Environnement Canada Environment Canada

Canada

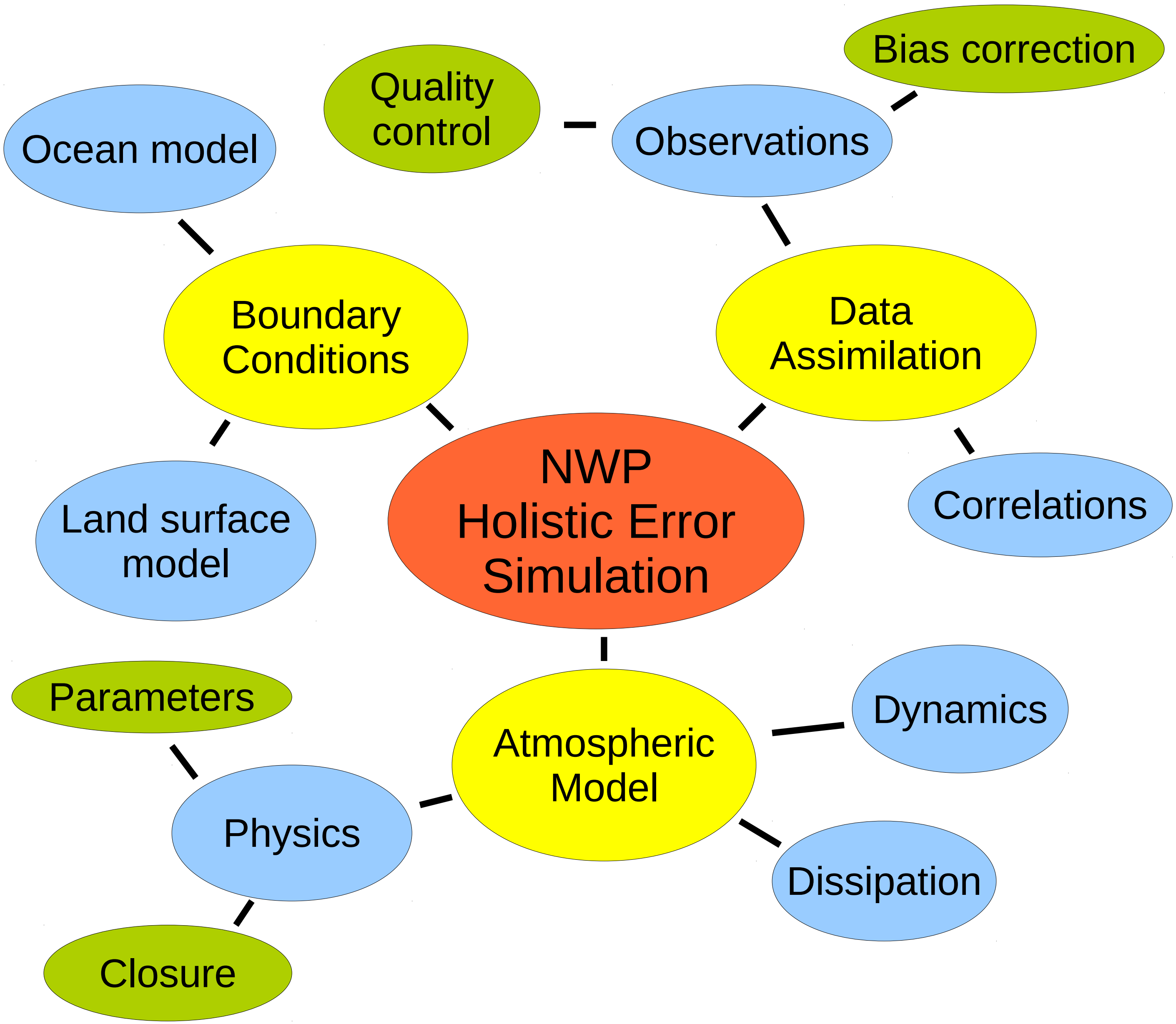
The Canadian approach to ensemble prediction

ECMWF 2017 Annual seminar:
Ensemble prediction : past, present and future.
Pieter Houtekamer
Montreal, Canada

Overview.

- The Canadian approach.
- What are the sources of error?
- Verification of the reliability of Canadian ensembles.
- For the future.

Consider the NWP system as a whole.

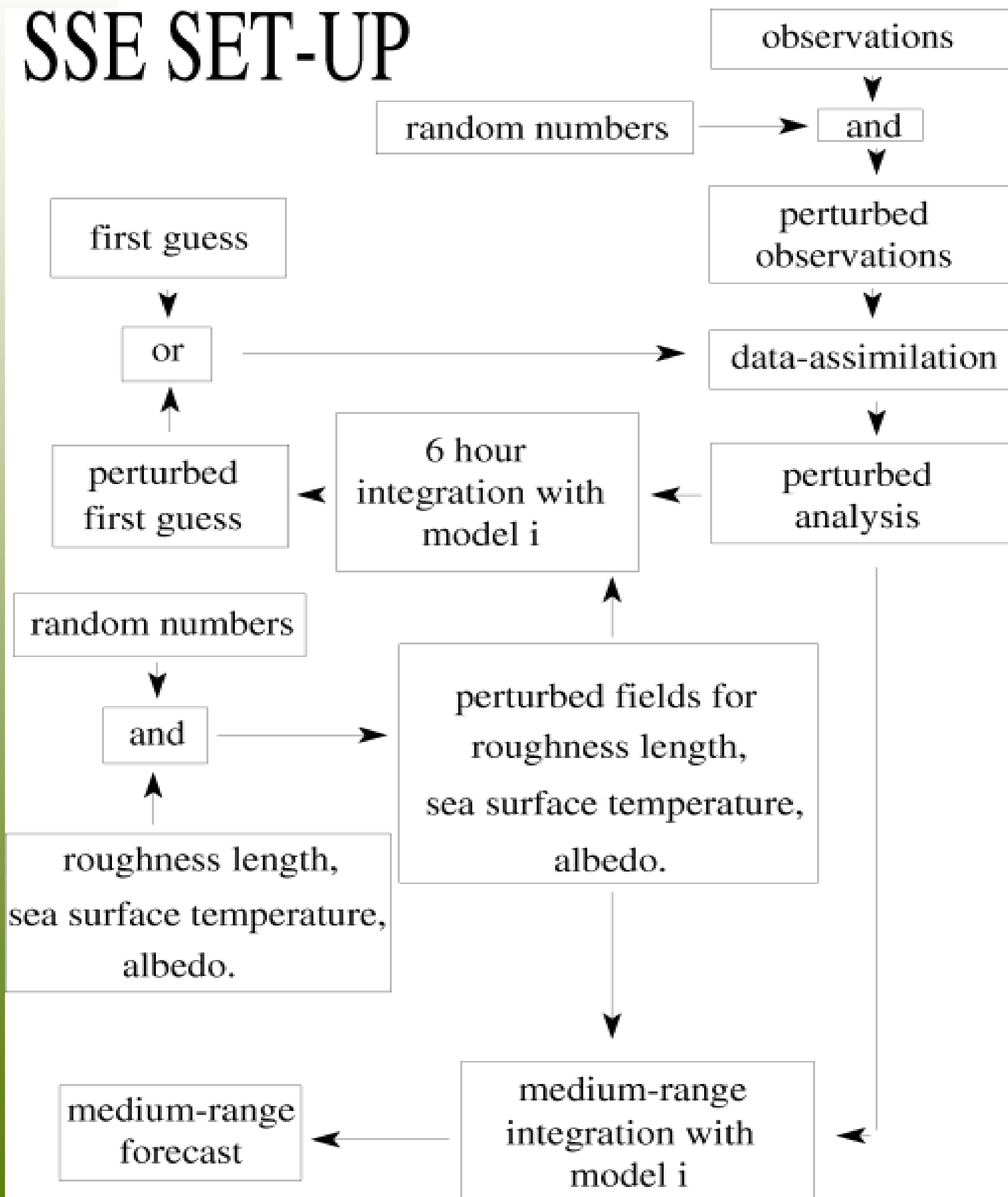


Throw the dice

If a component is uncertain, randomly select among possible alternatives.

To obtain one member of an ensemble, one may need to collect O(1 000 000) random numbers.

To get one member



To obtain one ensemble member i, the entire forecasting system is run with random perturbations to the observations, forecast model and surface fields.

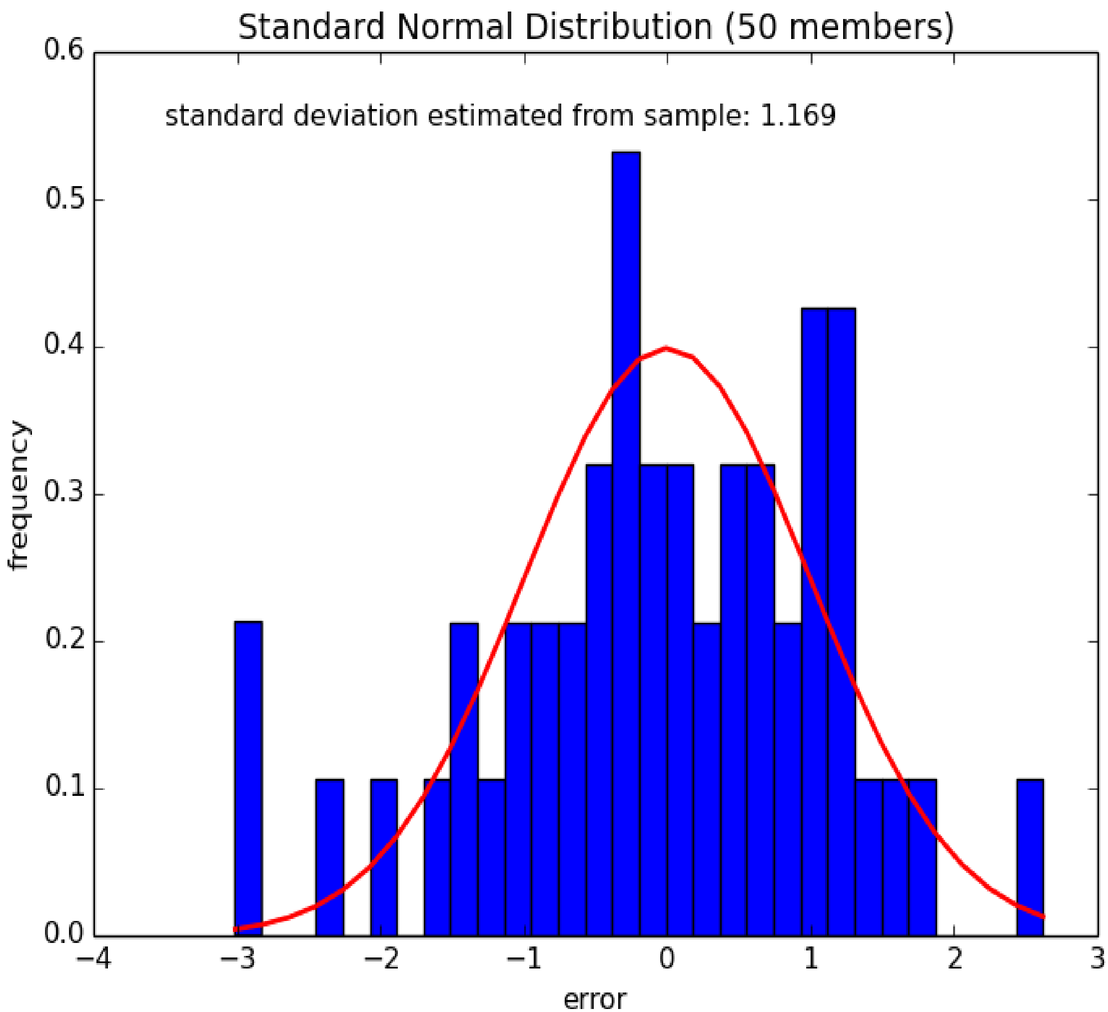
Random ensembles are available at all stages of the forecasting procedure to give information on uncertainty.

Figure from : Houtekamer et al., 1996, A system simulation approach to ensemble prediction. Mon. Wea. Rev.

With ensemble data assimilation (EDA), ECMWF now has a similar system.

Canada

Why are Monte-Carlo methods efficient?



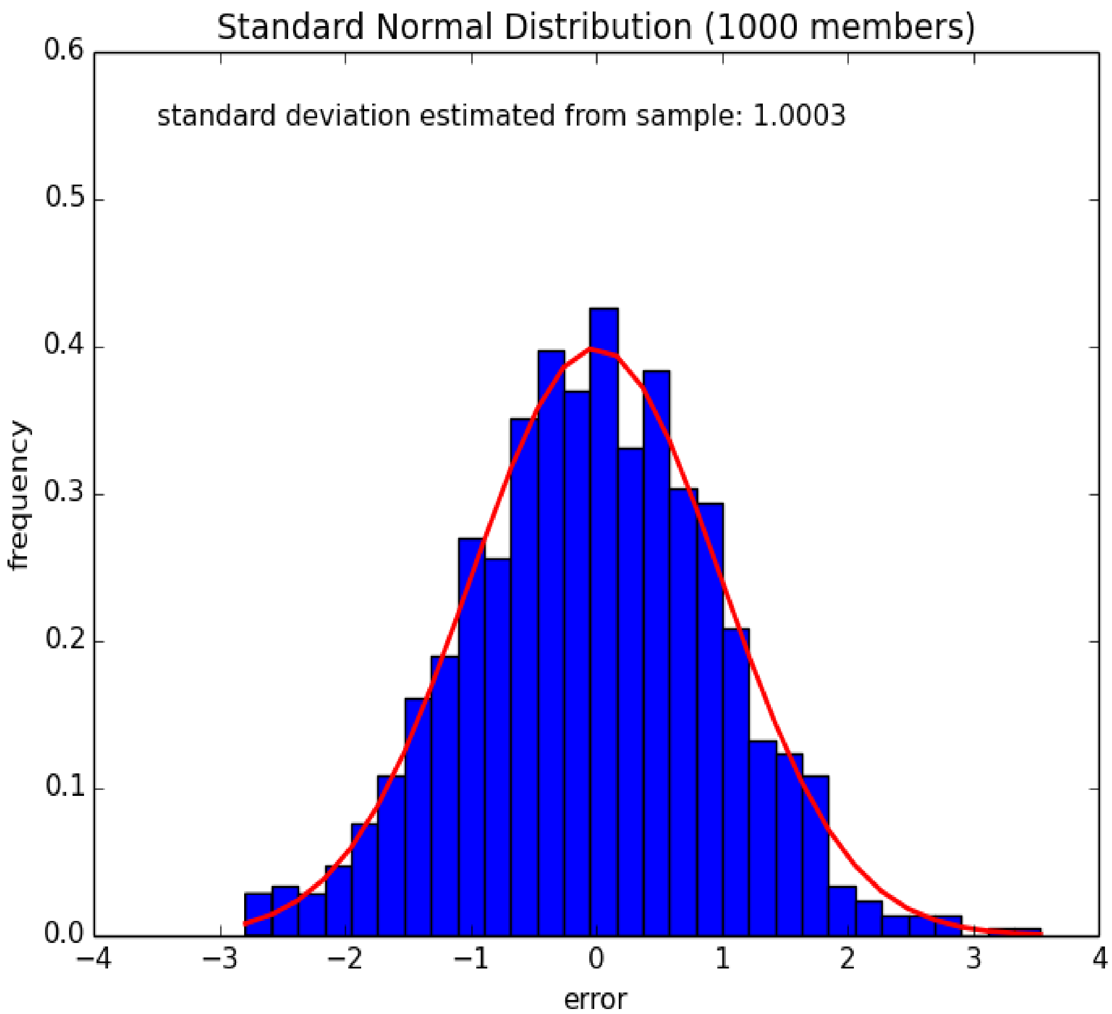
Take any scalar linear function f of say 1000000 random input variables, each with a Gaussian distribution. The random variable f will also have a Gaussian distribution.

Most users will not see benefits of having more than 50 members (Candille, Talagrand).

NAEFS has 40 members.
ECMWF has 50 members.

Canada

Need for large ensembles



Large ensemble have two main applications :

1) estimation of the probability of rare events by member counting.

2) estimation of covariance matrices (e.g. the Canadian Ensemble Kalman Filter uses 256 members).

Sources of error

To simulate the evolution of errors in the system, we first need an estimate of the importance of different possible sources of error.

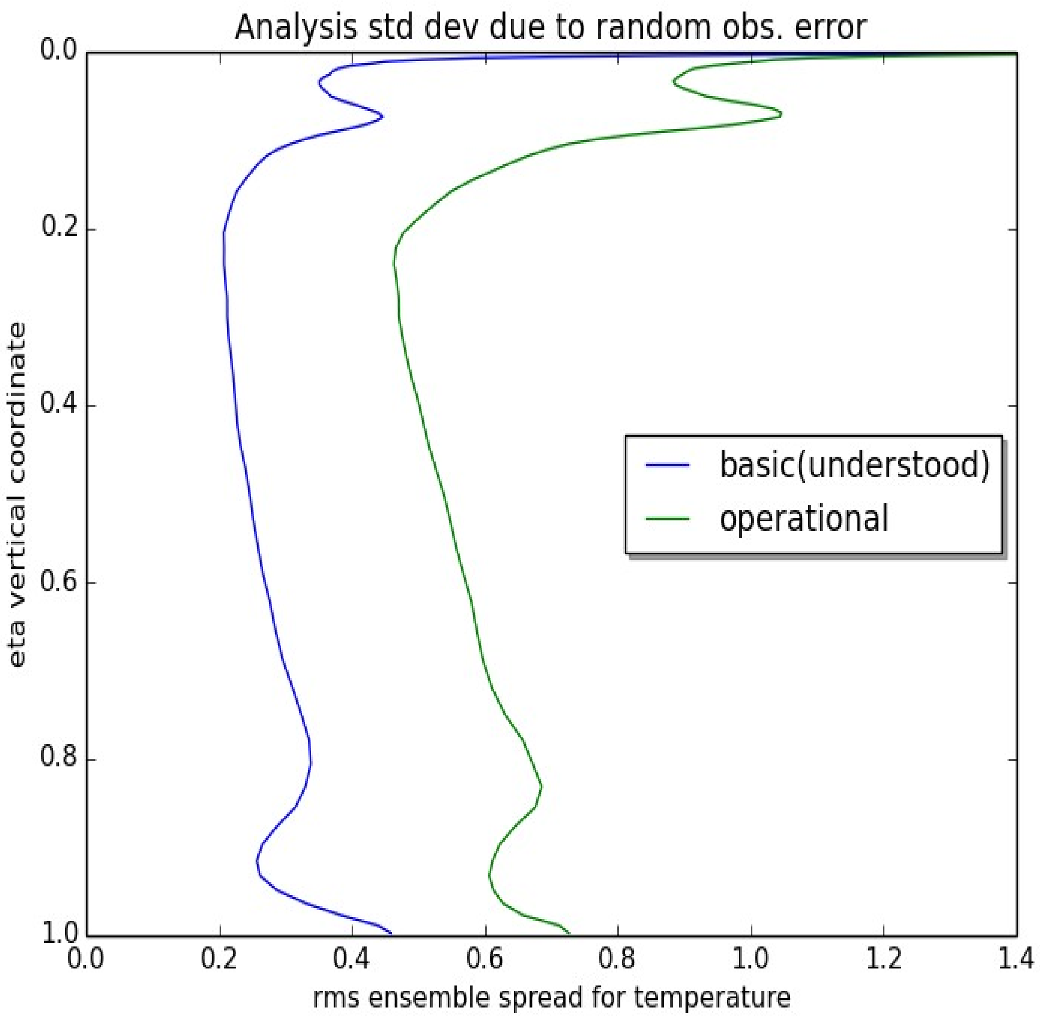
We will consider :

- Error due to limitations of the observational network
  (like having a limited number of observations, each with a random observational error).

- Error due to the forecast model
  (also known as model error). This error could be approximated by asking two independent groups to model the atmosphere.

- Error due to the data assimilation method
  This error could be approximated by asking two independent groups to design a data assimilation method.

Assuming that random observational errors are the only problem.

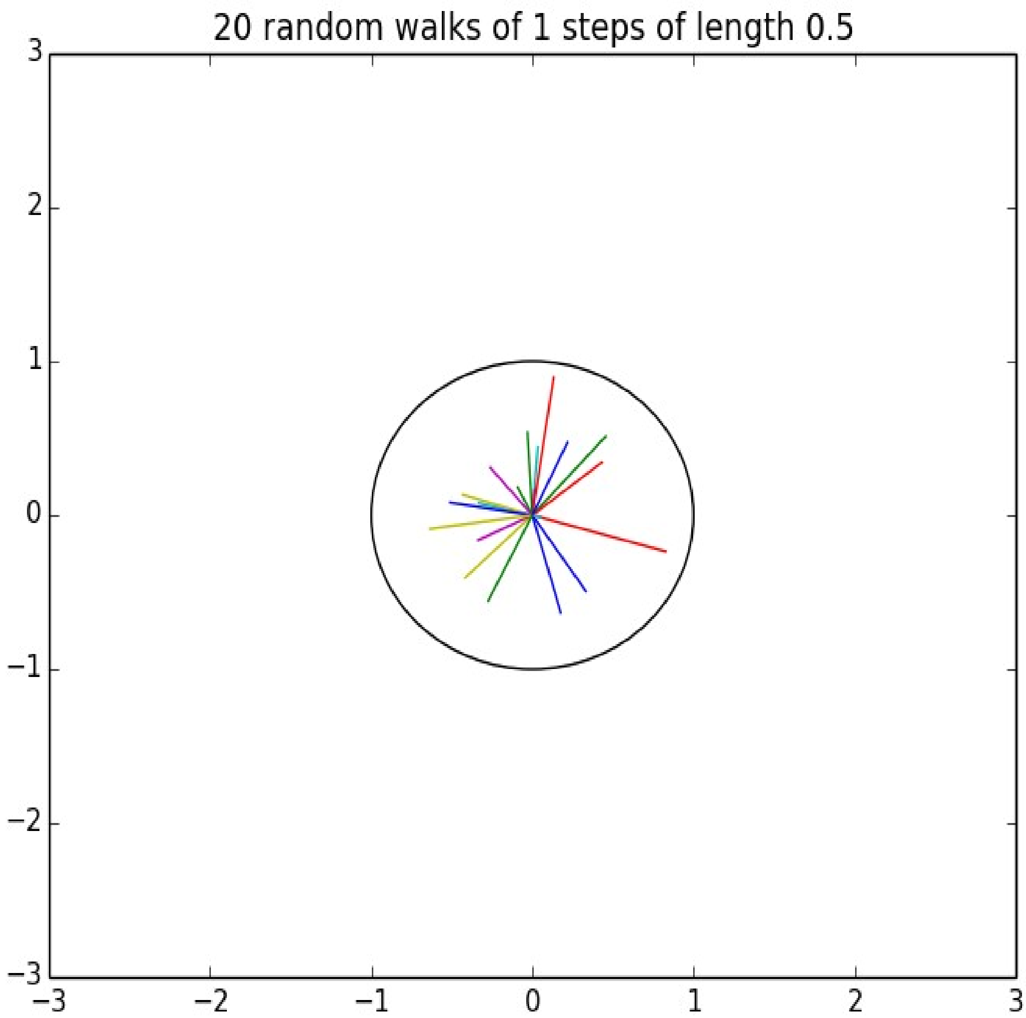


Experiment : Cycle the EnKF with all available observations. Observations are perturbed randomly.

Model error is not accounted for in any manner by the EnKF. A unique version of the model is used, there is no covariance inflation or addition.

Error levels, here shown for temperature as a function of height, are about half of what is thought to be realistic.

Covering distance with a random walk (1 step)

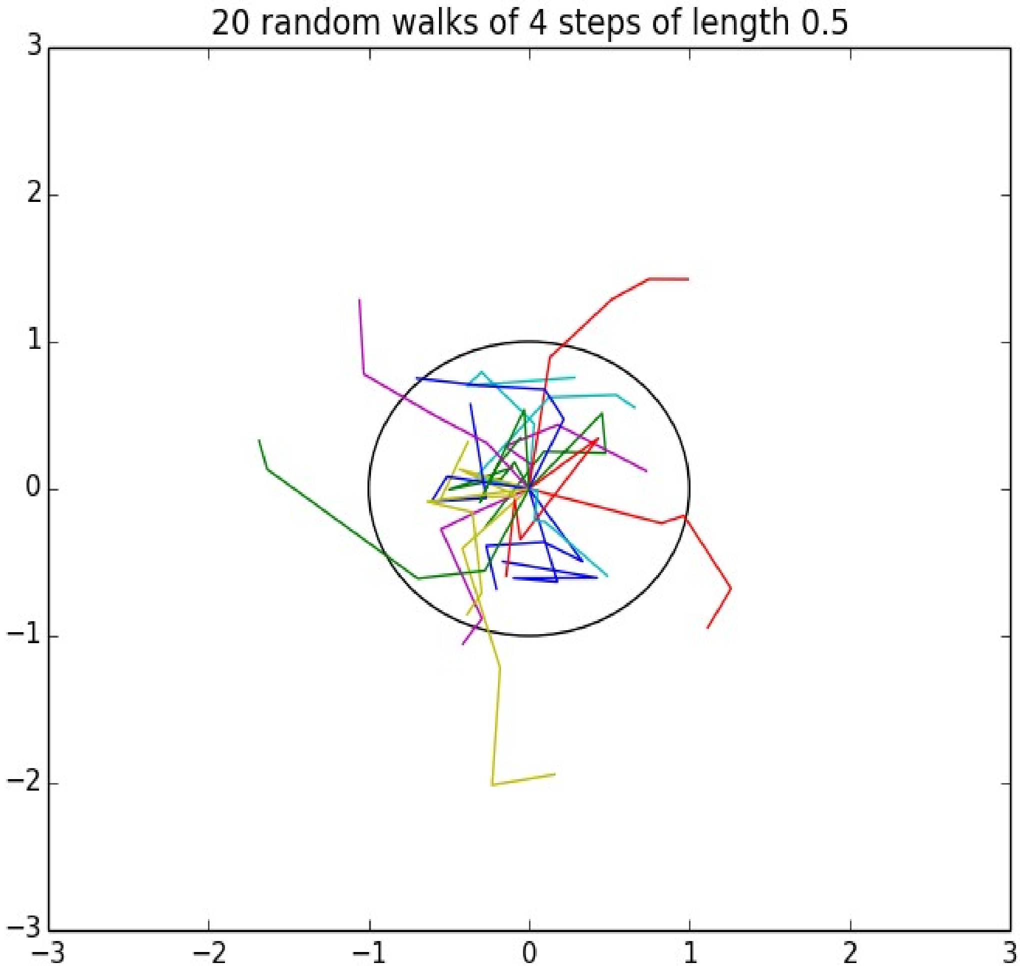


The x and y directions are independent. The standard deviation of the 2-dimensional step length = 0.5 (as in the example before).

The probability of arriving at a distance of 1.0 in one step is 1/20.

In the example, only one step is made and none of the 20 members arrives at a distance of 1.

Covering distance with a random walk of 4 steps

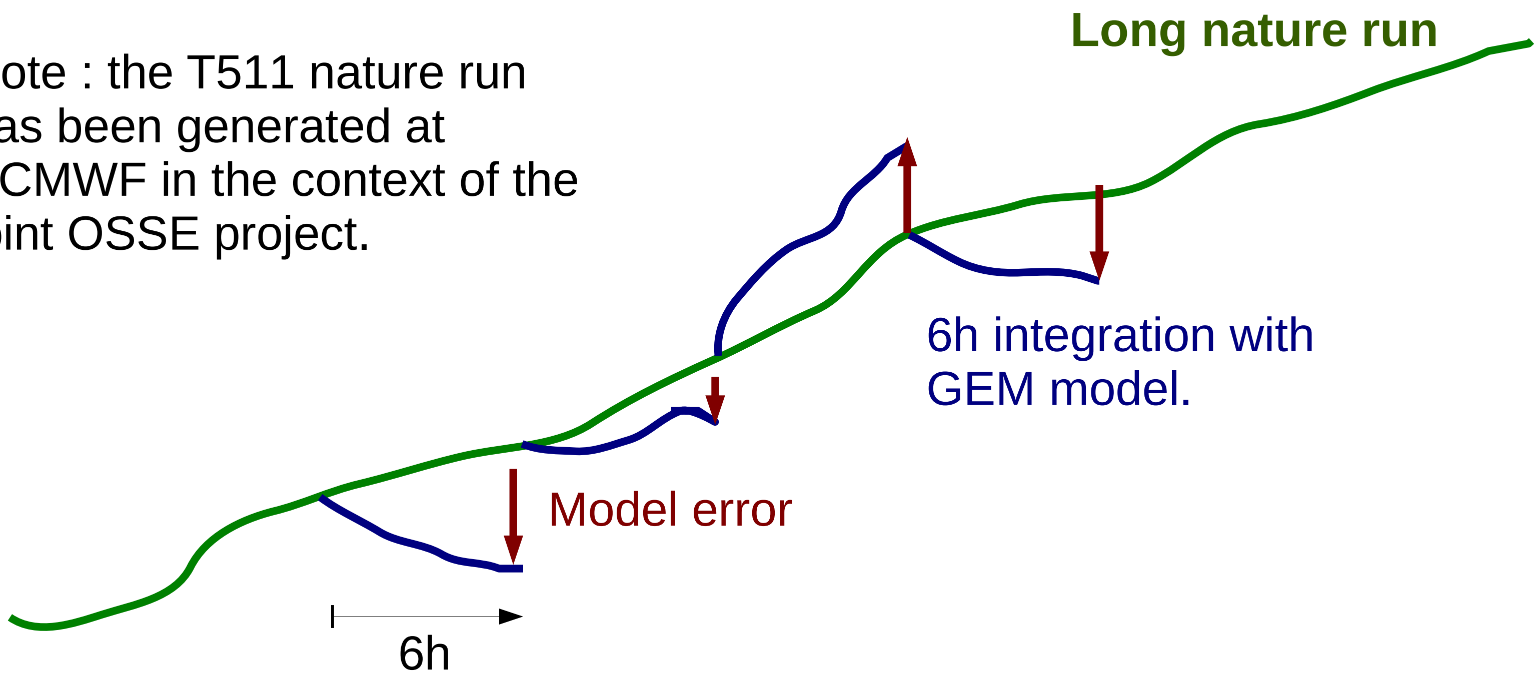


It takes 4 independent steps with a std dev of 0.5 to arrive at a total step with a std dev of 1.0.

Thus, with only the accurate simulation of the impact of observational uncertainty, an Ensemble Kalman Filter captures only one fourth of the error variance.

Sampling model error using the T511 ECMWF nature run

Note : the T511 nature run has been generated at ECMWF in the context of the joint OSSE project.



Every 6h, initial conditions for a 6h run with the Canadian GEM model are obtained from a long continuous nature run. At 6h, the difference is considered to sample model error. Over a one month period, we thus obtain 124 difference fields. From these we compute mean, standard deviation and rms.

The T511 ECMWF nature run

The nature run (NR) has been generated using version cy31r1 at T511 horizontal resolution (about 40 km), with 91 vertical levels. The model top is at 0.01 hPa.

Version cy31r1 was used at ECMWF operations between September 12 2006 and Dec 12 2006 with truncation T799 for the deterministic forecast.

The NR has been initialized using the ECMWF operational analysis on May 1 2005. It has been integrated until June 1 2006.

The NR was forced by daily Sea Surface Temperature and Ice Cover provided by NCEP.

It has been used for OSSE studies at Environment Canada (Heilliette et al. 2013). Here we use states from the NR every 6h in January 2006.

The GEM model integrations

The nature run (NR) provides initial conditions to the Canadian GEM model. The GEM model uses an 800 x 400 uniform grid (50 km resolution) with model top at 0.1 hPa. This configuration is currently used in R&D for the ensemble prediction system.

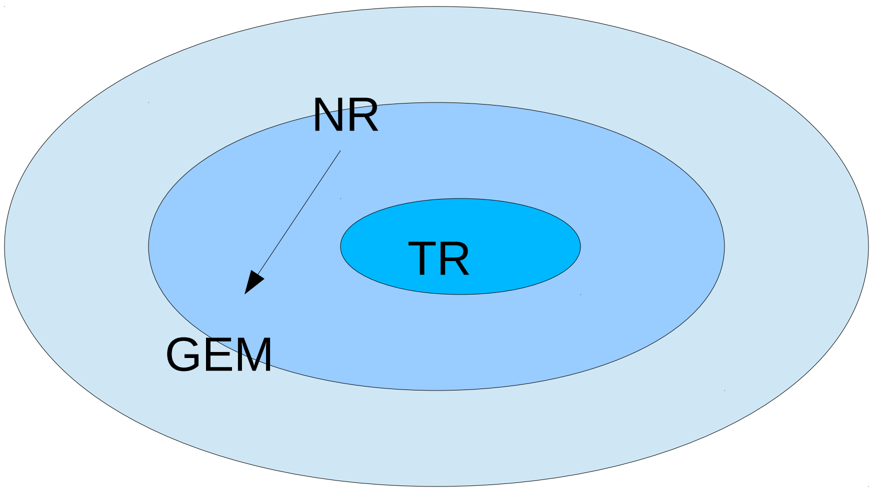
The model will interpolate to its grid and output the interpolated 0h field. This interpolated field is used in the computation of model error. A digital filter finalization procedure is used for balancing (Fillion et al. 1995).

Sea surface temperature and sea ice are from NCEP.

Other surface fields, such as for soil moisture and snow density, have been obtained from a multi-year run of a surface prediction system forced by ERA-Interim atmospheric fields.

Similar procedures are used for the Canadian reforecast system (developed by Normand Gagnon).

Hypothesis about model error : same quality of NR and GEM



Model error space

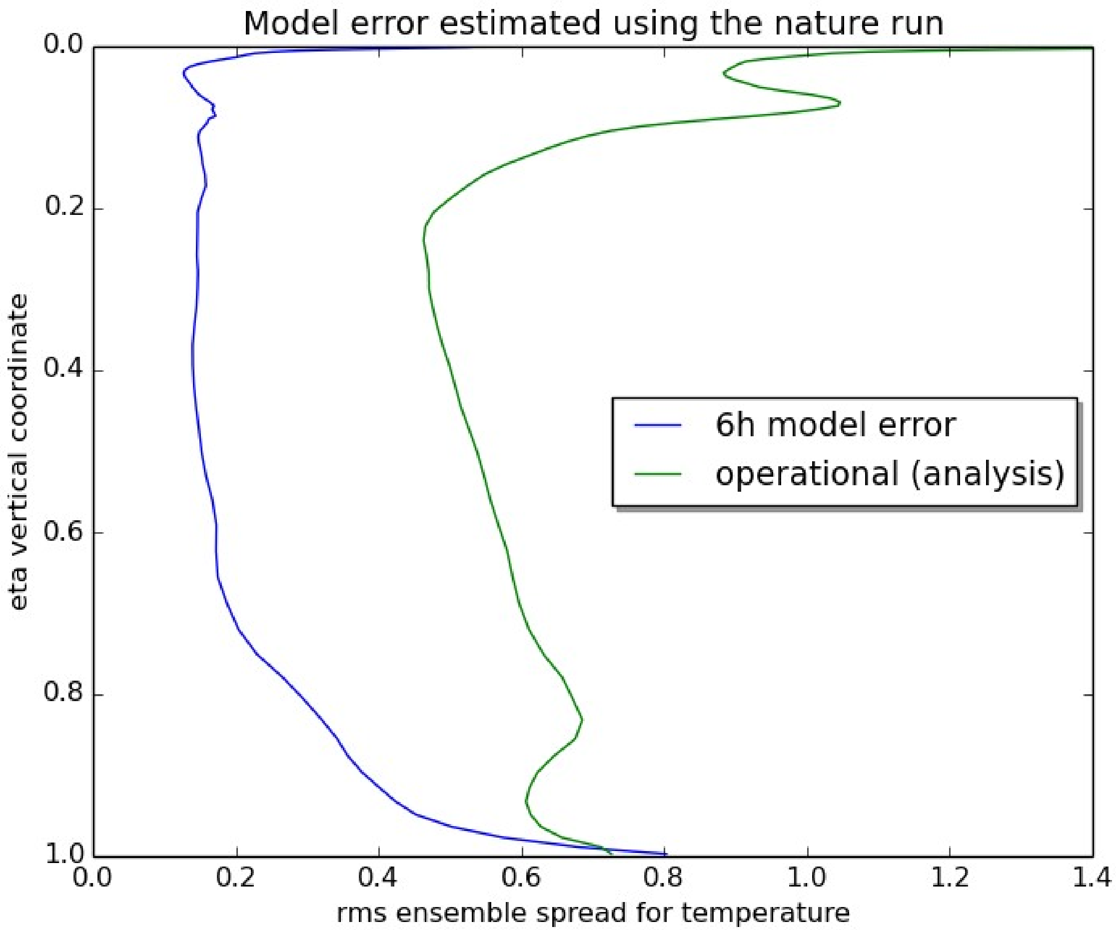
NR : Nature Run
(ECMWF T511 ~ 40 km).

GEM : Canadian model
(50 km).

TR : True atmosphere
(zero model error)

The working hypothesis is that the Nature Run and the GEM model provide samples of model error (with respect to the atmosphere) that are of *the same quality*. Thus the difference between NR and GEM will have to be scaled back by a factor root(2) to serve as an estimate of model error.

When model error is the only problem.



Experiment : Model error is estimated from the divergence over 6h between the ECMWF model and the GEM model (blue curve). It is compared with analysis error std dev in the EnKF (green curve).

Model error is a substantial error source. Near the surface, model error is a dominant error source.

Error due to the properties of the data assimilation method

At ECCC, we have two different operational data assimilation methods.

An EnKF is used to provide 20 initial conditions to the global medium-range ensemble prediction system.

An EnVar is used to provide a unique initial condition for the global deterministic high-resolution prediction system.

The EnVar procedure can be applied with the ensemble mean background as prior estimate (as opposed to using a high-resolution background trajectory). It will provide an analysis at the resolution of the background ensemble. This EnVar analysis can be compared with the ensemble mean analysis of the EnKF.

The difference between the EnKF and the EnVar is a measure of the uncertainty that is related to data-assimilation choices.

The EnVar implementation

•Background check : with respect to the high-resolution background of the high-resolution deterministic system.

•Quality control of observations : A variational quality control is used.

•Data selection : Only done to compensate for horizontal correlations.

•Balancing algorithm : An Incremental Analysis Update procedure forces towards an analysed *trajectory*.

•Localization is done in model space.

•Interchannel observation error correlations are important for channels that are sensitive to humidity.

•Analysis grid : Some operations are done in spectral space.

•Background error covariance : Isotropic covariances (B_NMC) are added.

The EnKF implementation

•Observations are obtained from the determinstic system. An additional background check is applied using the ensemble mean background state.

•Quality control of observations : A Huber norm is used.

•Data selection : Severe data selection limits computational cost as well as the impact of remote dense observations.

•Balancing algorithm : Incremental Analysis Update in which the increment is added using a smooth bell-shaped function centered at the analysis time.

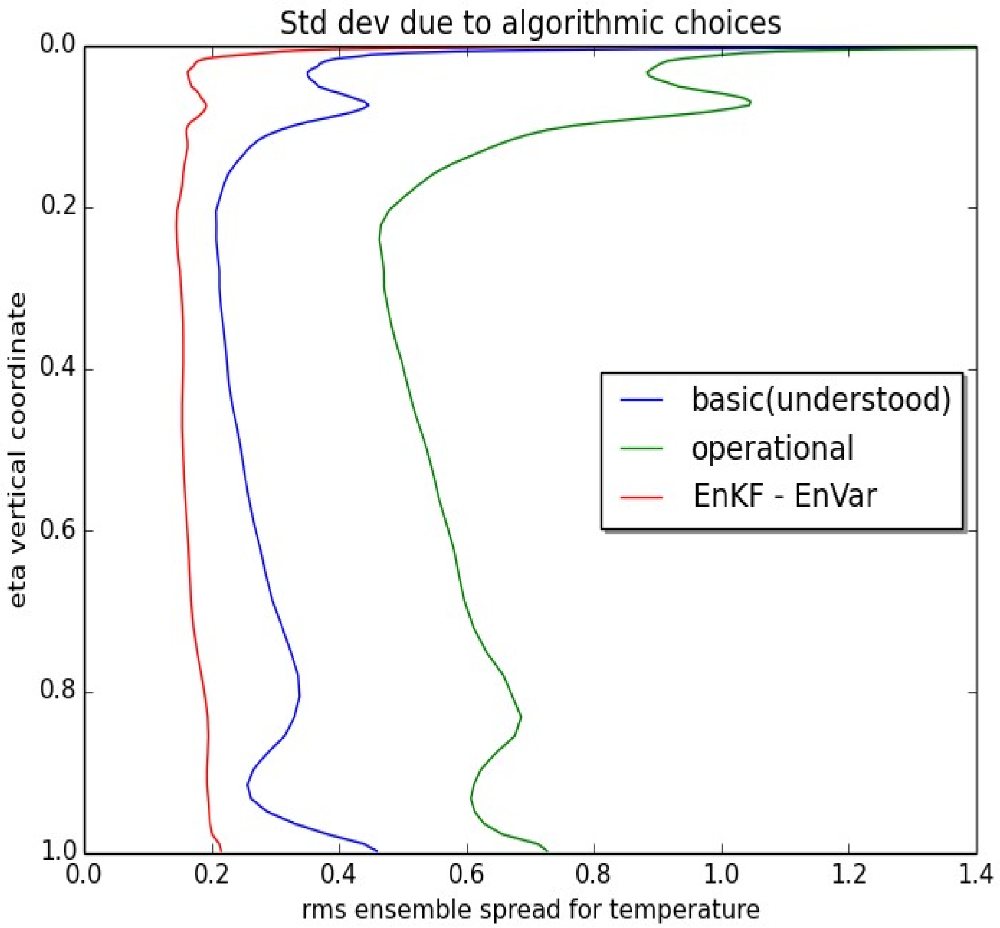
•Localization is applied to HBH^T and to the BH^T matrix.

•Interchannel observation error correlations are not used due to a lack of impact.

•Analysis grid : the analysis uses the Yin-Yang grid.

•Background error covariances are directly from the ensemble.

Errors due to choices in the data assimilation algorihtm.

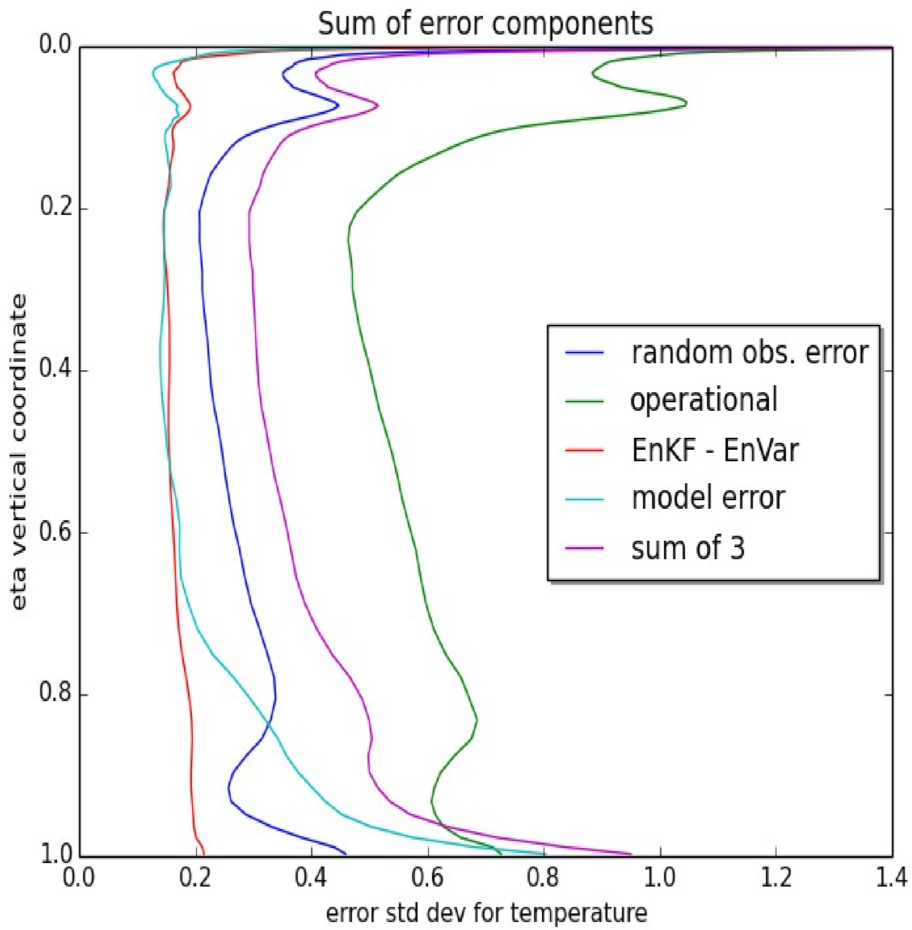


Experiment : For a 10-day period, compare the ensemble mean EnKF analysis with the EnVar analysis.

Assuming that the EnVar and EnKF are of similar quality, differences have been divided by root(2).

The uncertainty due to data assimilation assumptions is somewhat smaller than the uncertainty due to observational errors.

Does it add up?



The three previously discussed error components are considered to be independent and have been added (purple curve).

This curve can be compared with the operational estimate (green curve).

With an exception for the lowest levels, where model error dominates, *the estimated uncertainty is too small*.

Note that all the curves in the figure depend on hypotheses that may not be accurate.

Conclusion on error sources

We quantified the importance of (i) random observational errors, (ii) model error, (iii) weaknesses of data assimilation methods.

• We estimated (i) using an EnKF system (switching off any parameterization of model error).
• We estimated (ii) using the NR. Errors are large near the surface and near the model top.
• We estimated (iii) using differences of EnKF and EnVar analyses. Errors are somewhat smaller than in (i).

We have like to set-up an OSSE-type cycling algorithm that combines these 3 components.

Verification of ensemble forecasts

Ensemble forecasts are performed to obtain a probability distribution function.

It is desirable that the ensemble forecast is reliable :
1) On average, the predicted uncertainty should correspond with the average forecast quality. *Post-processing of a deterministic forecast could give a pdf that respects this criterion.*
2) The day-to-day variations of the predicted uncertainty should correspond with observed variations in forecast quality. i.e. *there should be a correlation between spread and skill.*

Error sources in the Canadian global ensemble

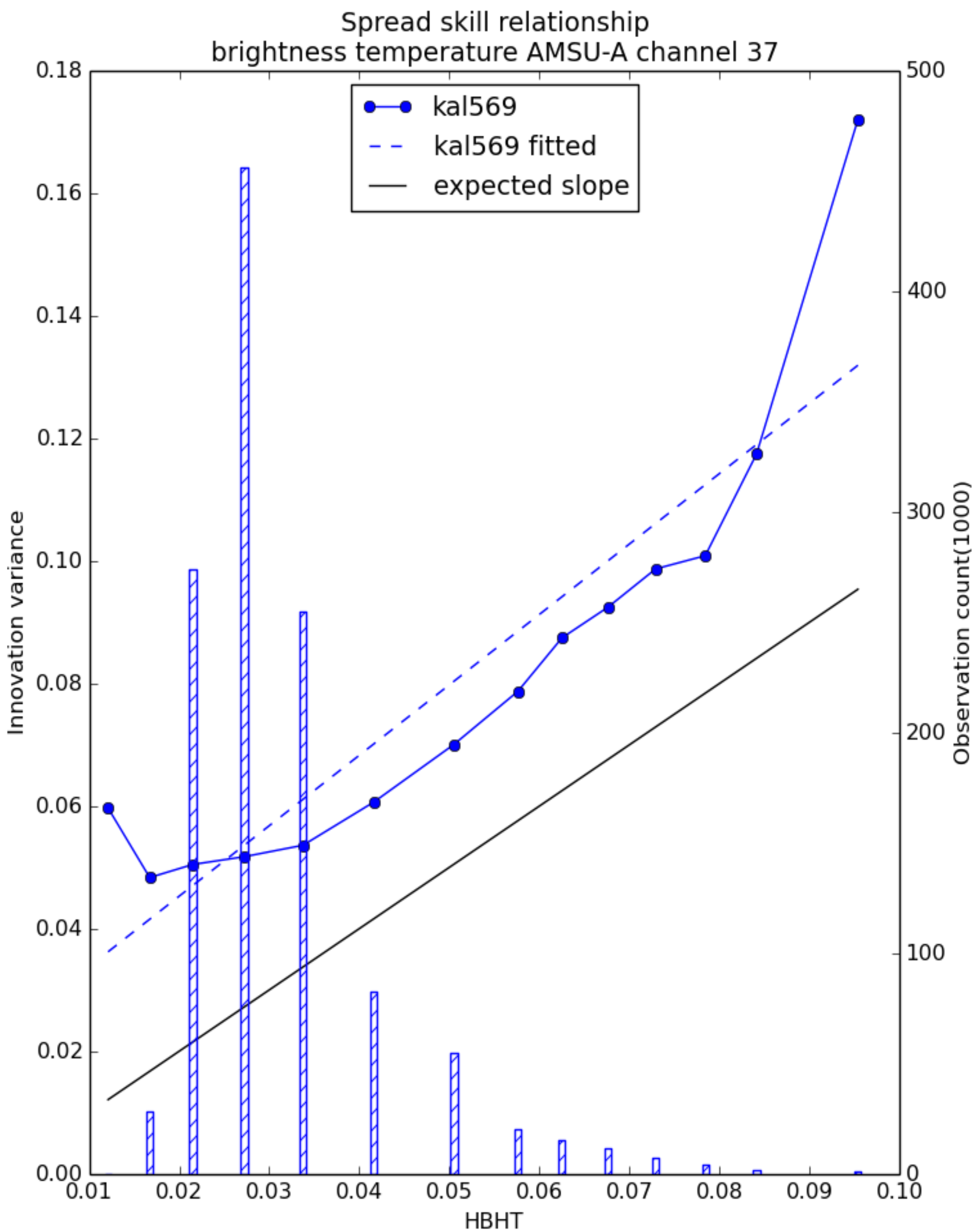
EnKF :

1) Random observational error
2) Multiple physical parameterizations
3) Isotropic additive error every 6h

Medium-range ensemble :

- Multiple physical parameterizations
- Isotropic additive error at the initial time
- Stochastic Kinetic Energy Backscatter (SKEB)
- Physical Tendency Perturbations (PTP)

Spread-skill relation for the EnKF AMSU-A channel 37



The ensemble spread is a predictor for the quality of the ensemble mean :

$$var(O\text{-}Hx) = HBH^T + R$$

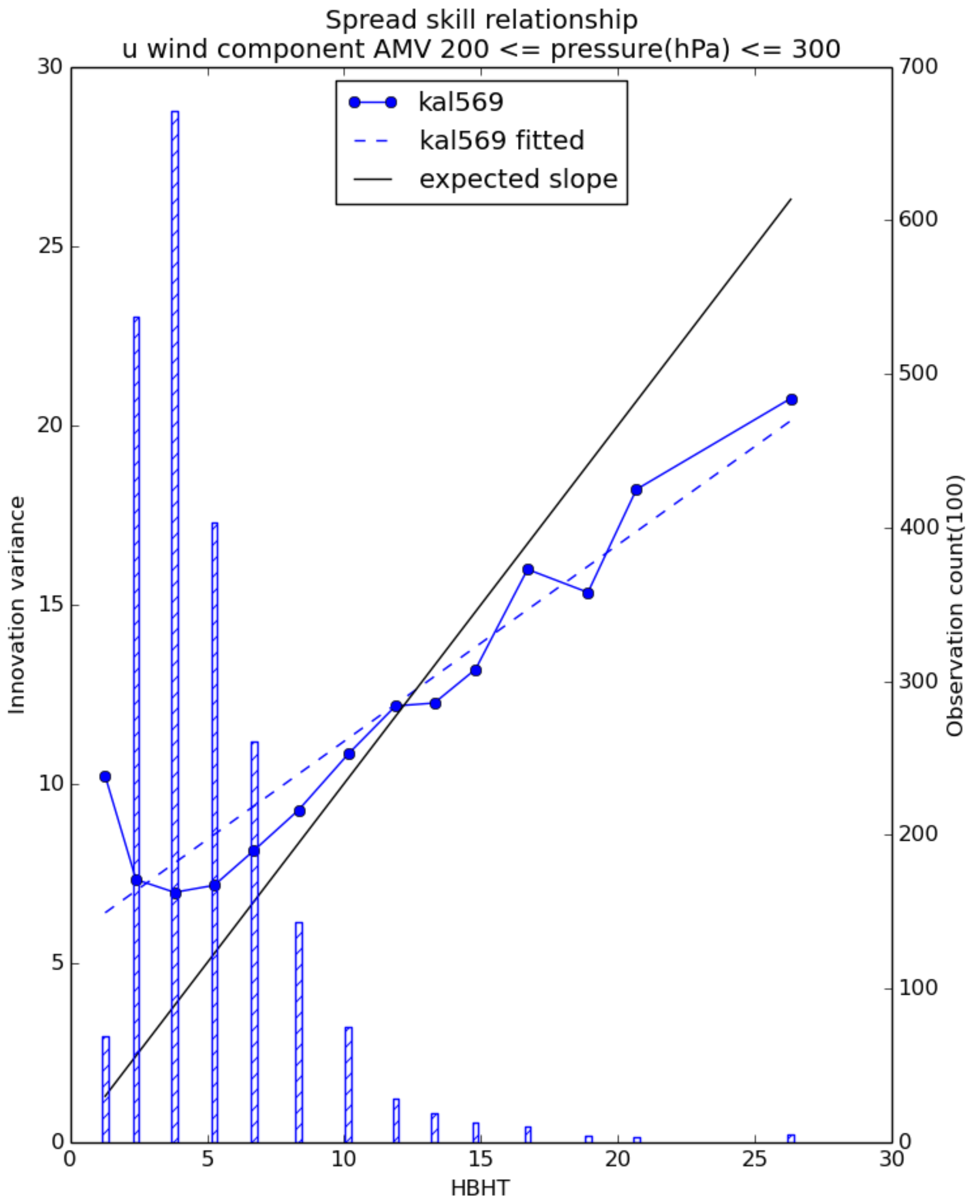
The quality of the mean is measured with O-Hx.

The interpolated ensemble spread is HBH^T and R is the observation error variance.

Results are shown for AMSU-A channel 37(10) which peaks at about 50 hPa.

Canada

Spread-skill relation for the EnKF AMV u-wind



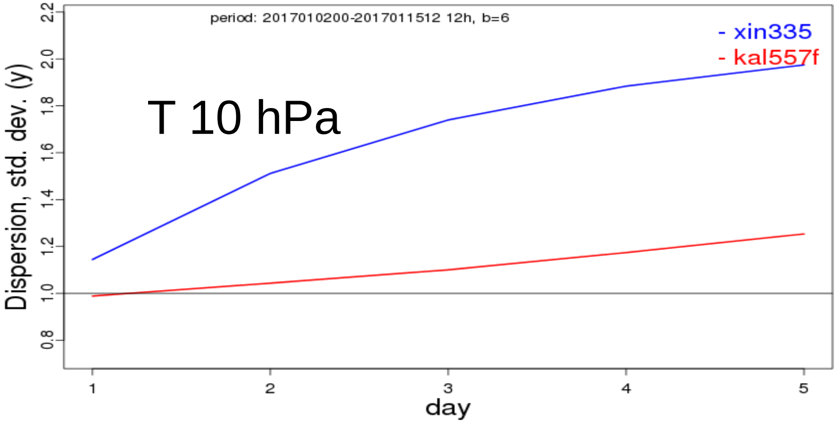
Results are shown for Atmospheric Motion Vectors for pressures between 200 and 300 hPa.

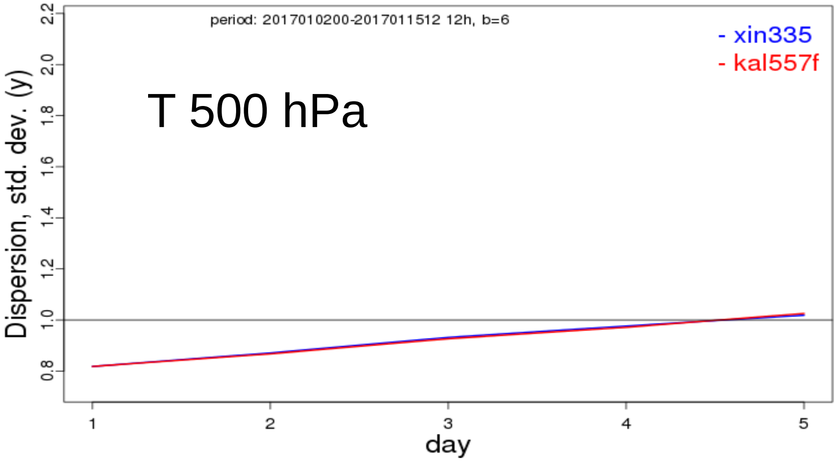
Larger interpolated ensemble spread corresponds with larger distance to the observations.

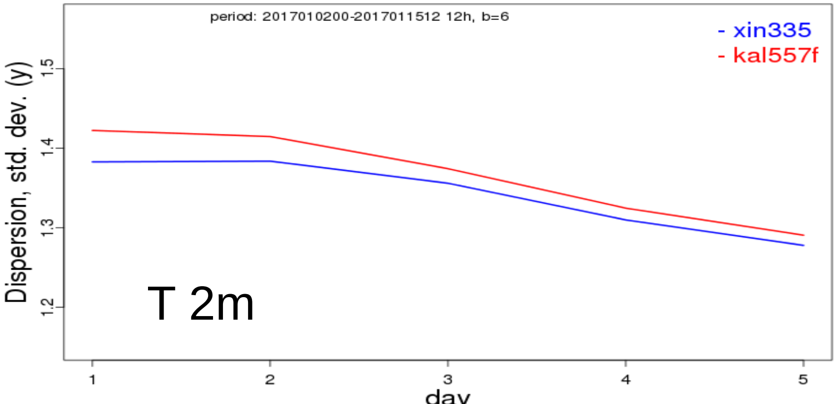
This suggests the EnKF ensembles are reliable.

Note, however, *the very small number of observations with large corresponding ensemble spread.* This makes verification difficult.

Canada







Reliability of the medium-range ensemble

Dispersion² = var(O-Hx)/(HBH^T + R)

Comparing the distance between the ensemble mean with observations with the ensemble spread and observational error, for many tropospheric variables the ensemble spread is nearly perfect (less than 20% over- or under-dispersive). Results, here, are for temperature measured by radiosondes.

The operational ensemble is underdispersive by a factor of 2 at 10 hPa. This problem will be alleviated by raising the model top from 2 hPa to 0.1 hPa.

The ensembles are also thought to be underdispersive near the surface.

For the future

We need to improve our understanding and simulation of error sources in the data-assimilation cycle.

We expect improvement near the surface from coupling with surface ensembles (e.g. with Caldas = Canadian Land Data Assimilation System).

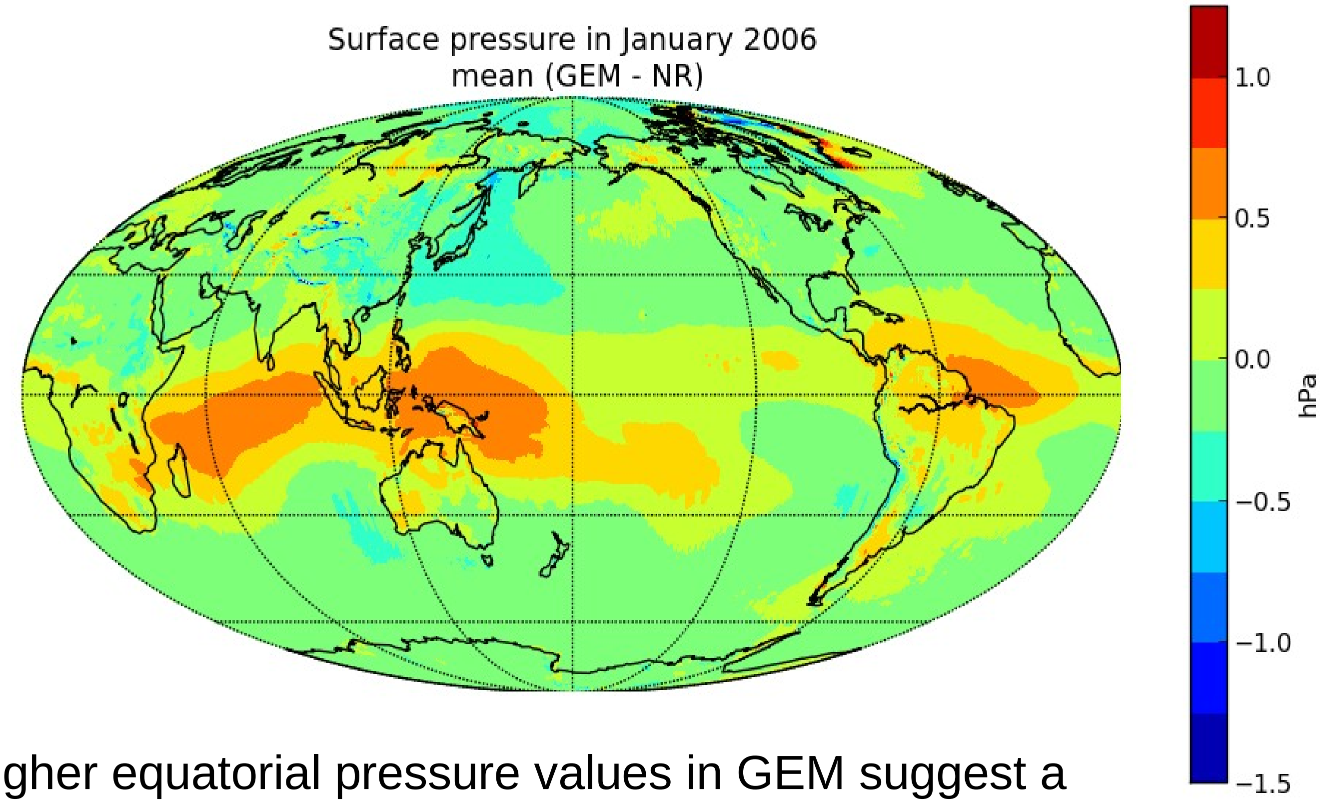
We would like to benefit from having two different data assimilation systems (with a Hybrid Gain type approach).

We would like to sample model error in the same way in the EnKF and in the medium range ensembles (i.e. using SKEB and PTP also in the EnKF).

Thank you for your attention

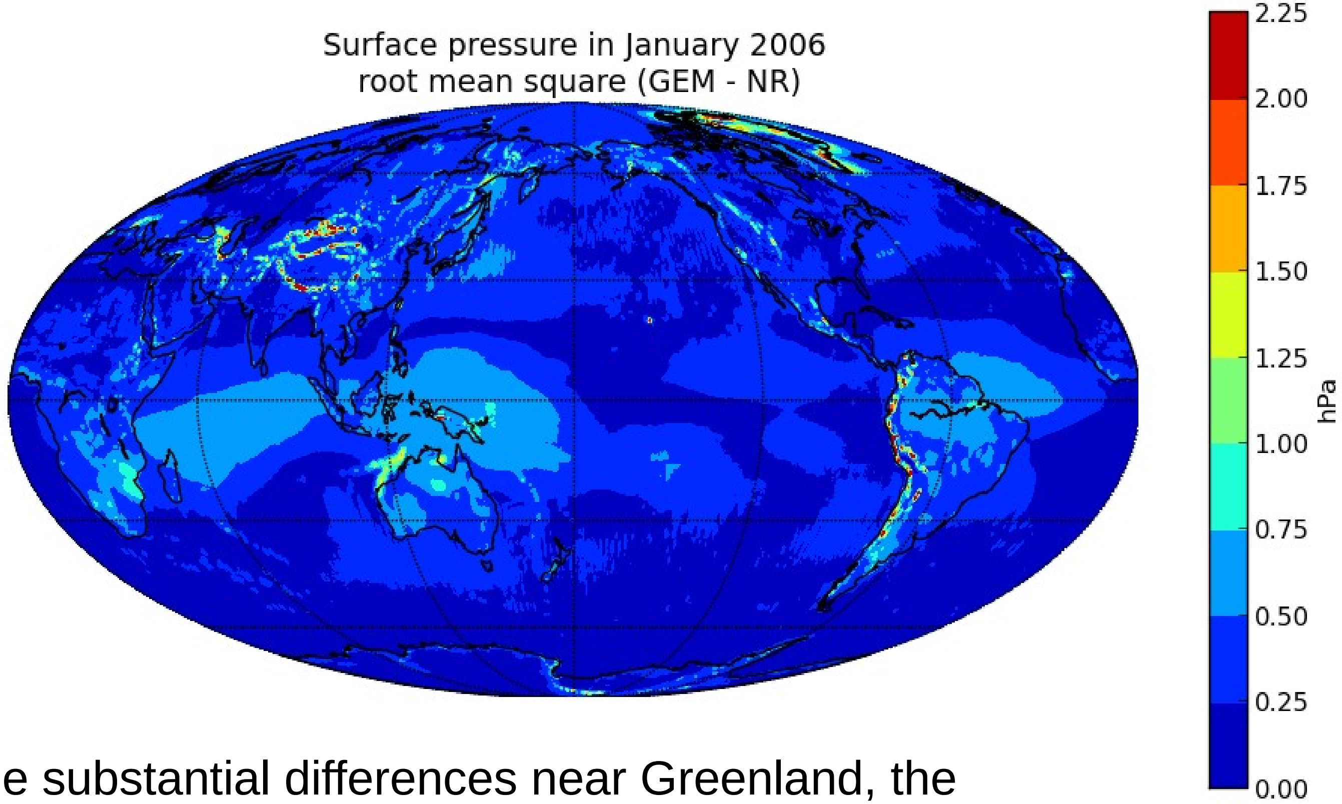
Supplementary information

Mean 6h surface pressure model error for January 2006



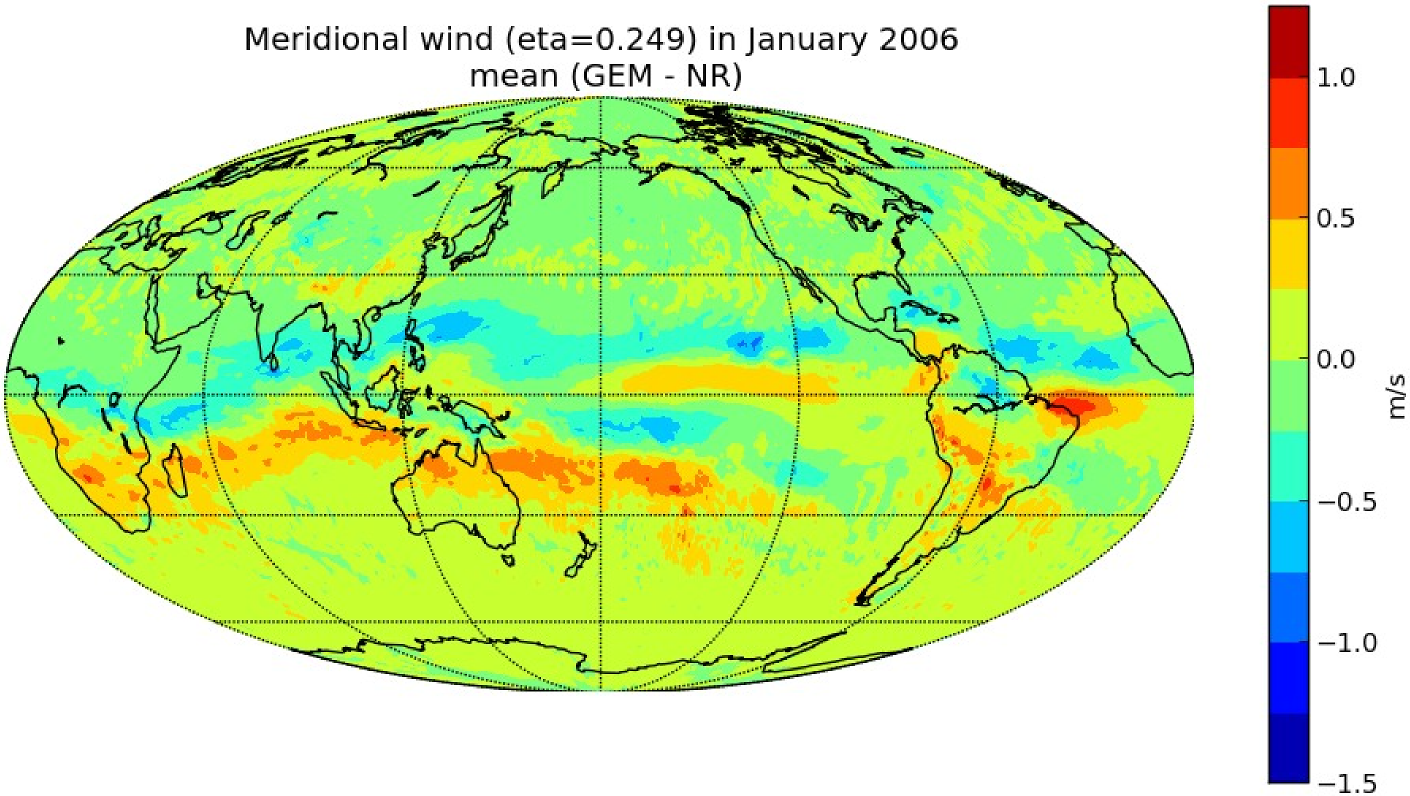
The higher equatorial pressure values in GEM suggest a weaker Hadley circulation in GEM.

RMS of 6h surface pressure model error for January 2006



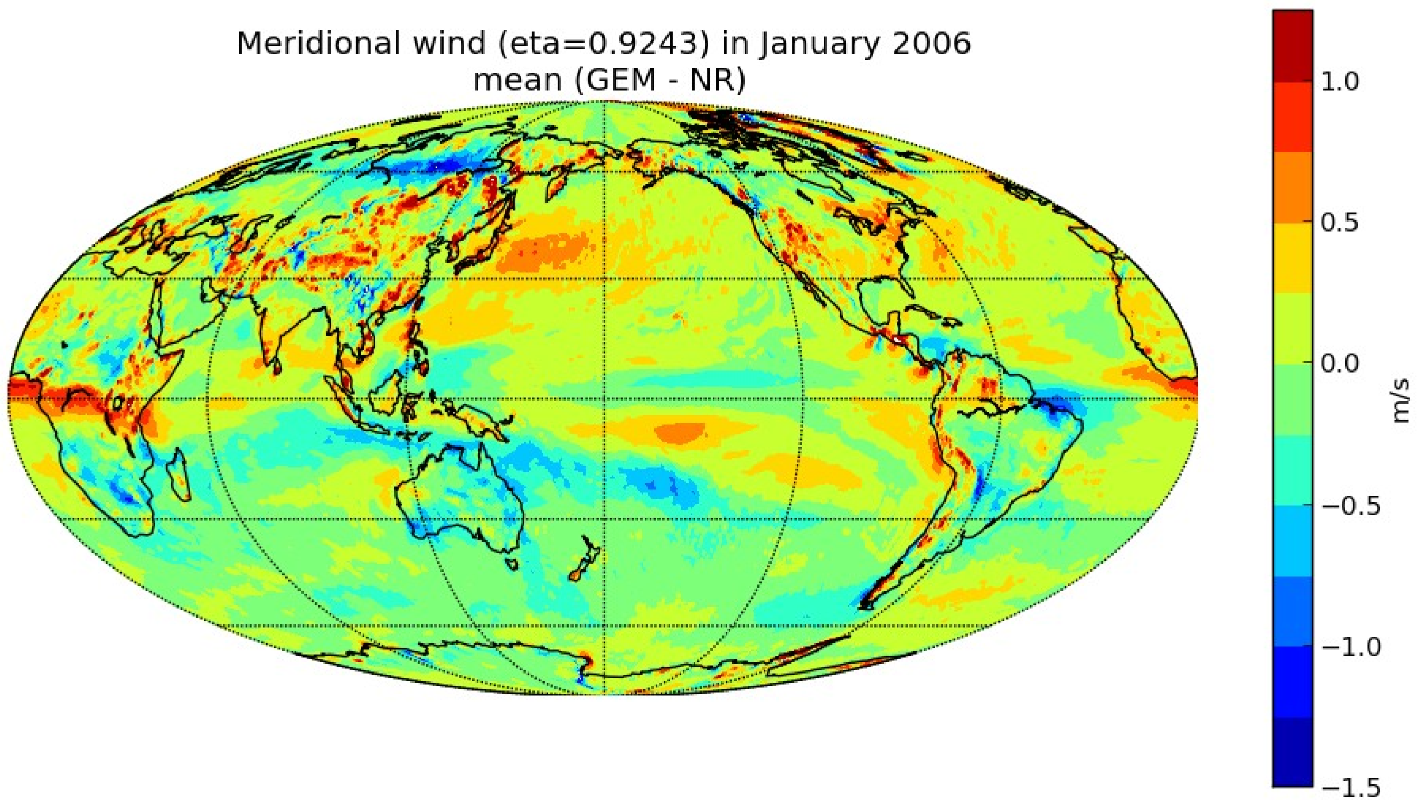
Note the substantial differences near Greenland, the Andes and the Tibetan plateau.

Mean 6h meridional wind (eta=0.249) model error for January 2006



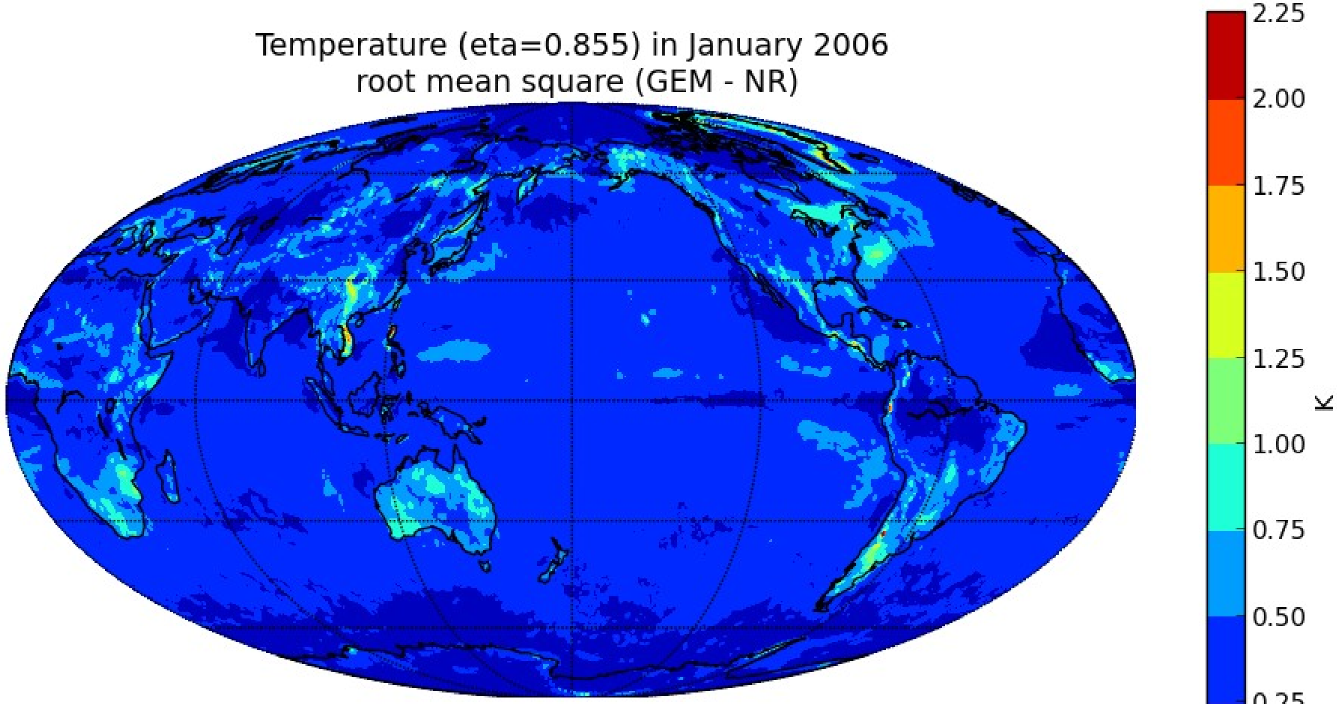
The bias pattern is consistent with the ITCZ being less intense in GEM.

Mean 6h meridional wind (eta=0.924) model error for January 2006

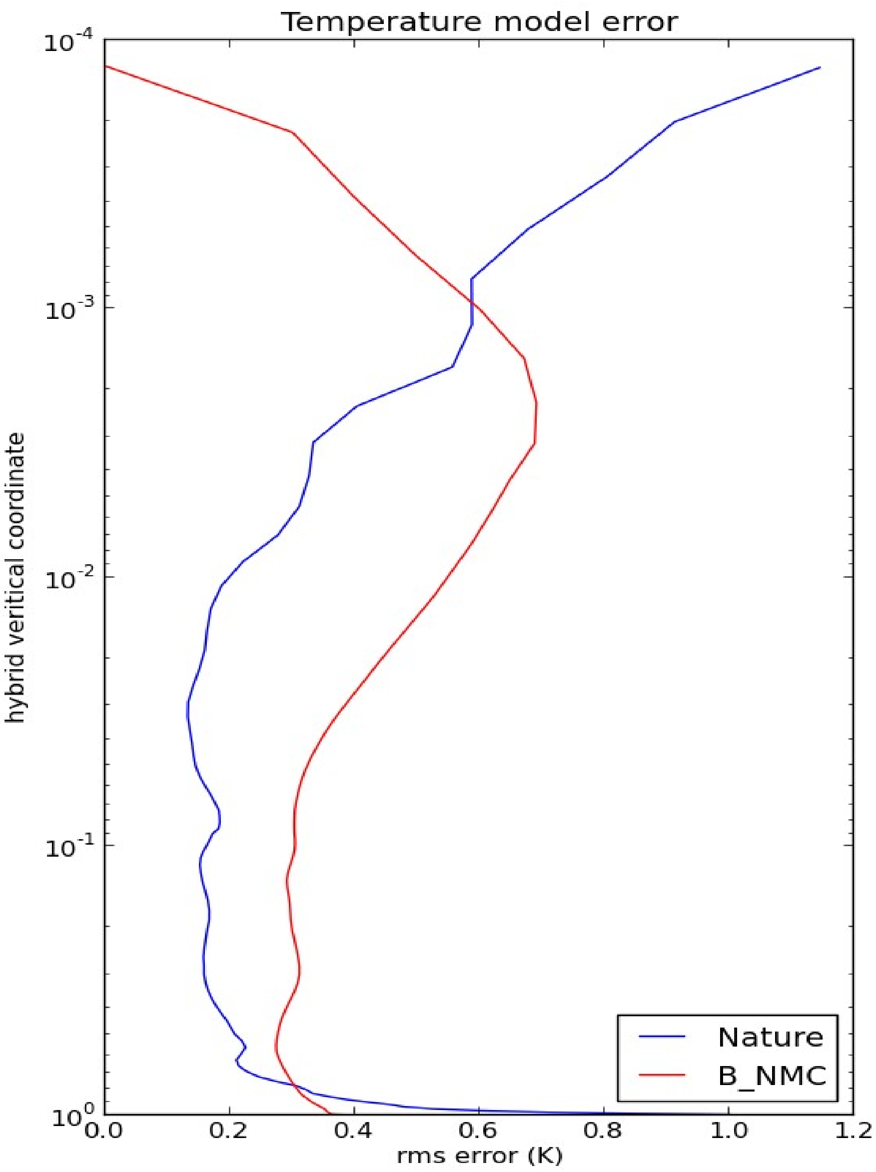


Note anti-correlation with previous map (suggestion an ITCZ intensity error).

RMS 6h temperature model error for January 2006



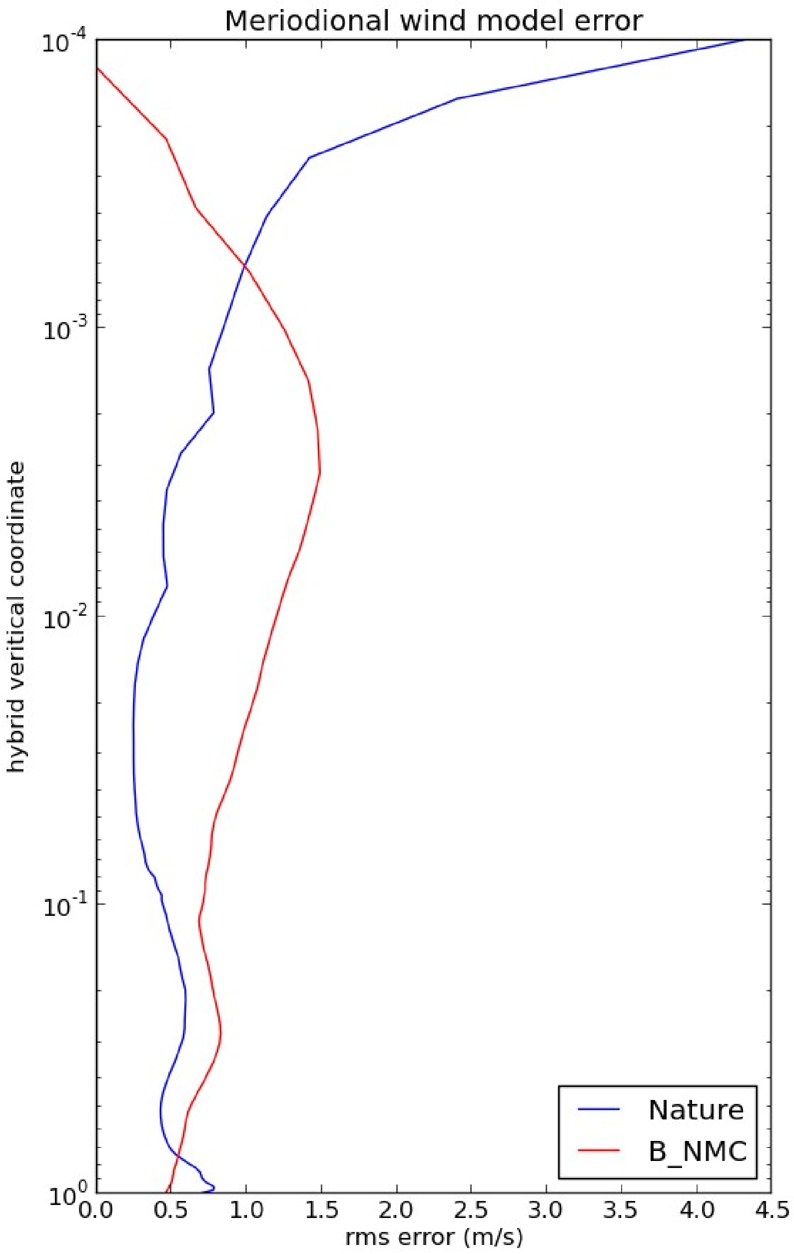
Issues are mainly over continents and with stratocumulus areas (south east Pacific).



Comparison with the B_NMC matrix

In the operational EnKF, the model error is assumed to be proportional to the B_NMC matrix. The B_NMC matrix has been obtain using differences between 24h and 48h forecasts that are valid at the same time.

The work with the nature run suggests that true error amplitudes are larger **near the surface** (below 800 hPa) as well as **above 1 hPa**. For intermediate levels, the approach based on B_NMC seems to give error amplitudes that are too large.

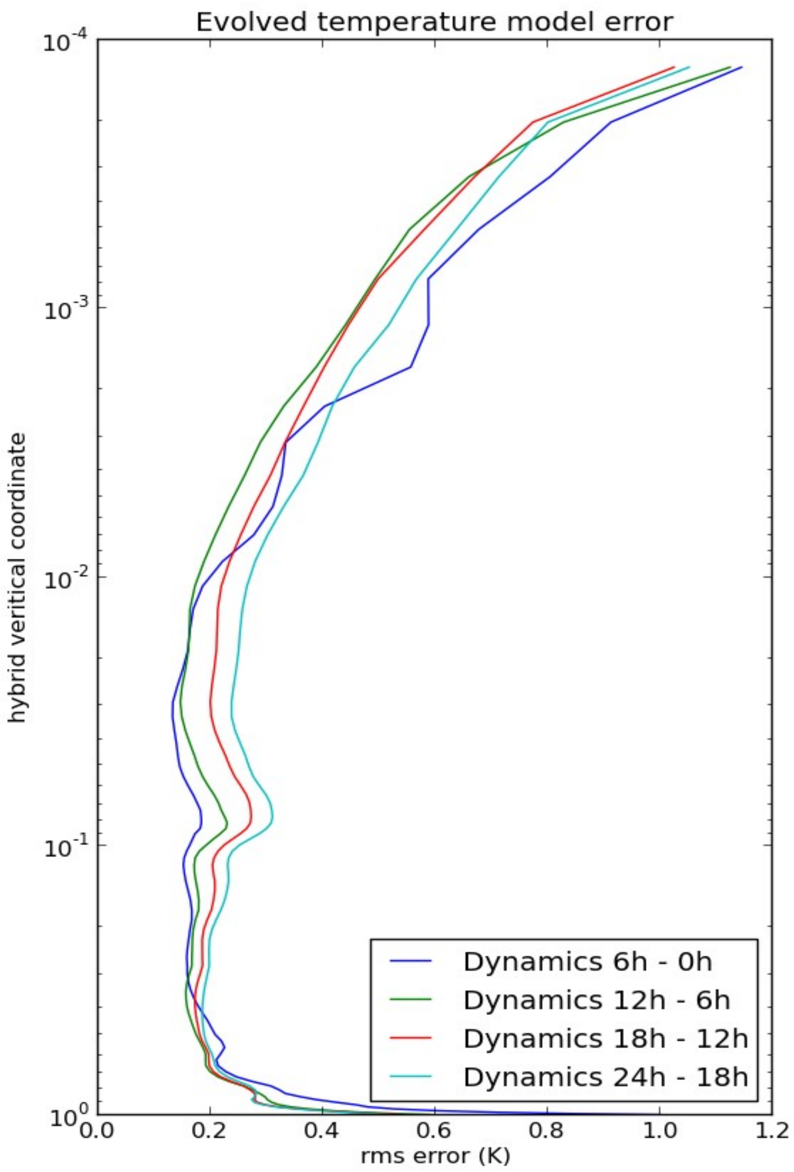


Comparison with the B_NMC matrix

Similar results were obtained for meriodional wind.

Note that the B_NMC matrix tapers to zero near the model top. This is done to avoid having analysis increments near the model top. It is thought preferable to let the model relax to its own climate near the model top.

In the atmospheric boundary layer, we have previously sought to have more spread in the EnKF by using multi-physics options.



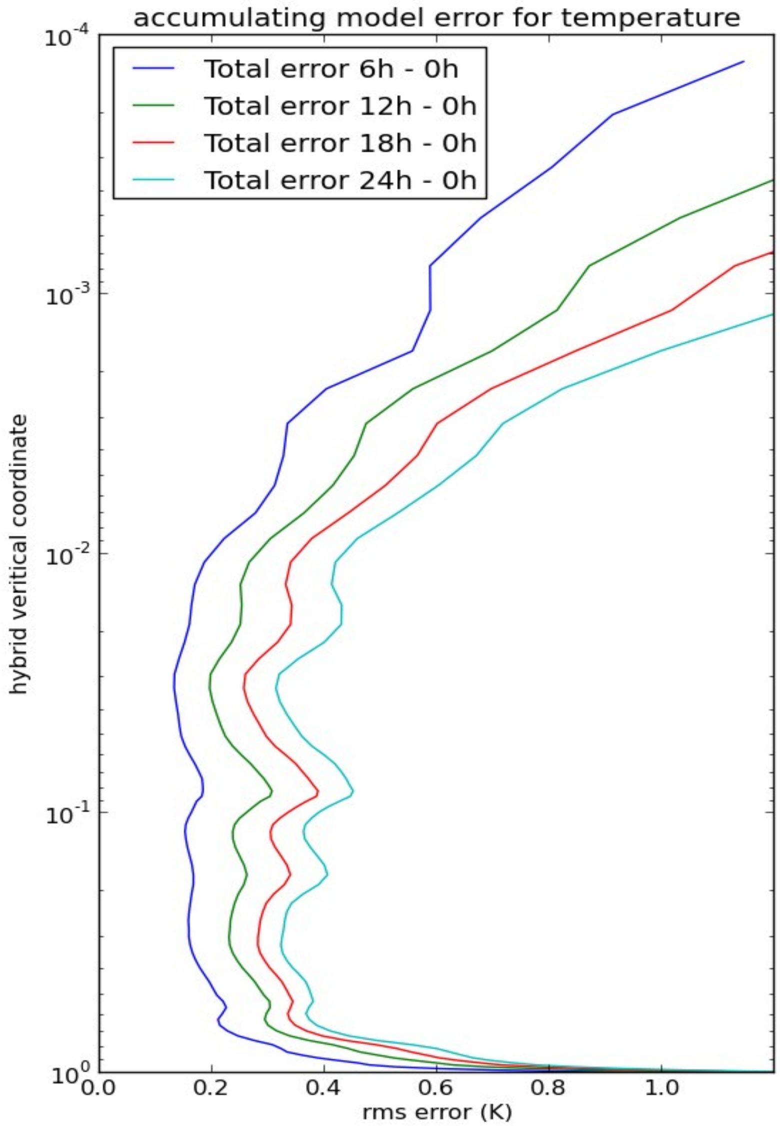
Evolving errors with GEM dynamics

For these experiments, the GEM model is used to evolve the 6h model error field. Thus the 6h forecast and the verifying state from the NR are both evolved with the GEM model dynamics.

As the model error sources in the atmospheric boundary layer and in the stratosphere are shut off, in these areas error amplitudes reduce. *The errors evolve towards dynamically unstable structures and grow between 600 and 2 hPa.*

The same would be expected if such error structures were to be used in an EnKF to sample model error.

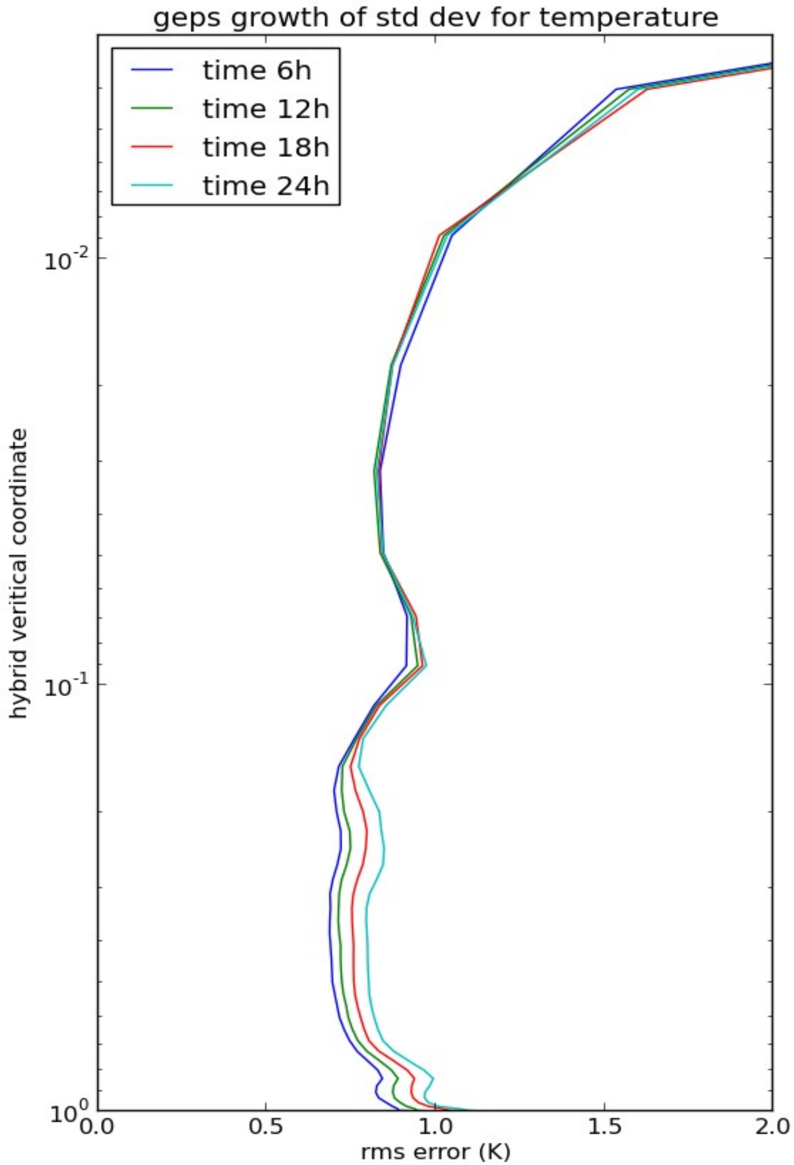
Divergence from the Nature Run.



For these experiments, longer range GEM integrations are verified against the nature run.

As the model error sources in the atmospheric boundary layer and in the stratosphere continue to be present, difference in these area continue to grow.

It could be considered how to sample similar differences in the Ensemble Prediction System.



Error growth in the global EPS.

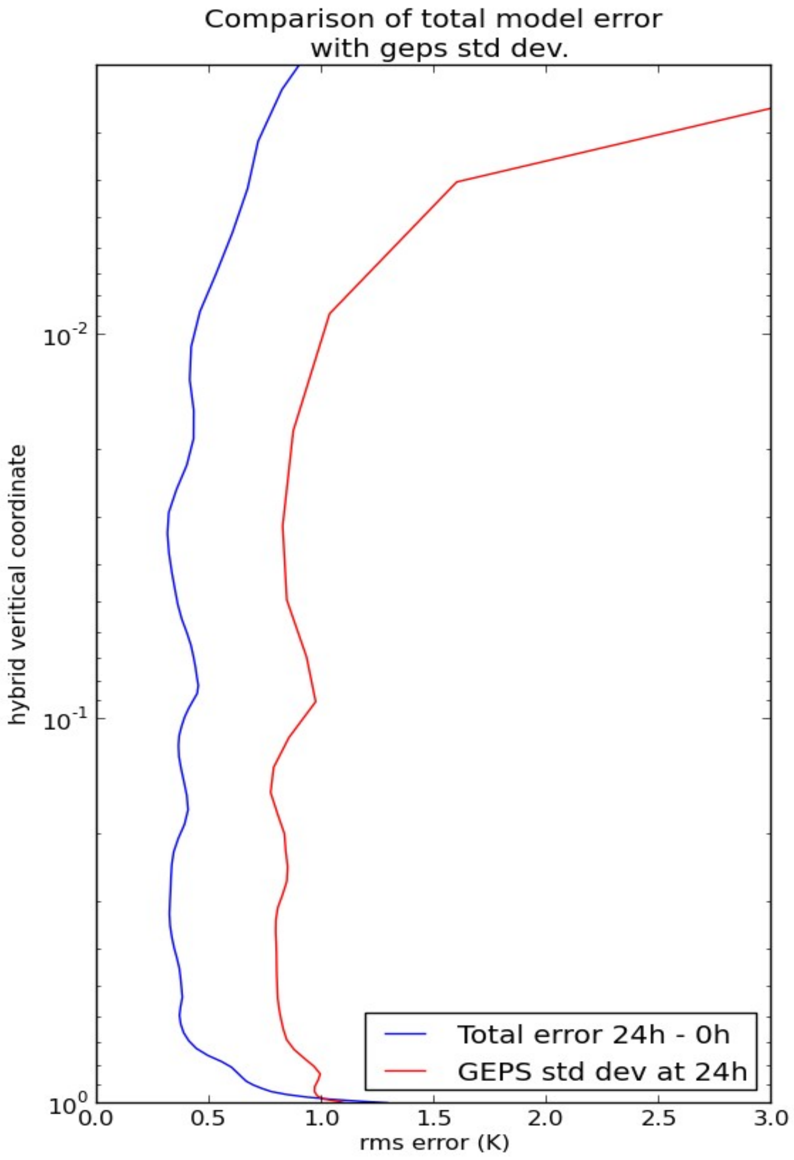
Ensemble std dev is shown for the first day of the operational ensemble forecast starting Jan 10 2017.

In the troposphere, the error growth is similar to the observed divergence from the NR.

In the stratosphere, std dev does not grow during the first forecast day.

Hypothesis : the errors are associated with unresolved topography and they are not currently simulated in the GEPS.

Relative importance of model error



For comparison, we show the temperature std dev in the operational global EPS at 24h (on Jan 10 2017).

At 24h, the error that is only due to the model, can explain about half of the required error amplitude.

Note that estimating model error involved an uncertain scaling by root(2).