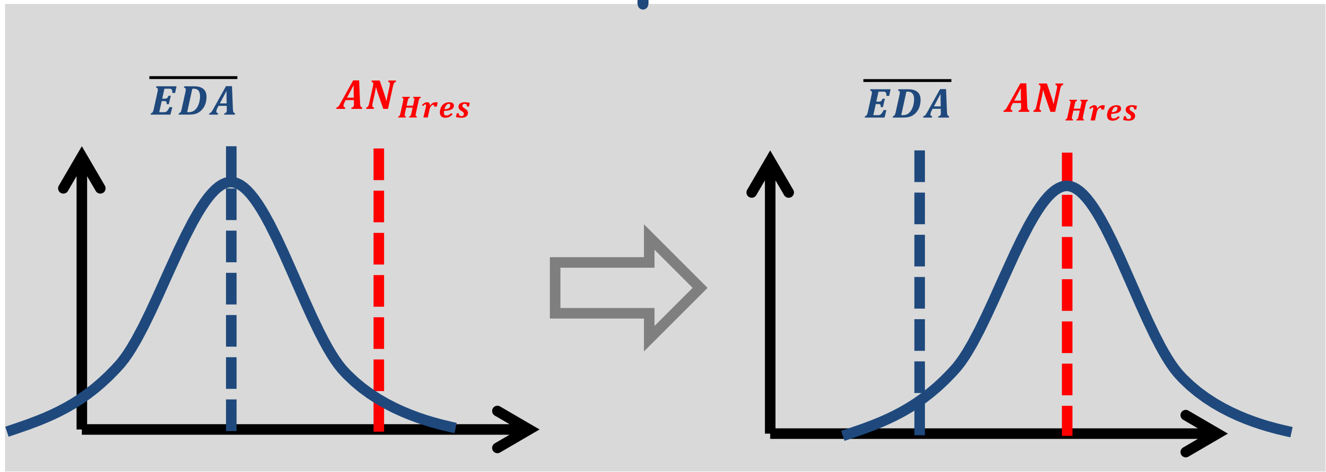
EDA StDev, 18 UTC, 6h Fcsts



Generation of initial conditions for the ensemble forecasts:

$$AN_{pf} = AN_{Hres} \pm \left(EDA_i - \overline{EDA}\right) \pm SVPERT_j$$

$i = 1..25$
$j = 1..25$



Re-centre EDA-Distribution on Hres-Analysis

EDA : 6h Forecasts

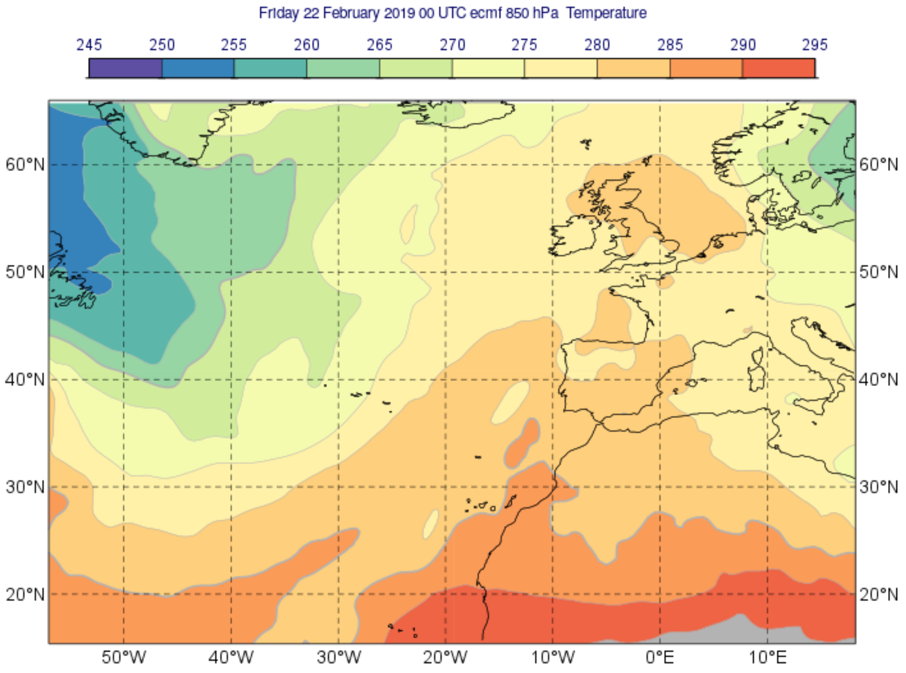
$$SVPERT_j = \sum_{l}^{NSET} \sum_{k}^{NSV_l} \alpha_{lk} SV_{lk}$$

α random number drawn from Truncated gaussian

NSET : nhem, shem, TCs1-6
NSV : 50 for nhem and shem, 5 for TCs

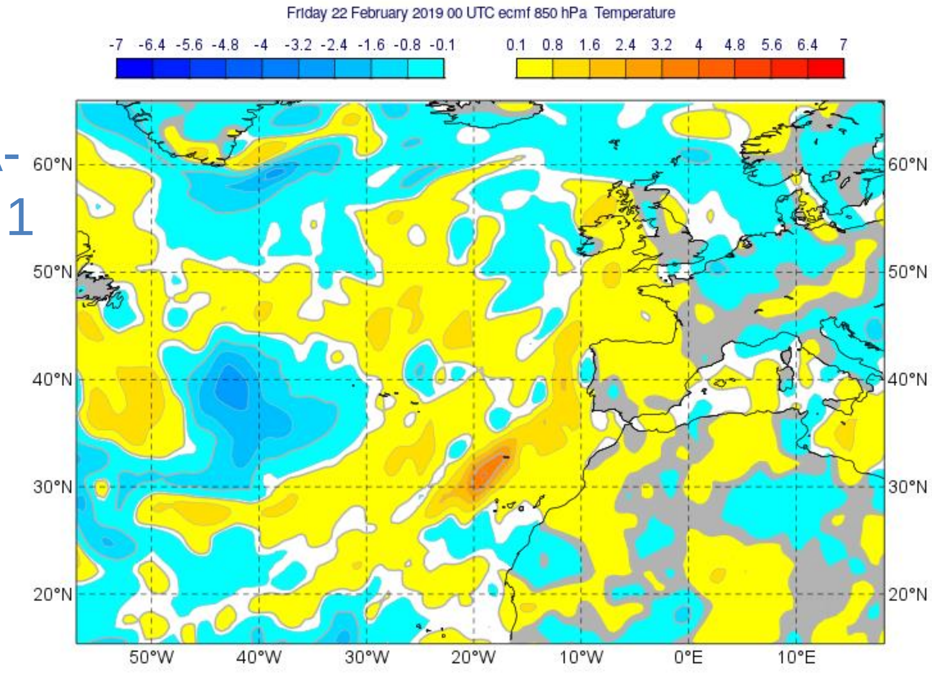
T850hPa

HRES Analysis 00 UTC



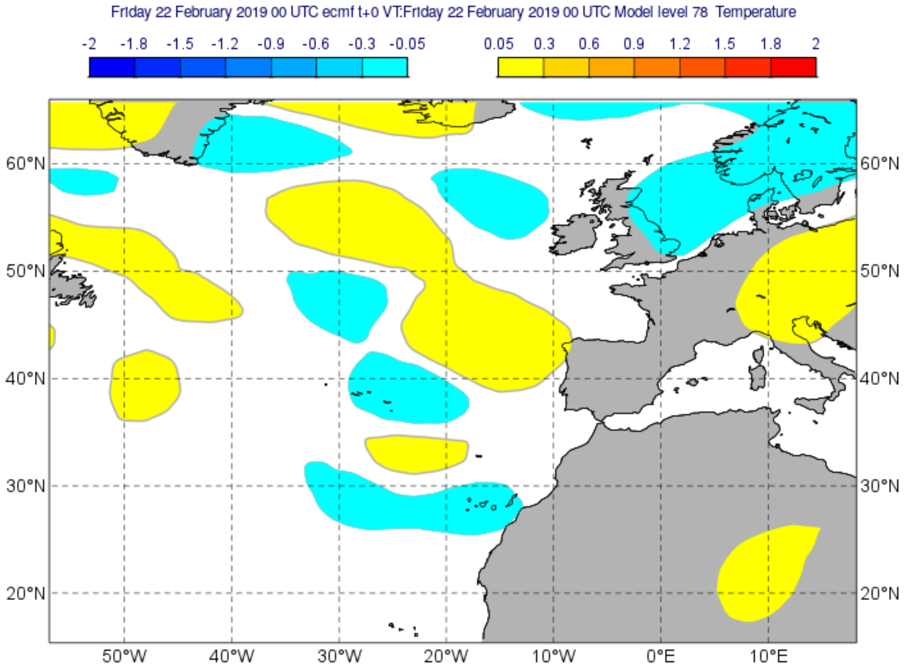
EDA-Pert 1

+



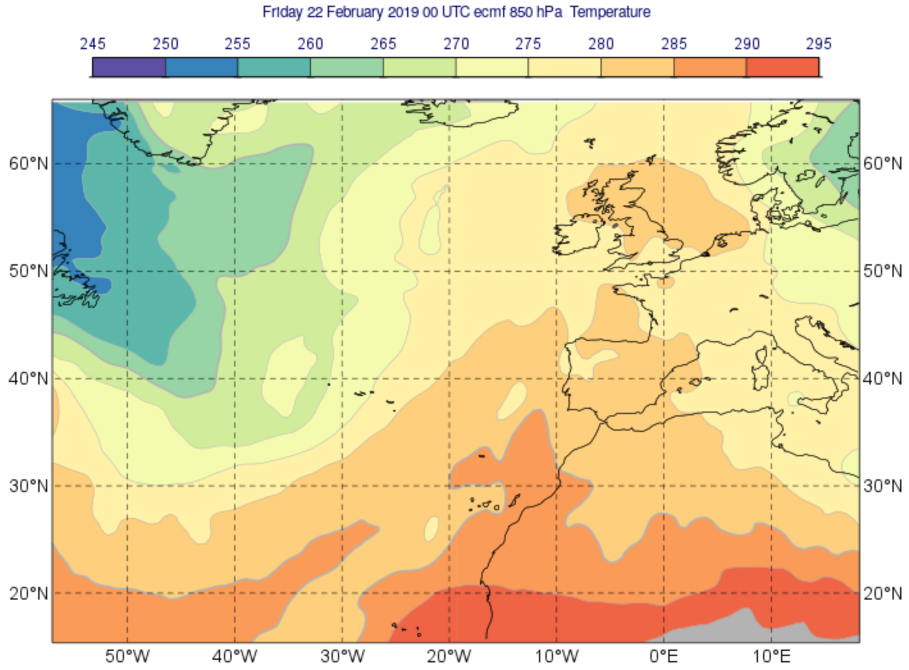
SV-Pert 1

+



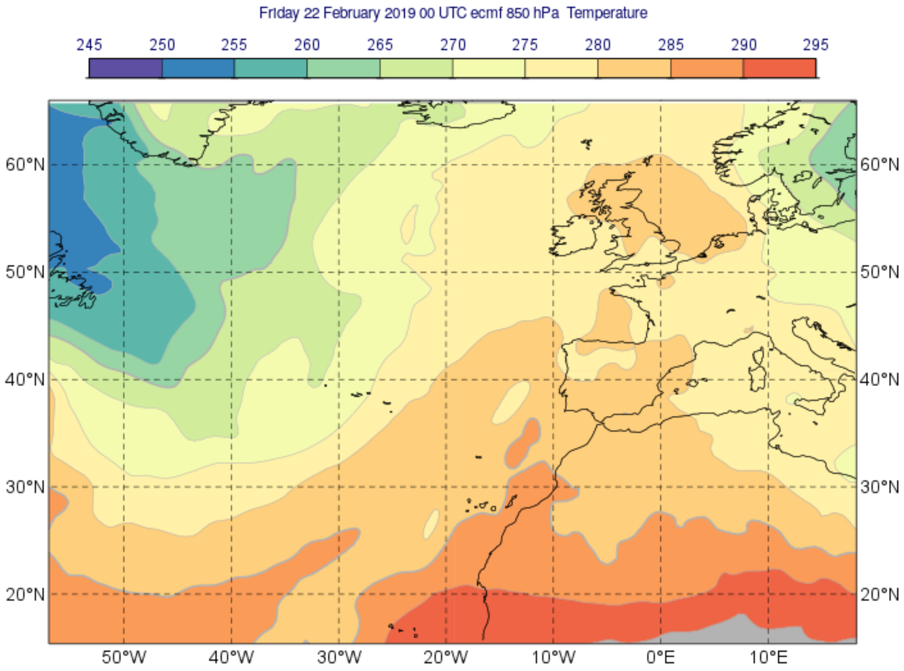
=

Initial conditions for ENS member 1



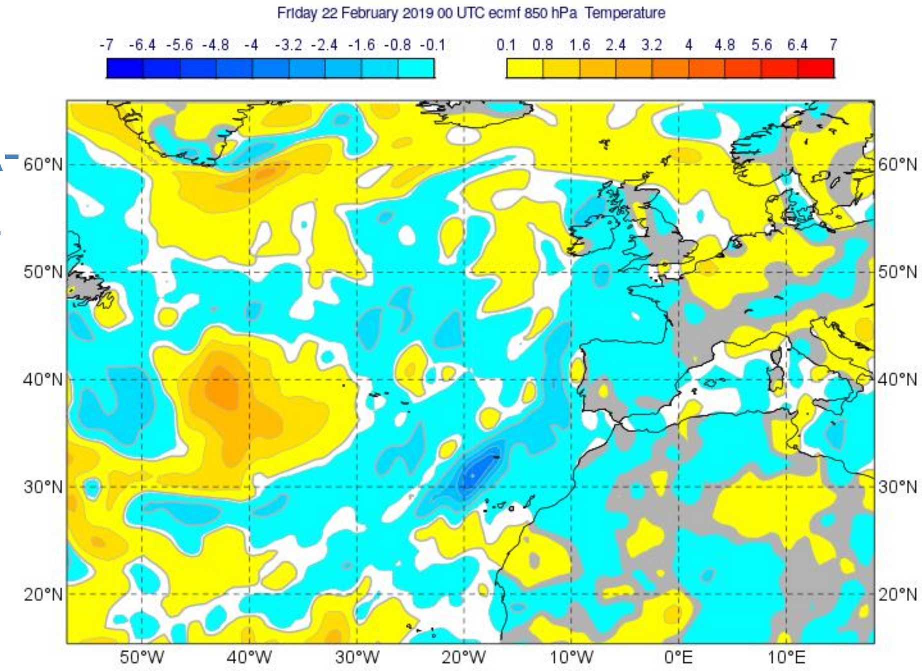
T850hPa

HRES Analysis 00 UTC



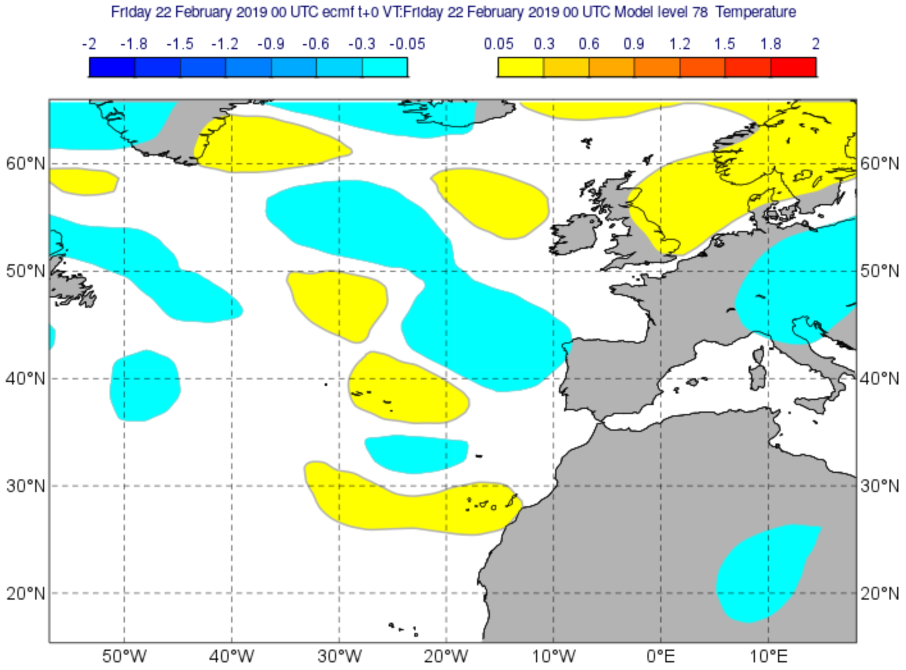
+

- EDA- Pert 1



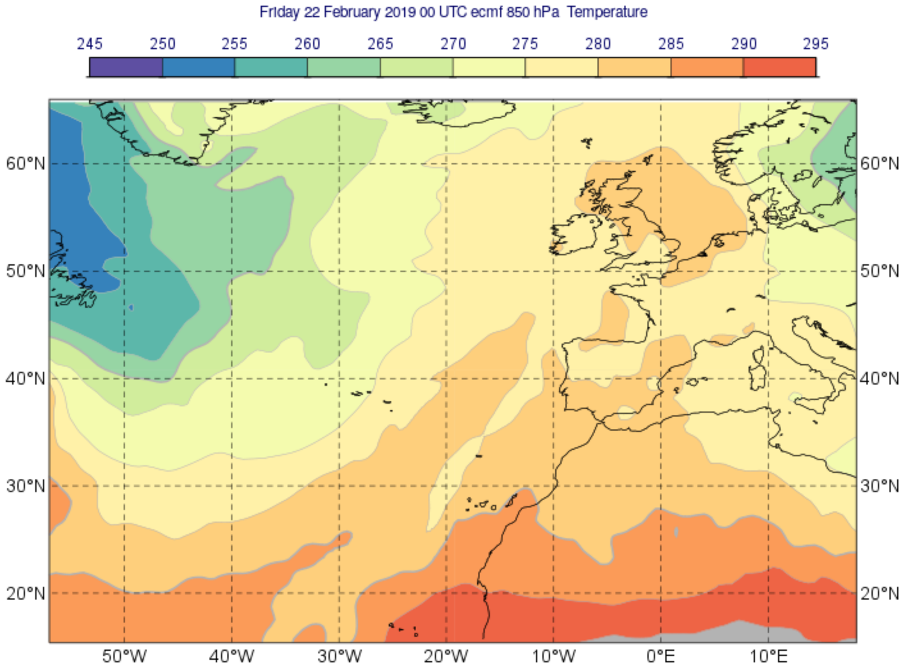
+

- SV- Pert 1



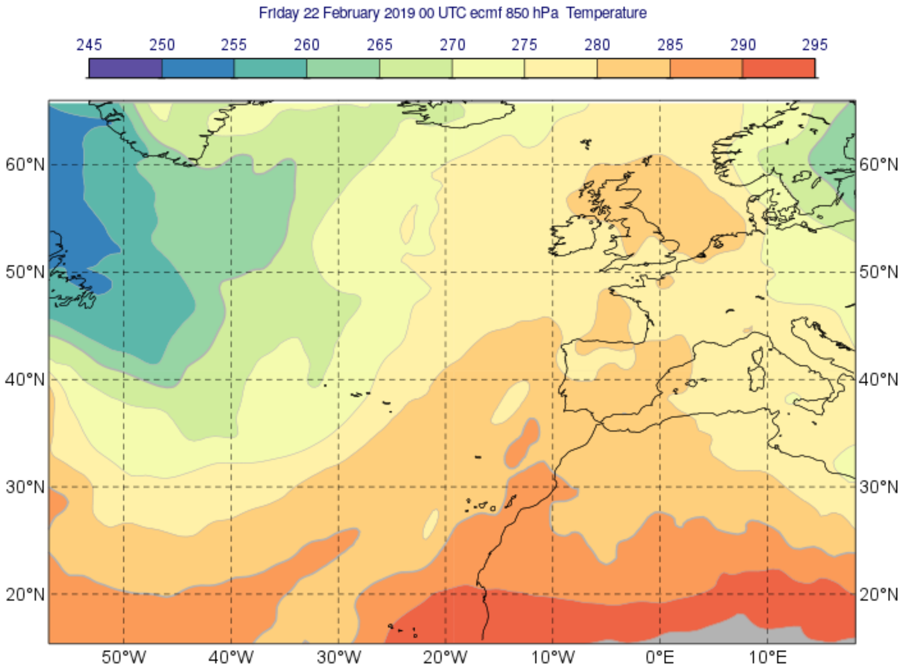
=

Initial conditions for ENS member 2



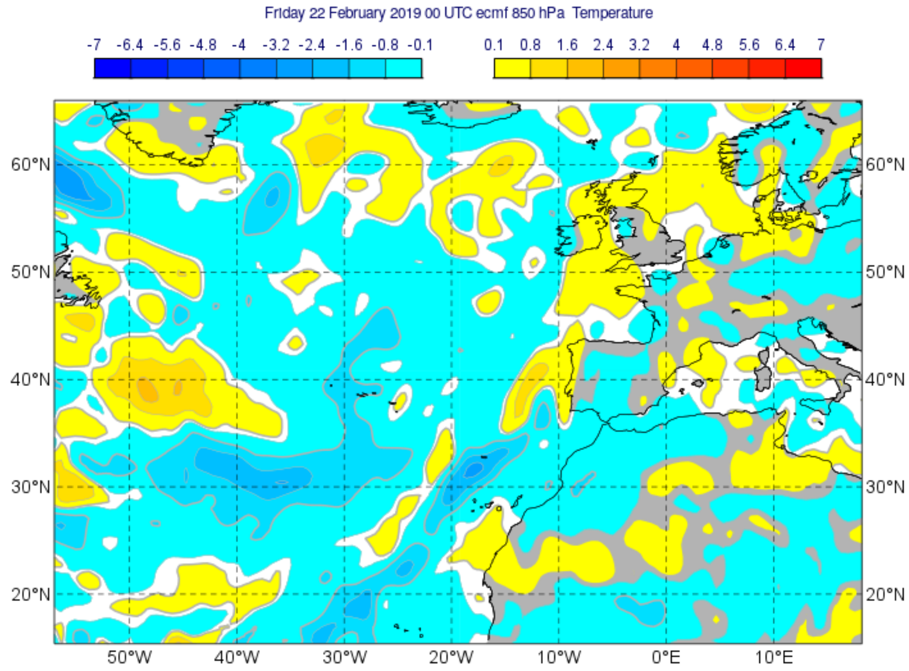
T850hPa

HRES Analysis 00 UTC



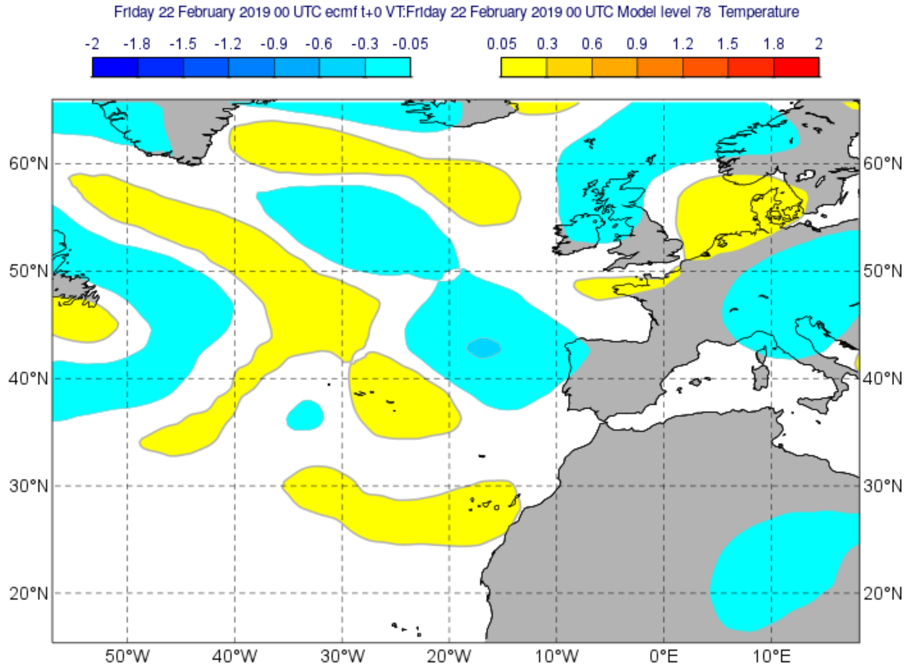
+

EDA-Pert 2



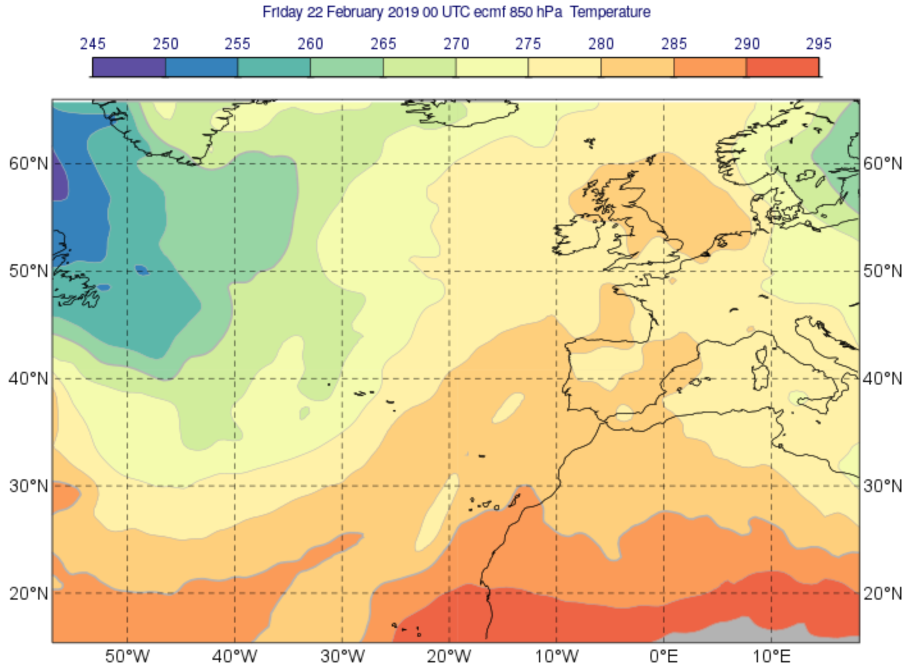
+

SV-Pert 2



=

Initial conditions for ENS member 3

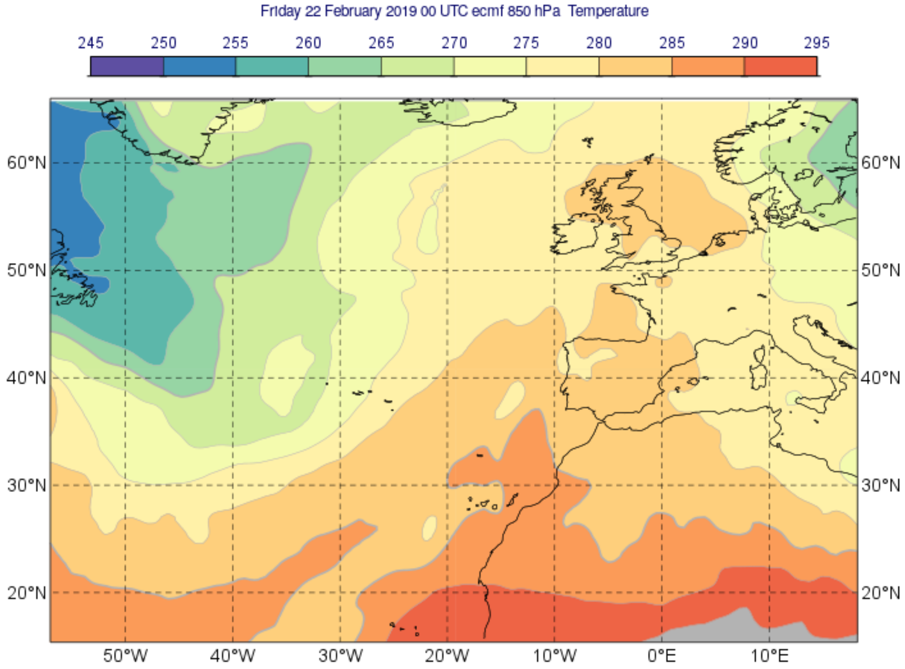


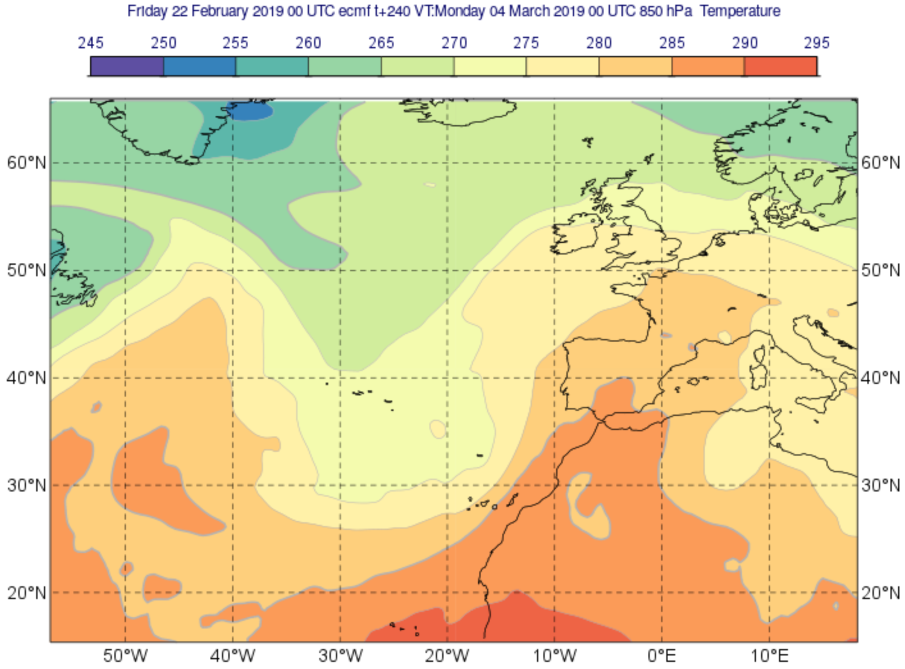
T850hPa

T+0h

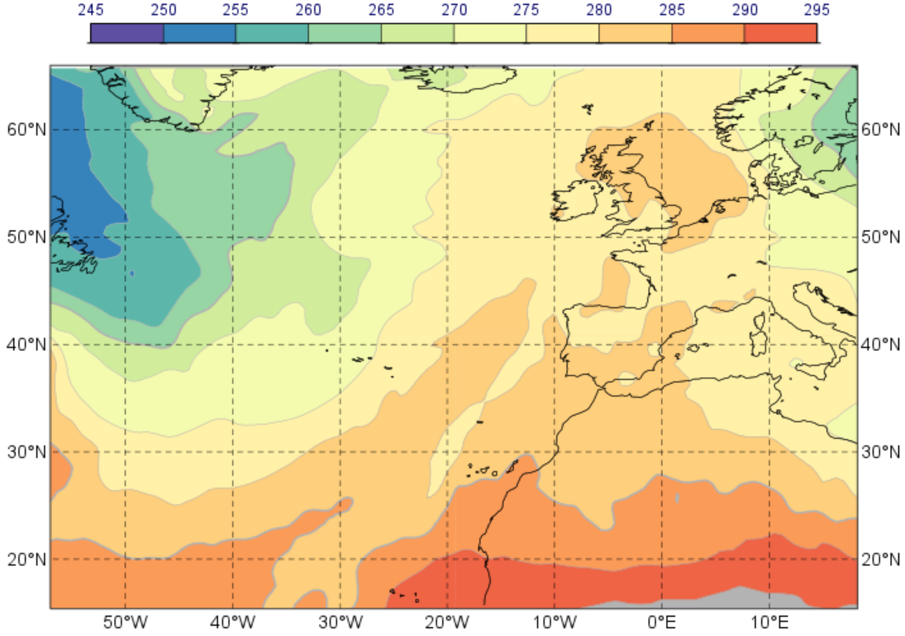
T+240h

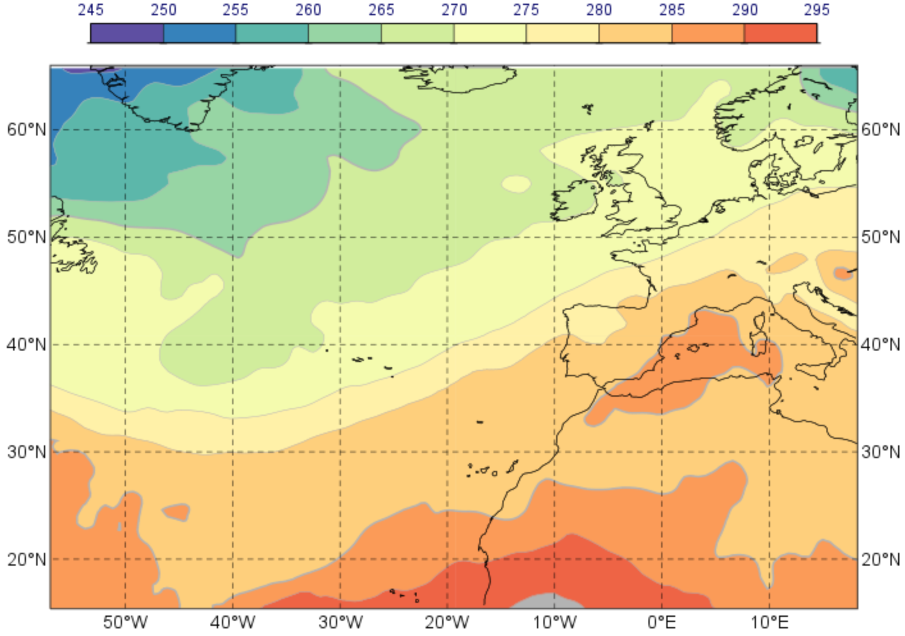
Member 1:





Member 2:



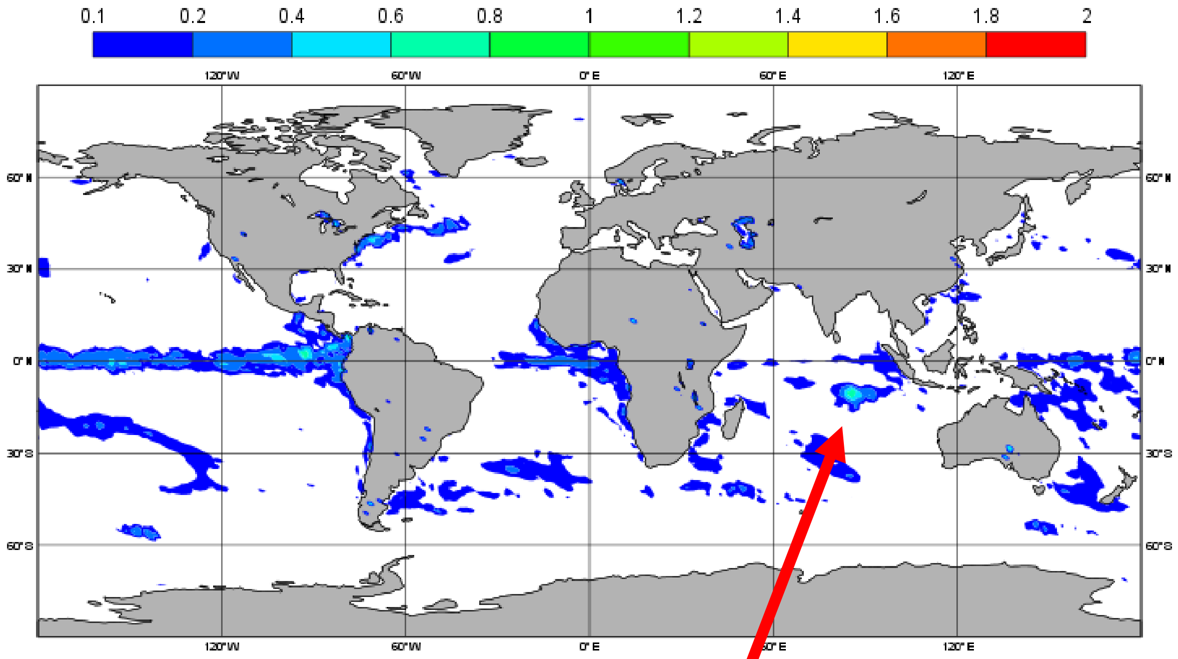


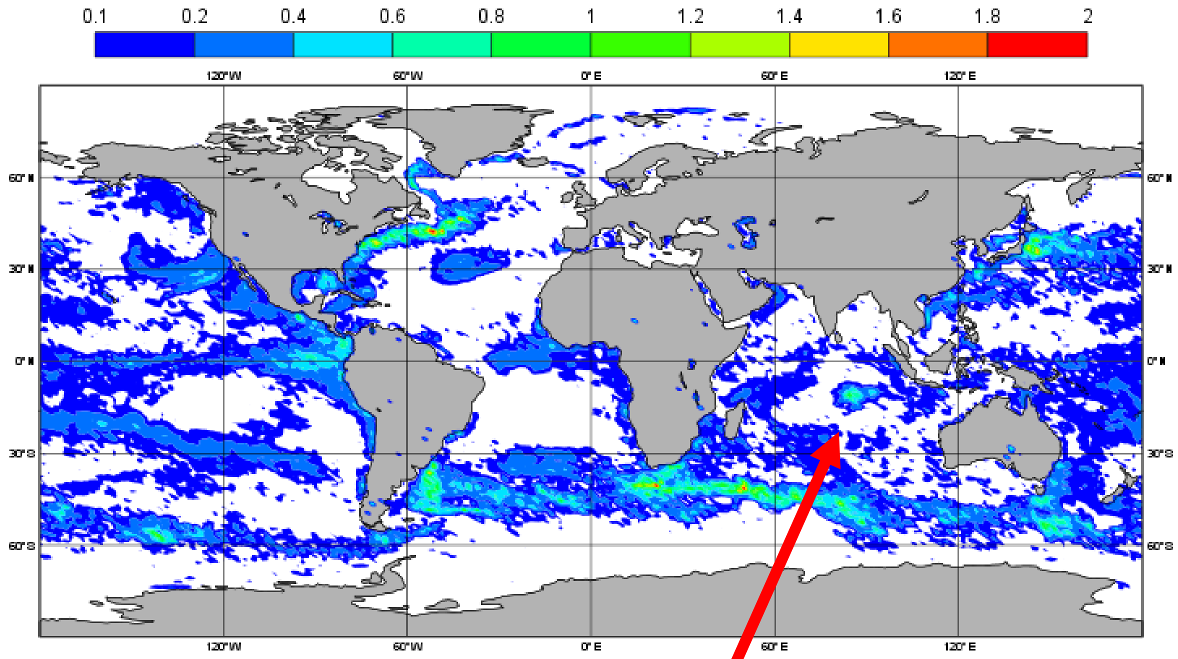
Ocean initial state:

50 Members + 1 Control, 5 Ocean analyses

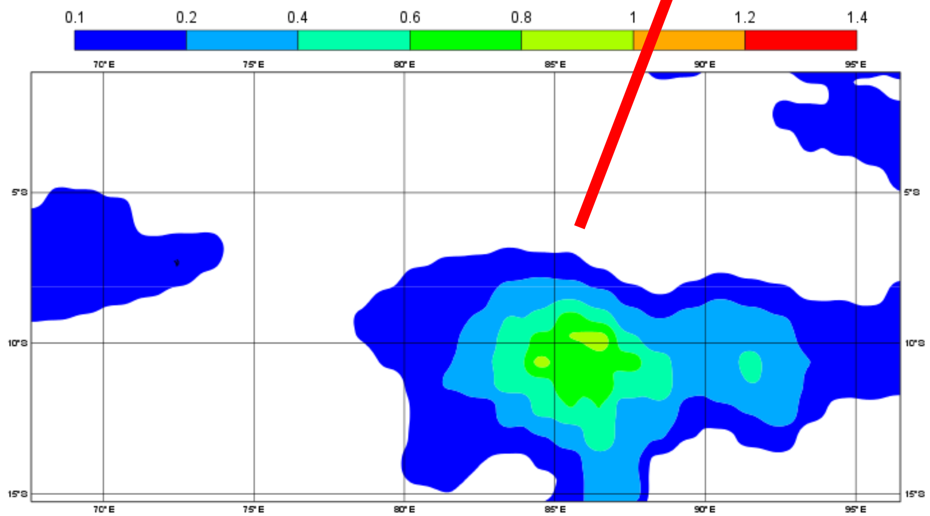
Member	Ocean analysis
Control	1
Member 1	2
Member 2	3
Member 3	4
Member 5	5
Member 6	1
Member 7	2
…	
Member 50	1

10 member, TCO639 realtime system, 1 initial date, SST StDev, fc step 120h



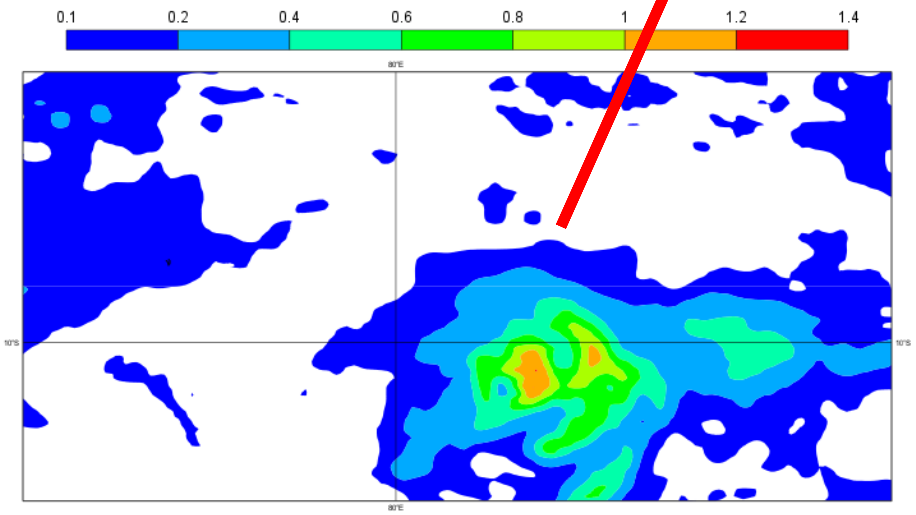


1 deg partial coupling



TC

¼ deg partial coupling



Model Error Representation: SPPT

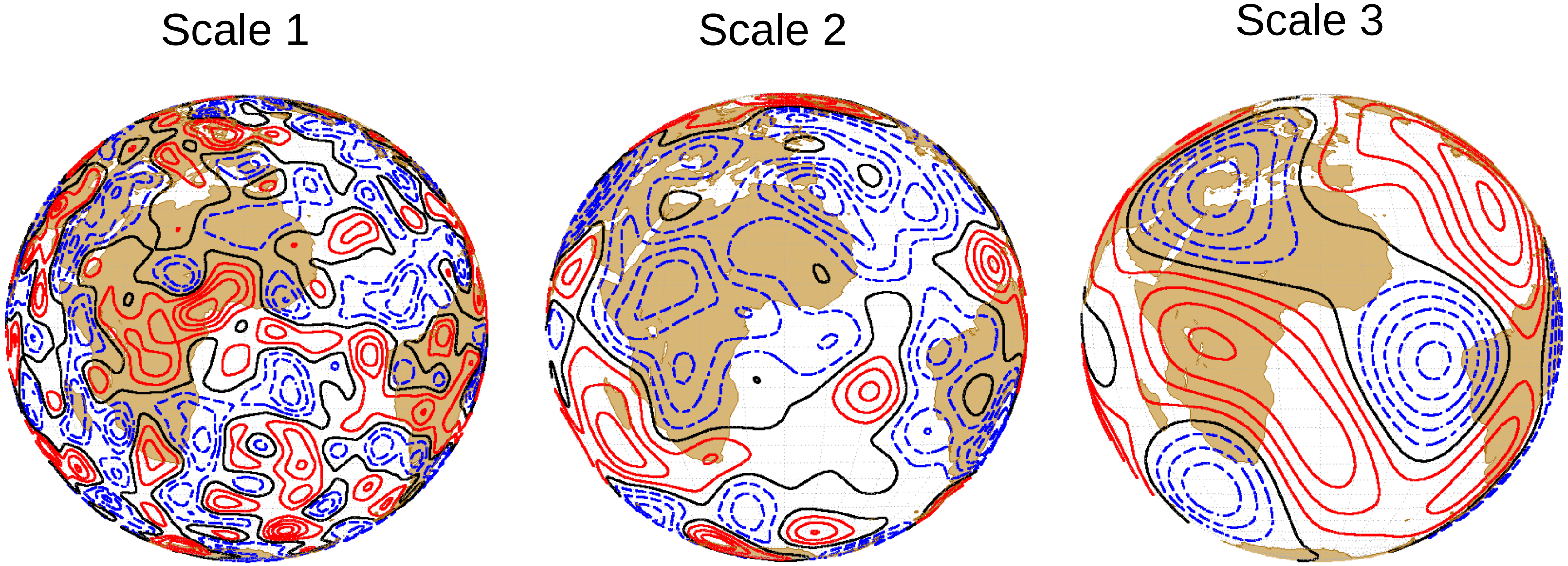
See Leutbecher et al., 2017 for details

Perturb model tendencies during the forecast:

$$\boldsymbol{x_p} = \boldsymbol{x} + \boldsymbol{\alpha x}$$

$\boldsymbol{x}$ sum of tendencies from parametrization schemes (convection, radiation, cloud etc.)

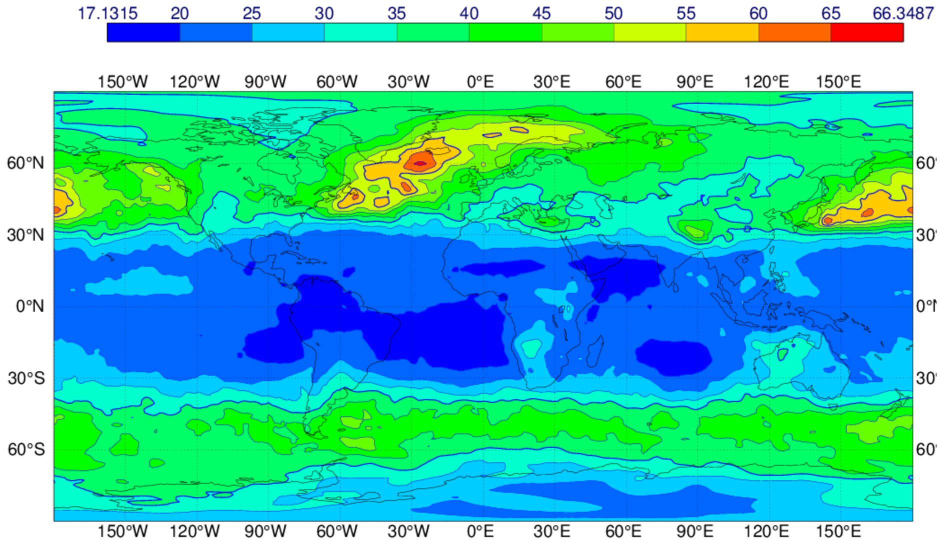
$\boldsymbol{\alpha}$ includes random time and space correlations, provided by a pattern generator



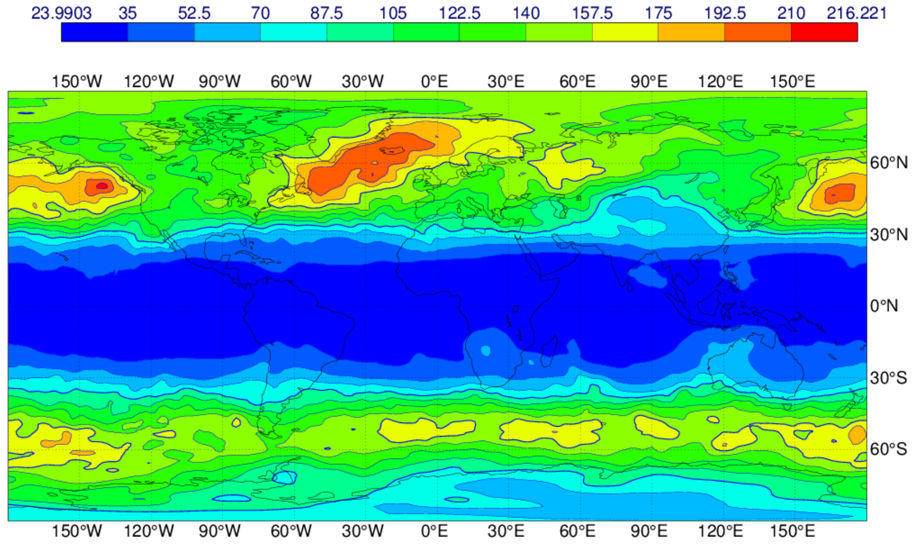
Same model uncertainty representation in ensemble forecasts and ensemble data assimilation

z500 hPa Ensemble StDev, averaged 2016112200 – 2017021300, 00 UTC Run

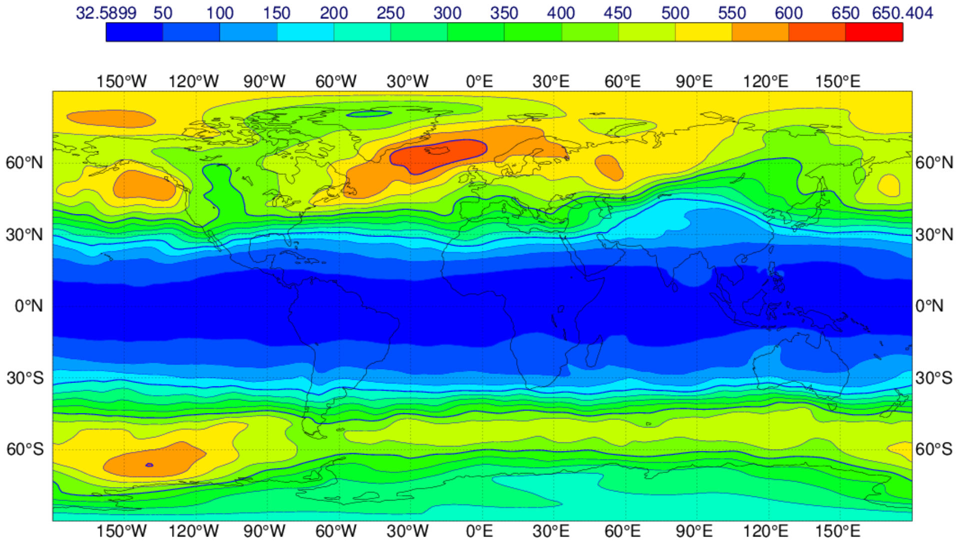
+12h



+48h



+120h



Reliability of the ensemble spread

- Consider ensemble variance ("spread") for an M-member ensemble

$$\frac{1}{M}\sum_{j=1}^{M}(x_j - \overline{x})^2$$

and the squared error of the ensemble mean

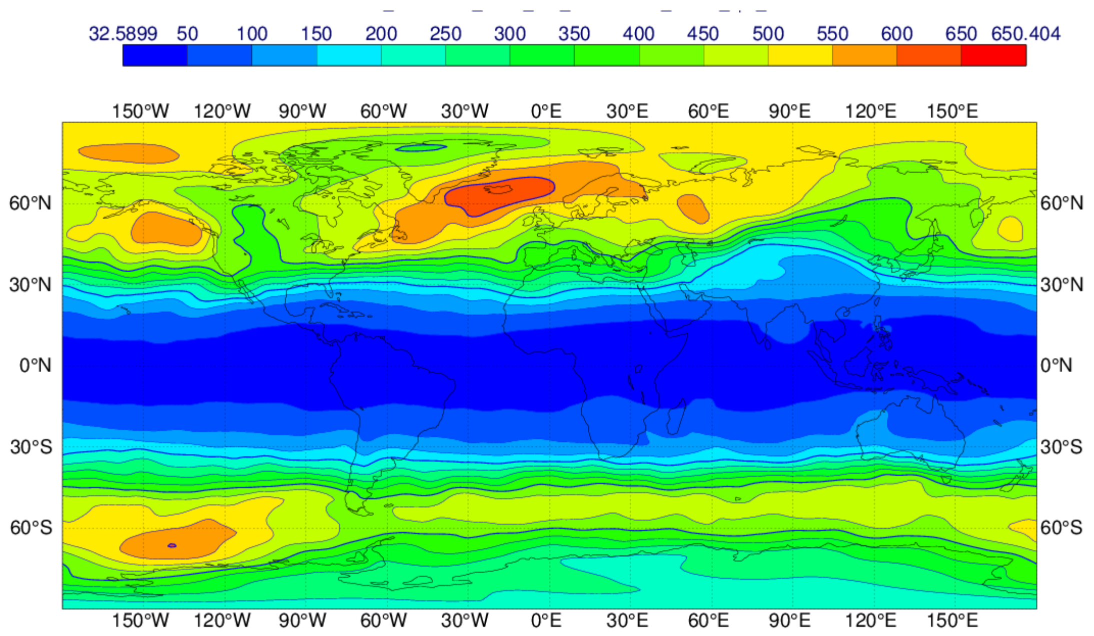
$$(\overline{x} - y)^2$$

- Average the two quantities for many locations and/or start times.
- The averaged quantities have to match for a reliable ensemble (within sampling uncertainty).

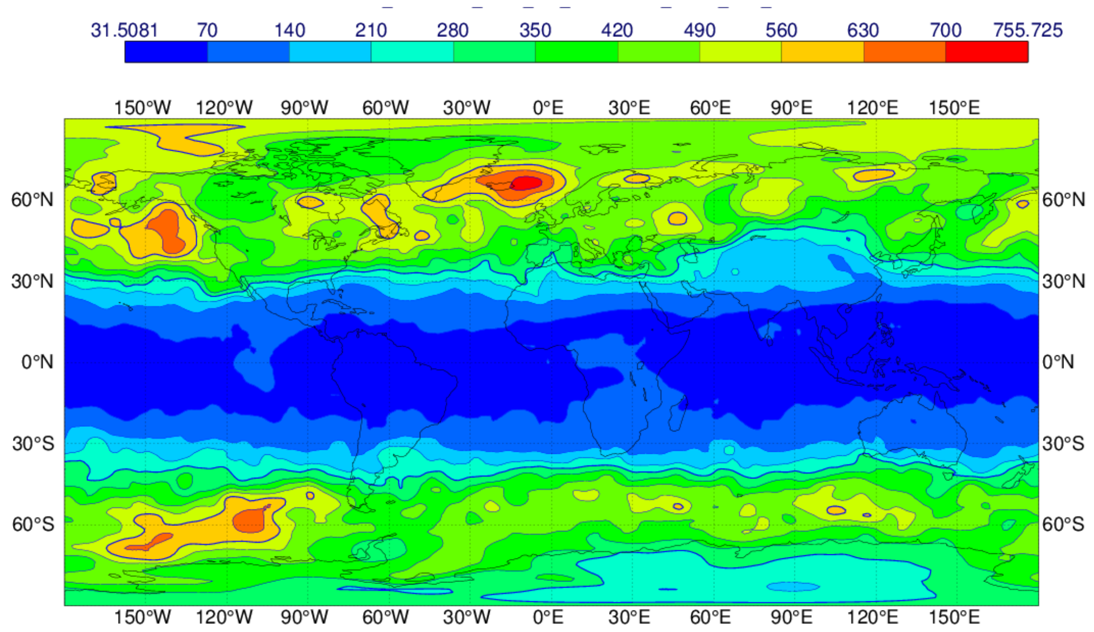
From Martin Leutbecher's lecture "Ensemble Verification 1"

Z500hPa
Ensemble StDev and Ensemble Mean RMSE, averaged 2016112200 – 2017021300
00 UTC Run

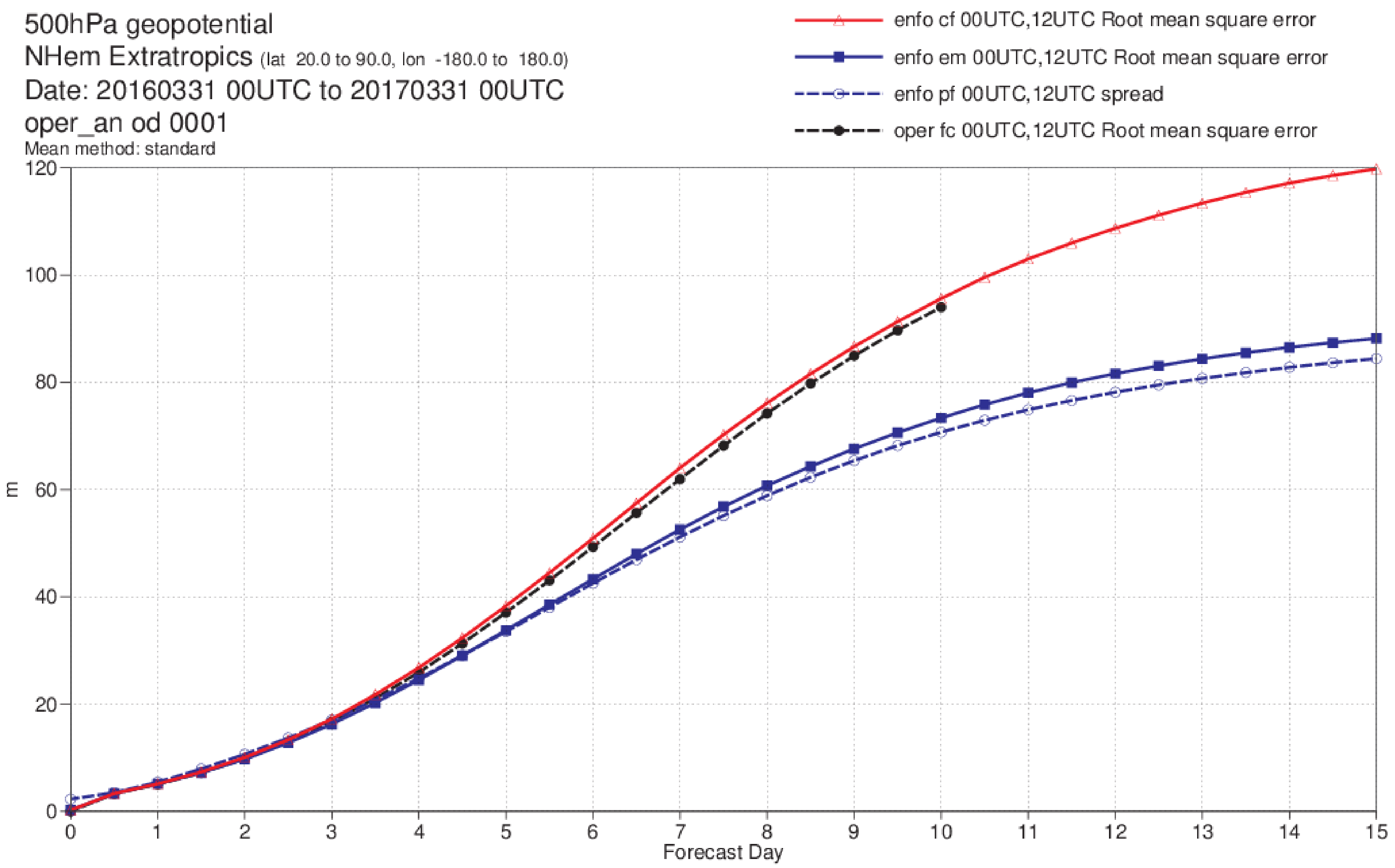
StDev T+120h



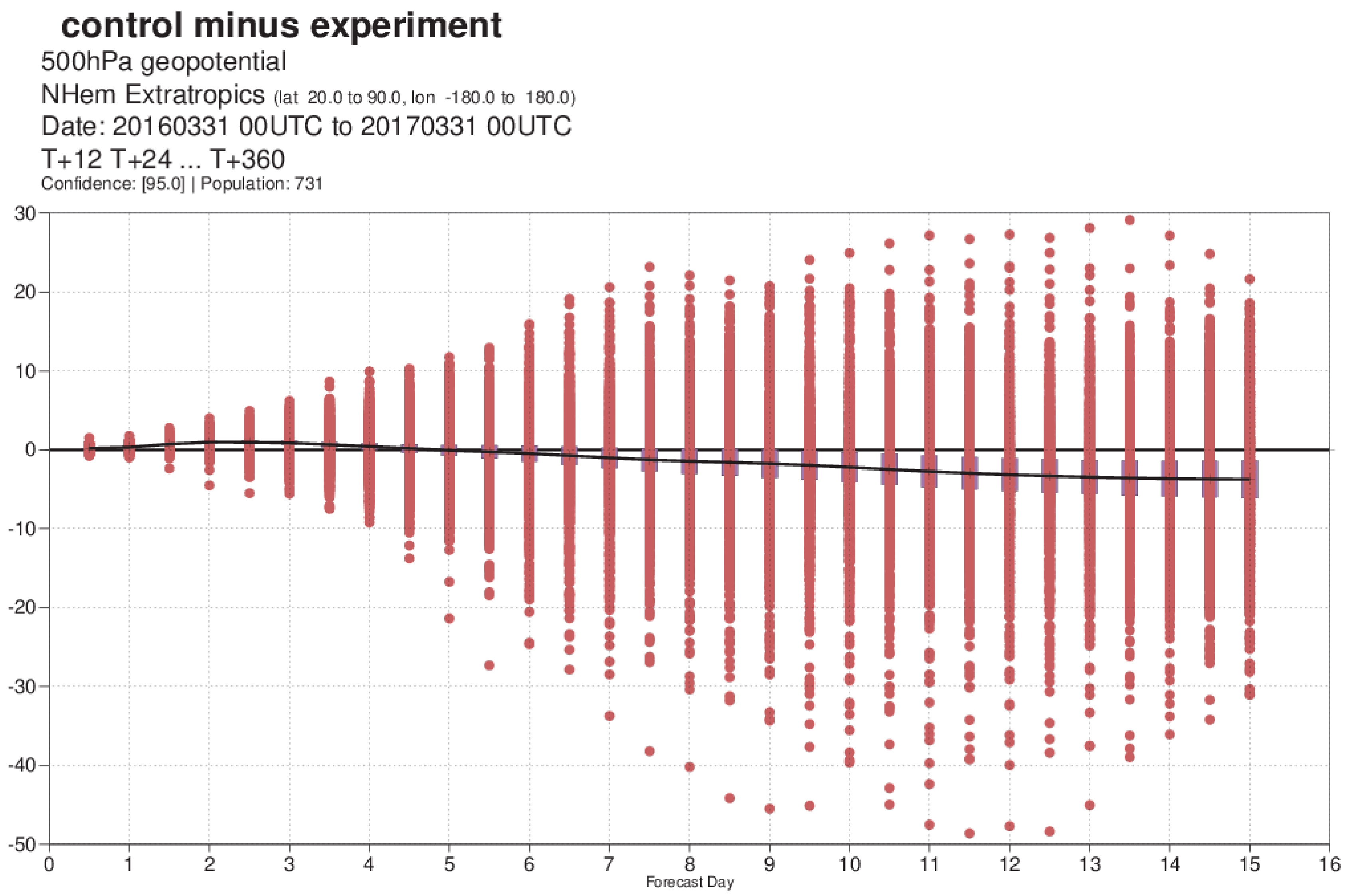
RMSE T+120h



Ensemble Spread vs Error

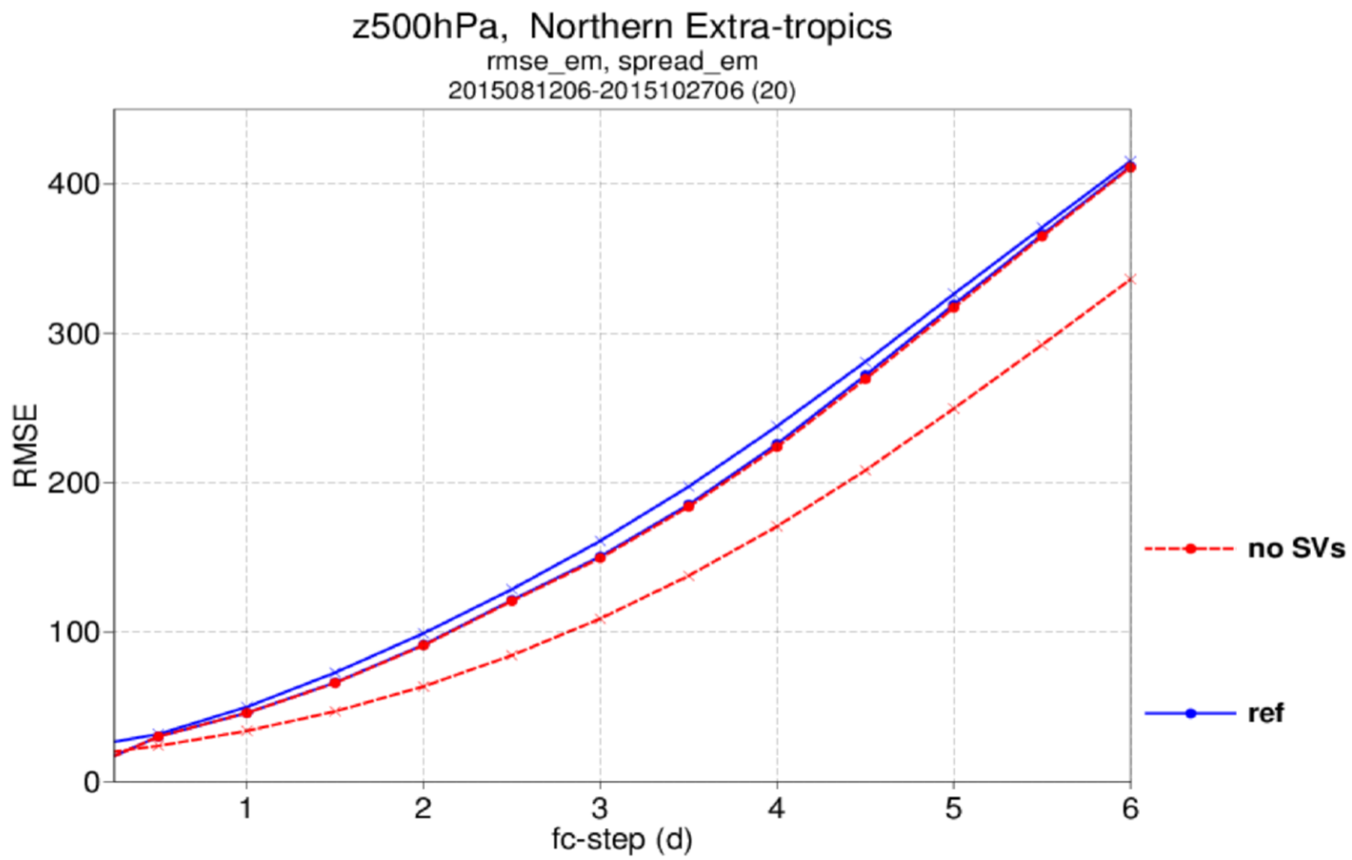


Ensemble Spread vs Error



Why SVs?

Impact of SVs on ENS



Oper like setup, TCo399, 20 Initial dates