Ensemble Forecasting at the Meteorological Service of Canada: Status, Research, and Plans

Collaborators:

Peter Houtekamer

Louis Lefaivre

Lubos Spacek

Bin He

Guillem Candille

Herschel Mitchell

Gérard Pellerin

Chantal Côté

Xiaoli Li

Martin Charron

Talk Outline

- ***Currently operational at the global scale***
  - *Ensemble data assimilation (EnKF)*
  - *Medium range ensemble forecasting*
  - *Improvements in the EnKF and global forecast system*
- ***Collaboration with NCEP***
  - *North American super-ensemble (NAEFS)*
- ***Research and plans***
  - *Short range ensemble forecasting*

Ensemble Data Assimilation

- *The ensemble Kalman filter has been operational since January 2005*
  - *2x48 members (we use two ensembles)*
  - *Impact of (perturbed) observations is localized to 2800 km in the horizontal, and to ~15 km in the vertical*
  - *Data are supposed valid at the synoptic times*

The Ensemble Kalman Filter

$$X_a^{\alpha,i} = X_b^{\alpha,i} + K^{\beta}\left(y^{\alpha,i} - HX_b^{\alpha,i}\right)$$

$$K^{\beta} = \rho \circ \left(P^{\beta}H^{T}\right)\left(\rho \circ \left(HP^{\beta}H^{T}\right) + R\right)^{-1}$$

$$P^{\beta}H^{T} = \frac{1}{N-1}\sum_{i=1}^{N}\left(X_b^{\beta,i} - \overline{X_b^{\beta}}\right)\left(HX_b^{\beta,i} - H\overline{X_b^{\beta}}\right)^{T}$$

$$\overline{X_b^{\beta}} = \frac{1}{N}\sum_{i=1}^{N} X_b^{\beta,i}$$

ρ *is the localization covariance matrix*

$\mathbf{X}^{\alpha,i}_{\mathbf{a}}$ *is one analysis in ensemble* α

$\mathbf{X}^{\alpha,i}_{\mathbf{b}}$ *is one perturbed background field in ensemble* α

$y^{\alpha,i}$ *is one perturbed observation in ensemble* α

R *is observational error covariance matrix*

N *is the number of members per ensemble*

The numerical model used for background:

- *Global Environmental Multiscale (GEM) model (Côté et al., 1998)*
- *Grid point model*
- *Very similar to model used for deterministic medium range weather forecasts*
- *300x150 horizontal grid points*
- *28 vertical levels with top at 10 hPa*
- *The background fields are obtained from adding a 6 hour model prediction and a model error term (to be changed)*

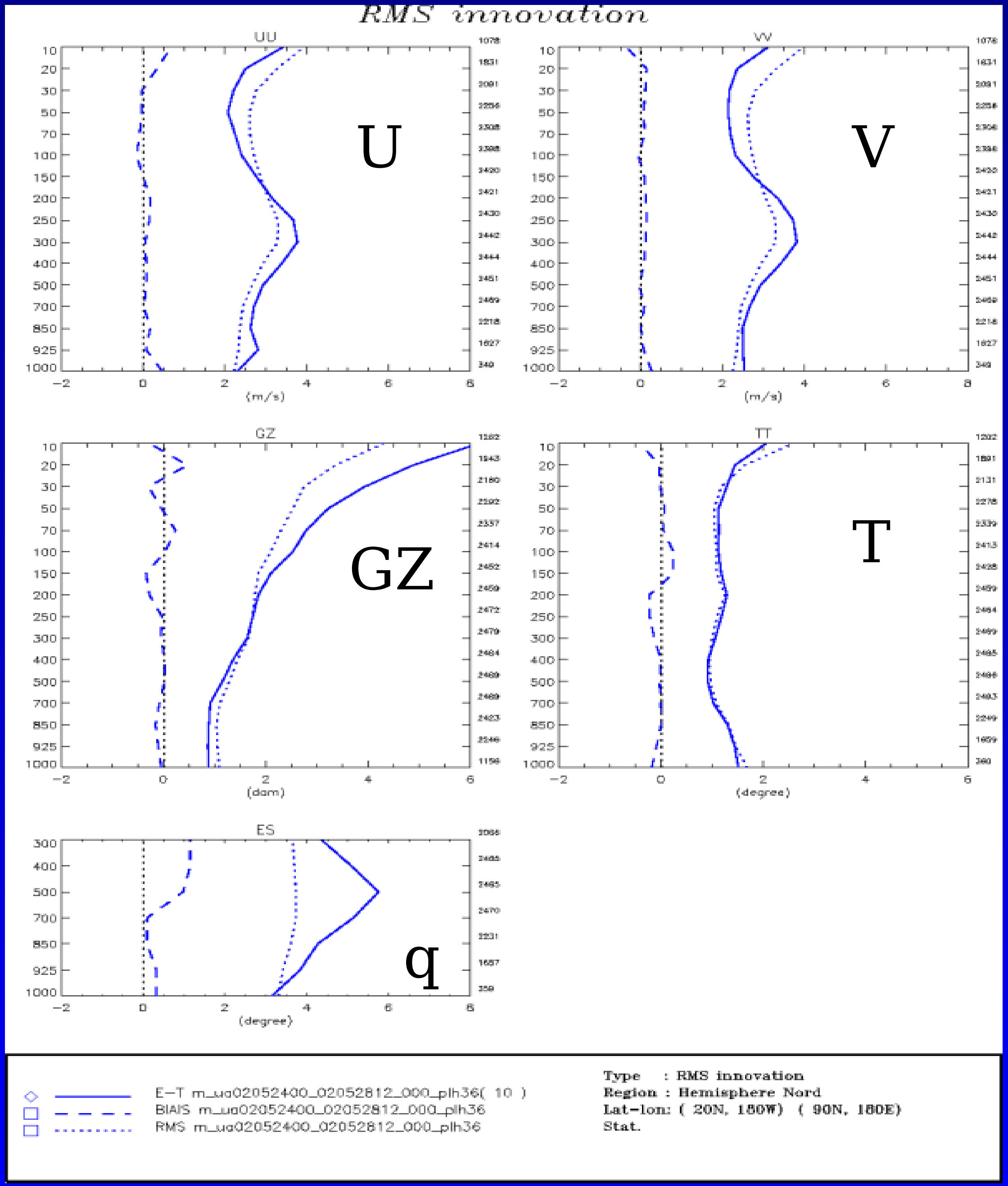
The "model error" component:

- *Hypothesis: the model error is similar in structure to the forecast error used in our centre's 4D-VAR*
- **$P(t+6h) = M\,P_a\,M^T + 0.25\,P_{4D\text{-}VAR}$**
- *Isotropic random error statistics for each member*
- *Currently the model error term includes:*
  - *A balanced component for wind, temperature and surface pressure*
  - *An unbalanced temperature component significant near the surface, in the tropics and near the top*

The observations:

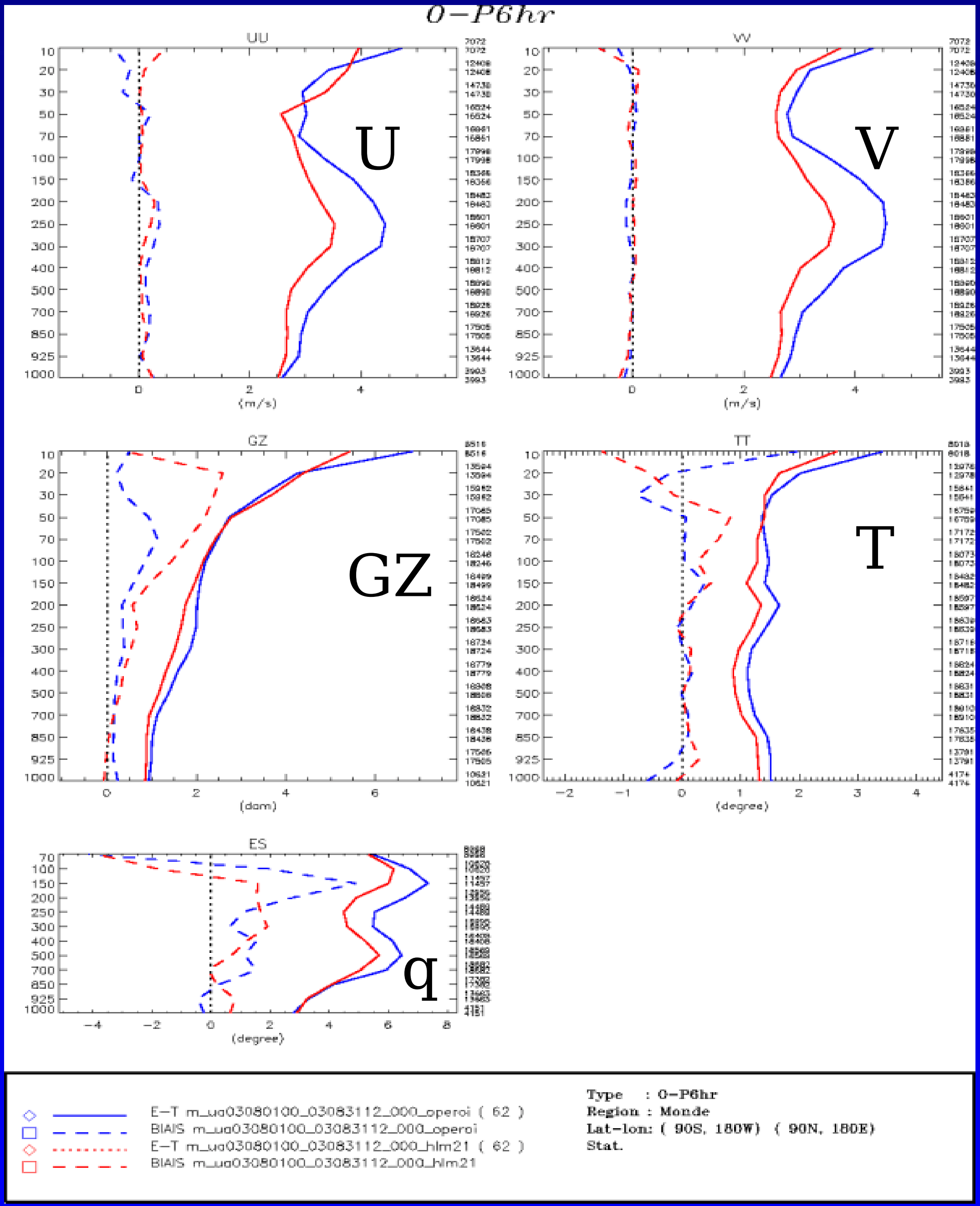
- *Try to assimilate same data as 4D-VAR*
- *Benefit from operational background check and variational quality control*
- *Same error statistic for marix* R *as in 4D-VAR*
- *Currently, we assimilate:*
  - *radiosondes: u, v, T, q, and surface pressure*
  - *aircrafts: u, v, and T*
  - *satellites: cloud track winds u, v, and AMSU-A radiances*
  - *surface observations: T, and surface pressure*
- *Surface humidity not yet assimilated*

Quality of error statistics



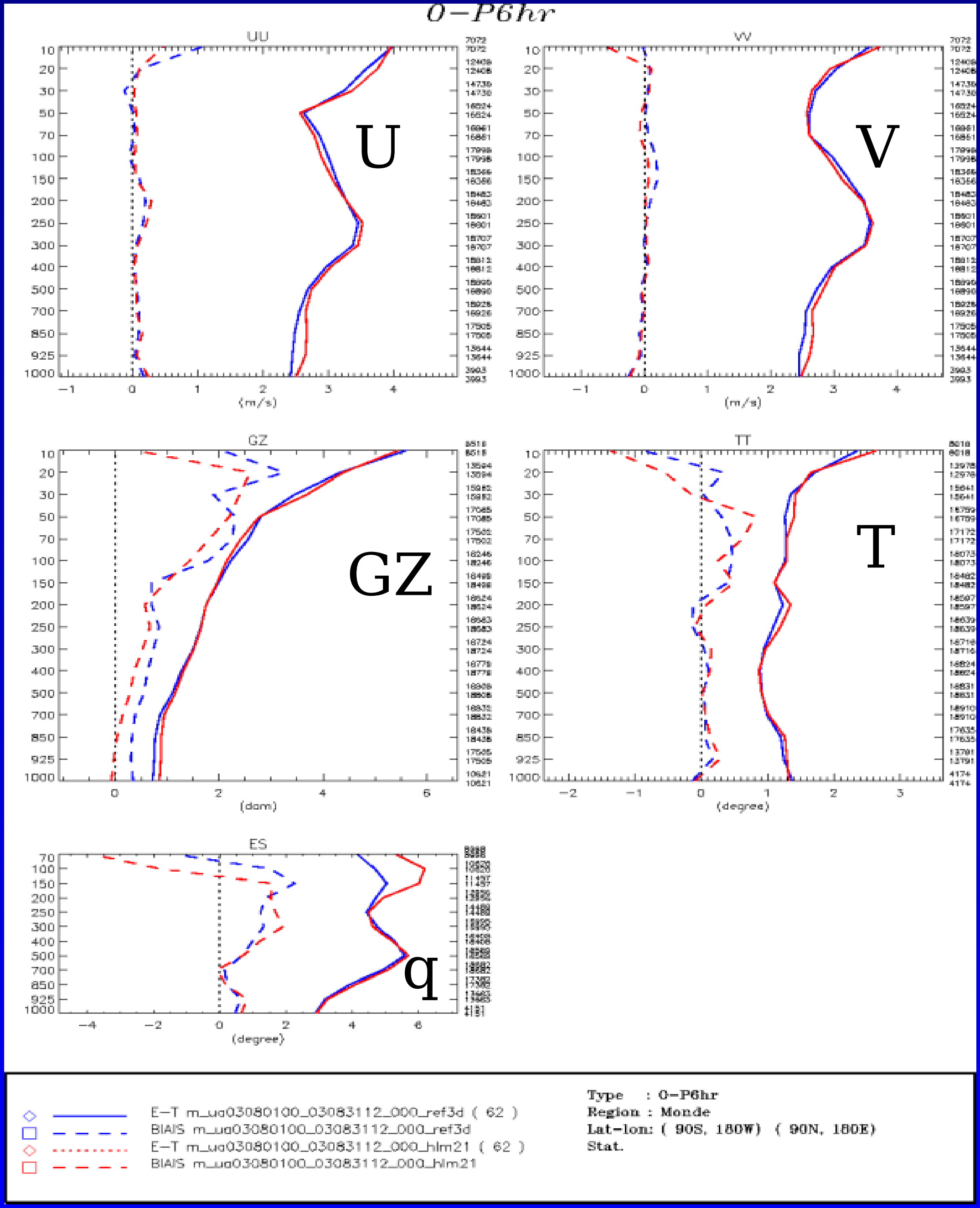
- *The solid line is the rms amplitude of the innovations corresponding with radiosondes.*
- *The dashed line is the ensemble based prediction of the innovation amplitude. It is the root of the sum of observational variance and ensemble spread.*
- *There is excellent agreement for the temperature. For winds the ensemble spread is too large near the model top. The spread is too small for humidity.*

Comparison with optimal interpolation



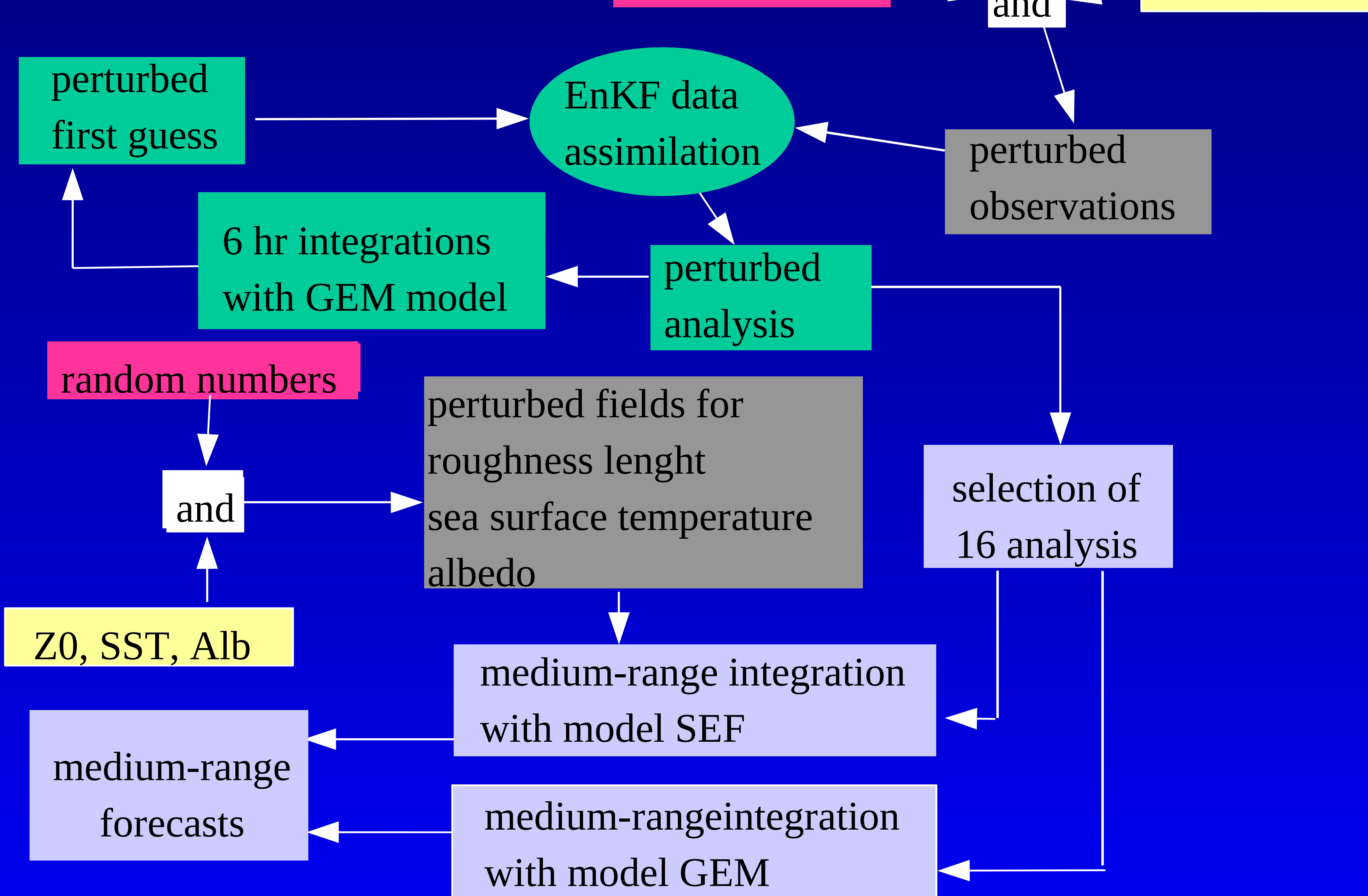
- *The solid line is the rms amplitude of the innovations corresponding with radiosondes.*
- *The dashed line is the bias*
- *Red: EnKF*
- *Blue: Optimal Interpolation*
- *Scores are for summer*

Comparison with 3D-VAR

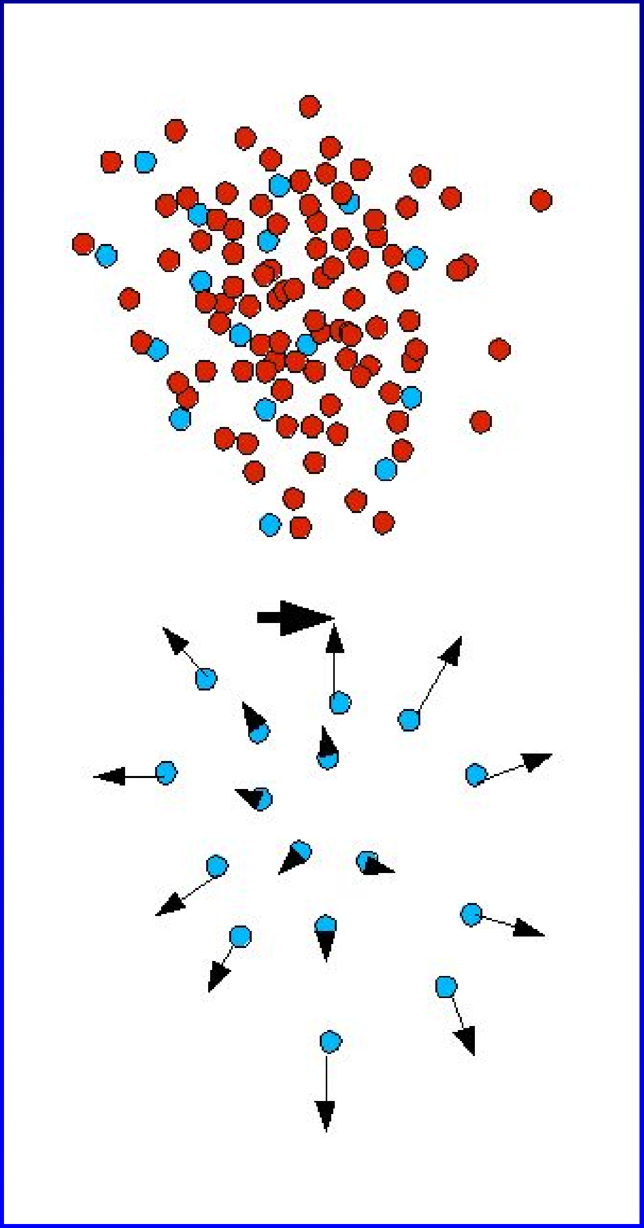


- *The solid line is the rms amplitude of the innovations corresponding with radiosondes.*
- *The dashed line is the bias*
- *Red: EnKF*
- *Blue: 3D-VAR*
- *Scores are for summer*

ENSEMBLE SET-UP



From an Ensemble of 96 Initial Conditions to an Ensemble of 16 10-Day Forecasts



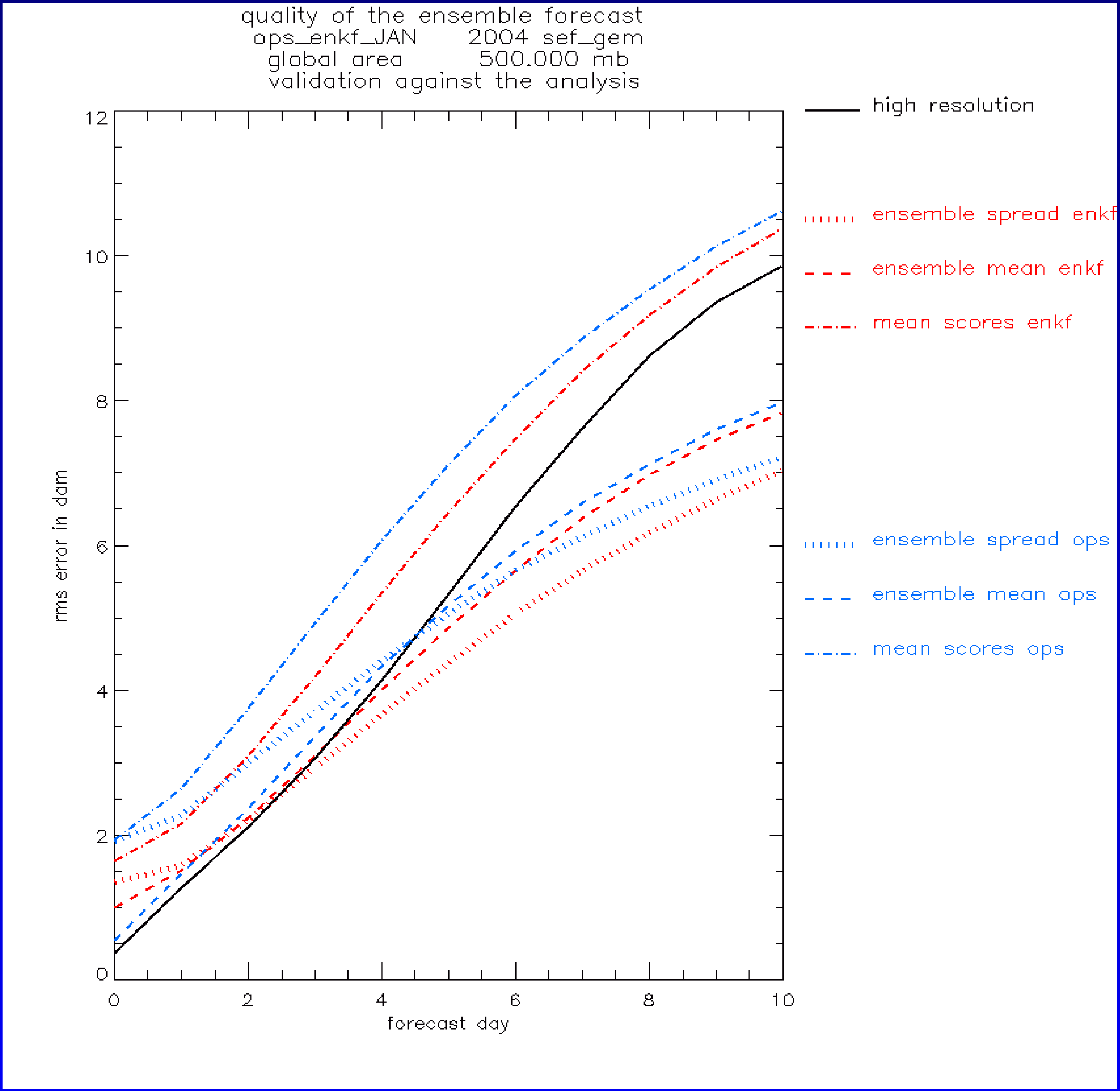
- *Choose randomly 16 members out of 96, preserving the ensemble mean.*
- *Inflate the ensemble standard deviation by a factor of 1.5 for u, v, T, q, and p_{surf}*
- *Correct q to avoid supersaturation and negative specific humidity*

Medium Range Ensemble Forecasting at CMC

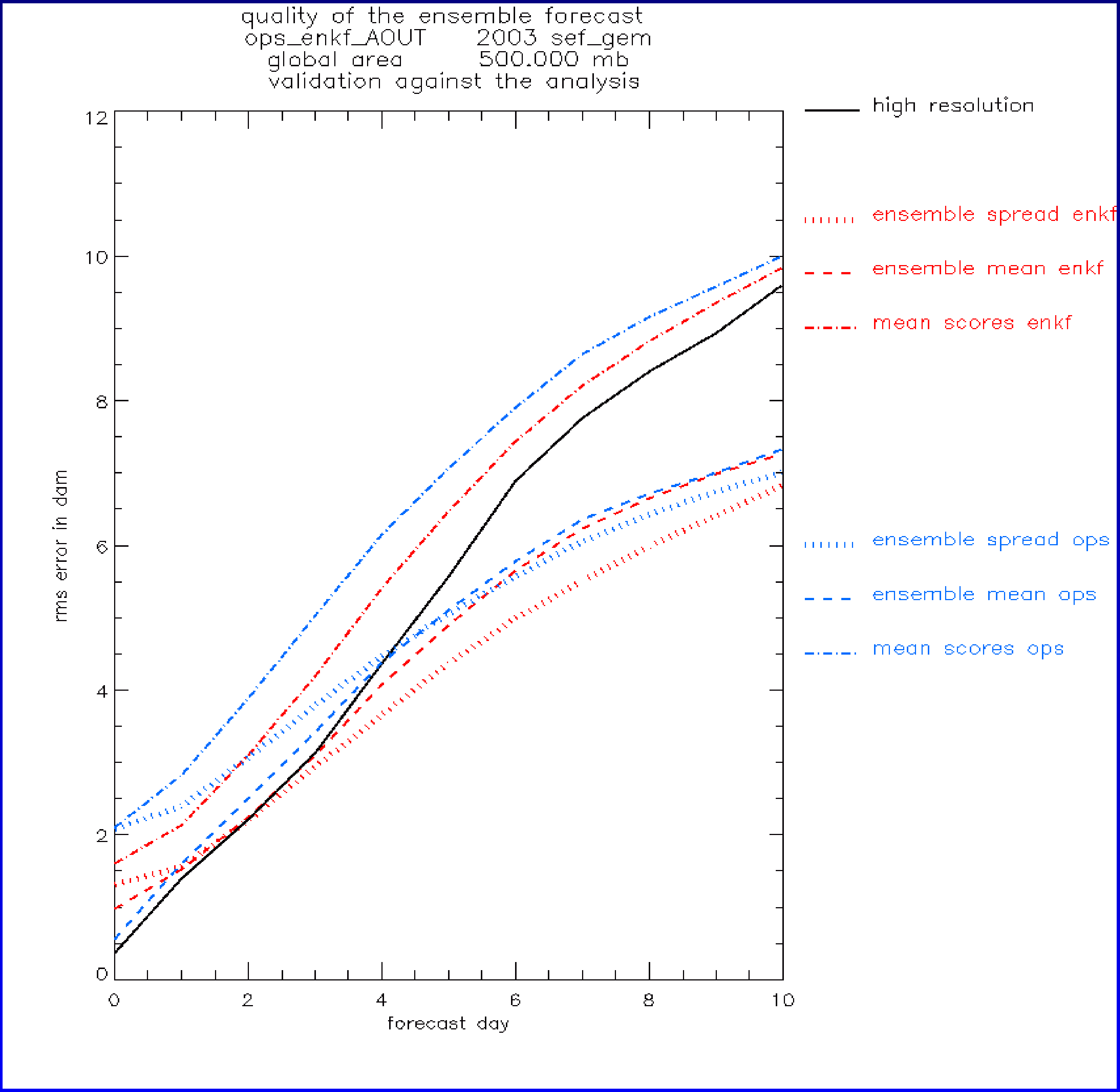
- *Use of multi-model multi-parameterization approach*
  - *This is our current way of representing the “model error” component of the 10-day forecasts*
  - *The initial condition uncertainties are provided by the EnKF*
- *8 members from the GEM model (grid point) at 300x150*
- *8 members from the SEF (spectral finite element) at T149*

Description of the different models

SEF (T149)	Radiation scheme	Convection Schemes deep	Convection Schemes shallow	Surface scheme	Number of levels	Time level
Control	Garand	Kuo	conres	Fcrest	27	3
1	Garand	Kuo	conres	ISBA	27	3
2	Garand	Ras	ktrsnt	Fcrest	27	3
3	Garand	Kuo	conres	Fcrest	27	3
4	Garand	Ras	ktrsnt	ISBA	27	3
5	Garand	Ras	ktrsnt	Fcrest	27	2
6	Garand	Kuo	conres	ISBA	27	2
7	Garand	Ras	ktrsnt	ISBA	27	2
8	Garand	Kuo	conres	Fcrest	27	2
GEM (1.2^0)	**Radiation Scheme**	**Convection Schemes deep**	**Convection Schemes shallow**	**Surface scheme**	**Number of levels**	**Time level**
9	Garand	Kuosym	ktrsnt	Fcrest	28	2
10	Garand	Ras	conres	ISBA	28	2
11	Garand	Ras	conres	Fcrest	28	2
12	Garand	Kuosym	ktrsnt	ISBA	28	2
13	Garand	Kuostd	ktrsnt	Fcrest	28	2
14	Garand	Kuostd	ktrsnt	ISBA	28	2
15	Garand	Kuosym	conres	ISBA	28	2
16	Garand	Kuo	conres	Fcrest	28	2



quality of the ensemble forecast
ops_enkf_JAN 2004 sef_gem
global area 500.000 mb
validation against the analysis
high resolution
ensemble spread enkf
ensemble mean enkf
mean scores enkf
ensemble spread ops
ensemble mean ops
mean scores ops
rms error in dam
forecast day
0
2
4
6
8
10
12

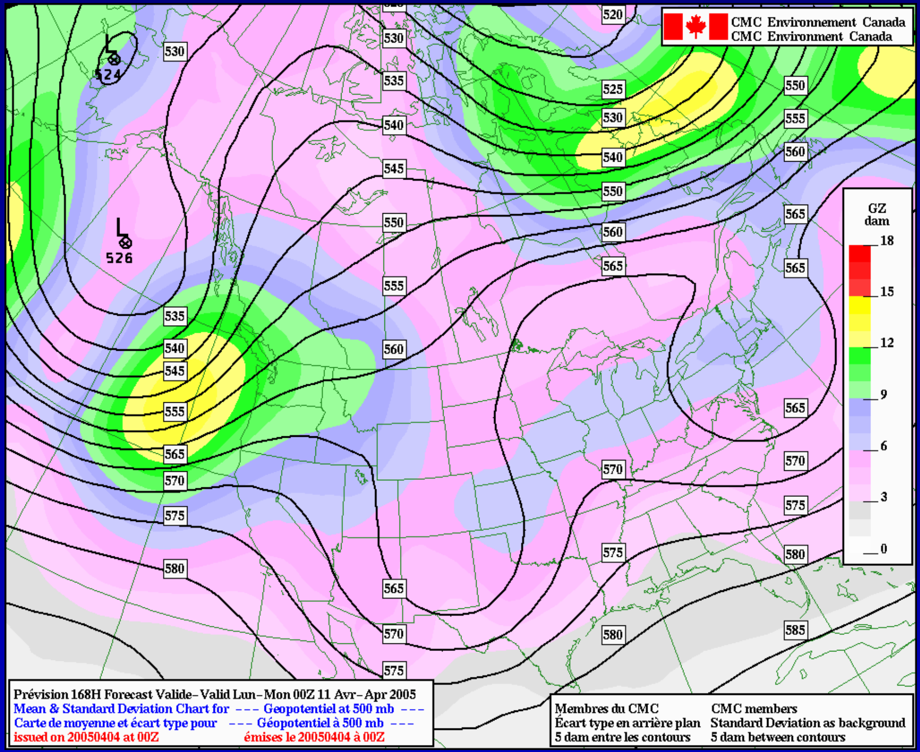

quality of the ensemble forecast
ops_enkf_AOUT 2003 sef_gem
global area 500.000 mb
validation against the analysis
high resolution
ensemble spread enkf
ensemble mean enkf
mean scores enkf
ensemble spread ops
ensemble mean ops
mean scores ops
rms error in dam
forecast day
0
2
4
6
8
10
12

In Development

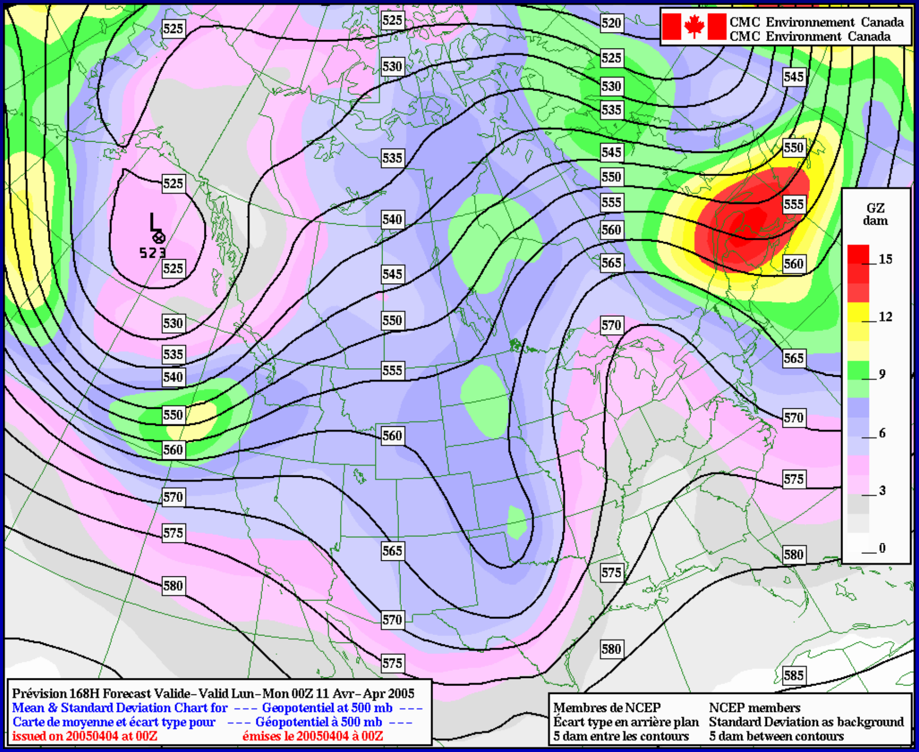
- EnKF
  - 4 ensemble configuration (allow weaker localization)
  - Later stage, assimilate data at non-synoptic times
- Forecast models
  - Test stochastic backscatter scheme based on Shutts
  - Reduce number of parameterizations
  - Extend forecasts to 15 days twice daily

The North-American Ensemble Forecast System (NAEFS)

- *Participants: MSC (Canada), NWS (USA), and NMSM (Mexico)*
- *High level agreement: Feb. 2003 (Mexico became involved in Oct. 2004)*
- *Product generation made routinely at MSC: early 2005 (in development mode, still not operational)*
- *Major tasks:*
  - *Exchange ensemble data between two centers (MSC and NWS)*
  - *Bias correction of each set of ensemble*
  - *Develop products based on joint ensemble (high impact weather?)*
  - *Verify joint product suite, evaluate added value*



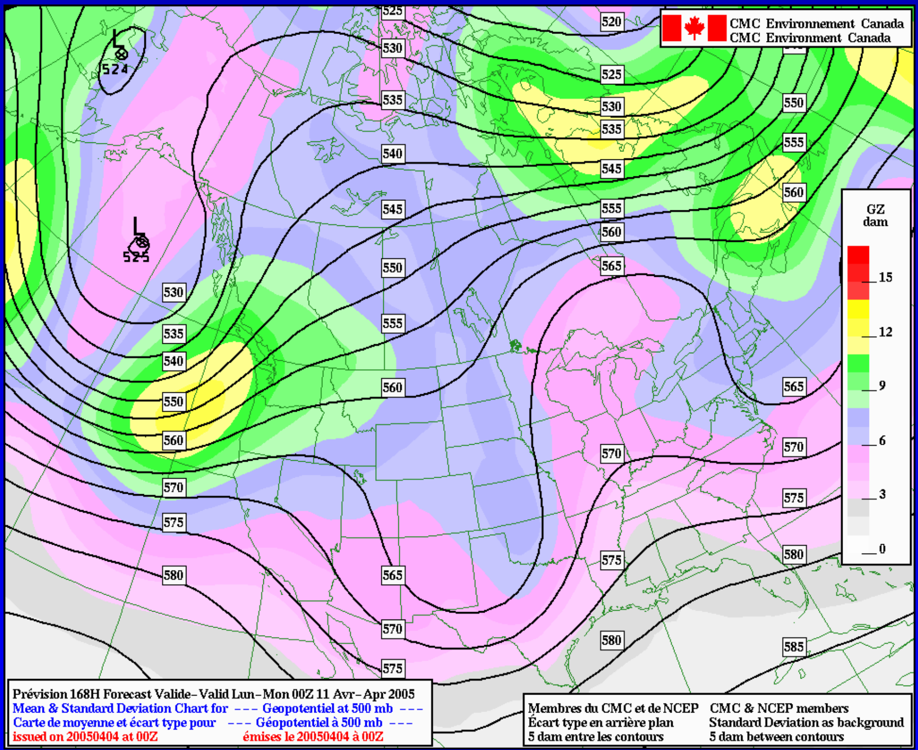
MSC



NWS

GZ at 500 hPa ensemble mean (black lines)

Ensemble standard deviation (color code)



7-day forecast over North America valid April 11 2005

Joint

Development of a Regional EPS

M. Charron (MSC), L. Spacek (MSC), Li Xiaoli (McGill)

- *Regional EPS based on targeted singular vectors*
  - *A version at 28 km resolution with 20 members over North America (LAM)*
  - *A version at 15 km resolution with 16 members over Eastern Canada (LAM)*

The 15 km REPS at MSC-McGill

- *8 singular vectors are calculated on a low resolution global grid (240x120, or about 150 km)*
- *Initial norm is global; final norm is located over a domain covering Eastern Canada (will be adapted to be over Beijing)*
- *Optimisation period is 24h and singular vector calculation includes*
  - *Vertical diffusion*
  - *Gravity wave drag*
  - *Deep Convection (Kuo)*
  - *Stratiform precipitation*
- *SVs are interpolated to the resolution of the pilot model (100 km res.)*
- *SVs are used to perturb the pilot runs producing lateral boundary and initial conditions of 16 LAM integrations*

Physics perturbations with Markov processes

- *Physical parameters/tendencies can be perturbed by a function F(λ,φ,η,t) given by:*

$$f(\lambda,\varphi,\eta,t)=\sum_{l=0}^{L}\sum_{m=-l}^{l}\sum_{k=0}^{K}a_{lmk}(t)Y_{lm}(\lambda,\varphi)e^{ik\eta}$$

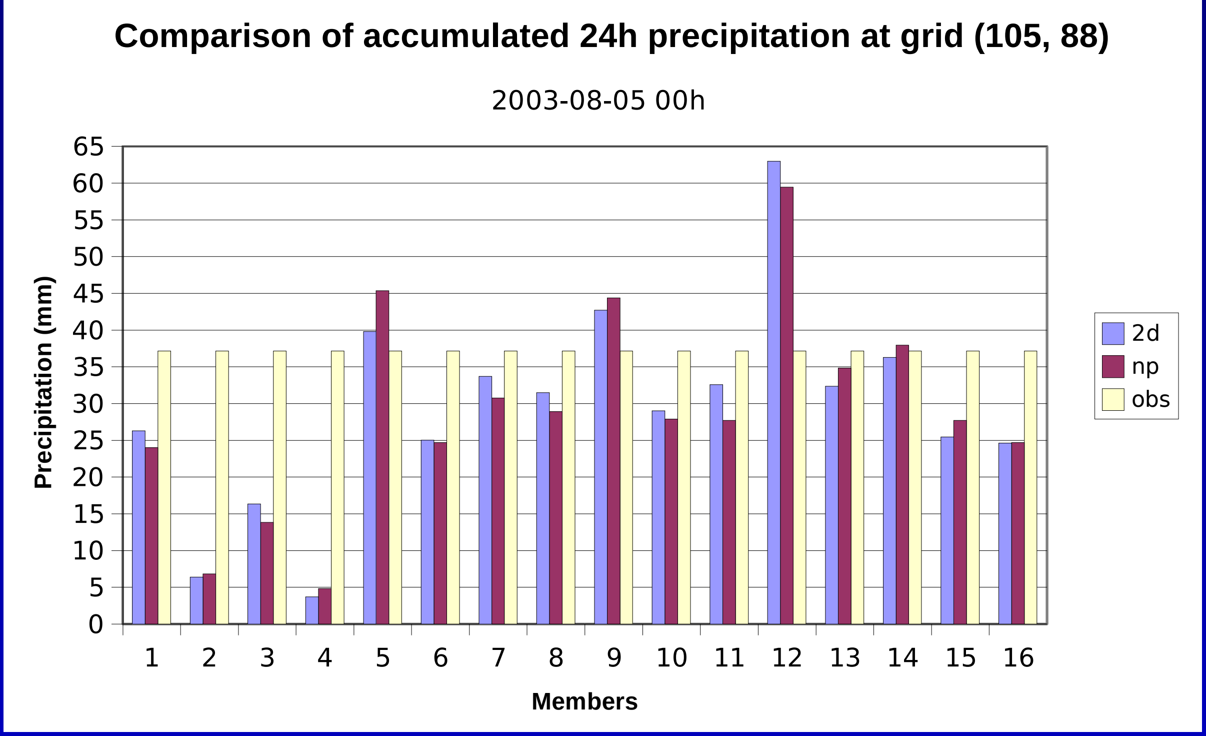
$$a_{lmk}(t)=e^{-\Delta t/\tau}a_{lmk}(t-\Delta t)+R(t)$$

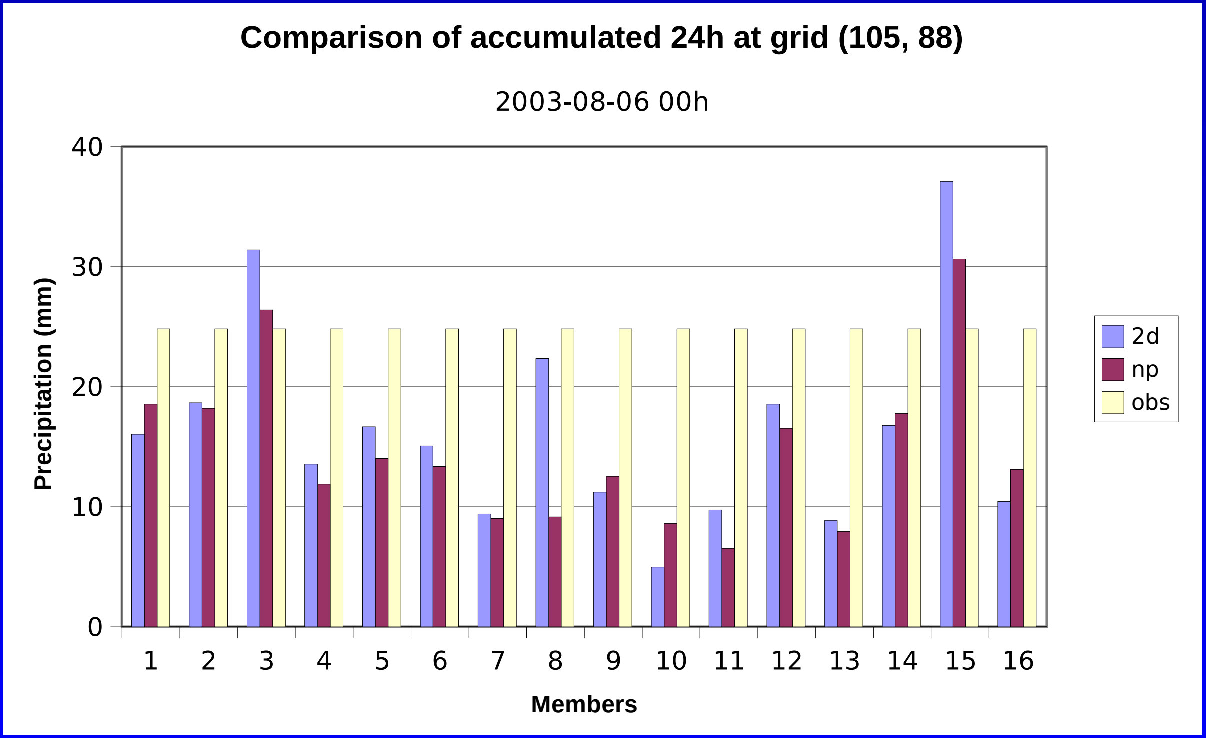
$$F(\lambda,\varphi,\eta,t)=Sf(\lambda,\varphi,\eta,t)$$

Perturbation of CAPE in the Kane-Fritsch convection scheme

- *Only the LAMs are perturbed with Markov chains*
- *CAPE perturbation similar to Lin and Neelin (GRL 2000), except*
  - *CAPE becomes CAPE**$F(\lambda,\varphi,t)$
- *Decorrelation time scale: 12 and 6 hours*
- *Truncation of the perturbed field: T7 and T14*

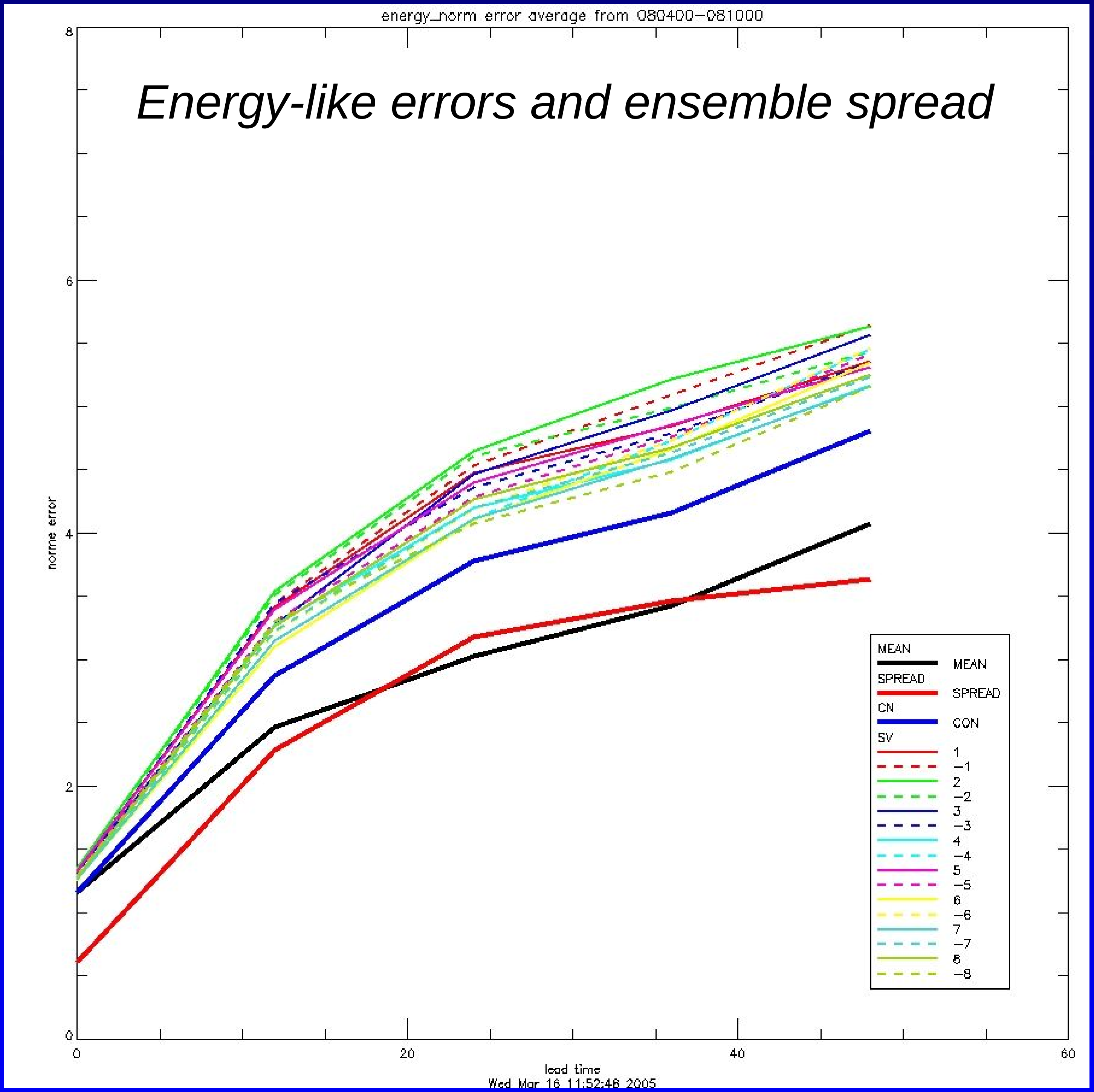
Impact of perturbing CAPE on precipitation





- ➔ *Perturbing the ICs with SVs has more impact on precip than perturbing CAPE*

Energy-like errors and ensemble spread

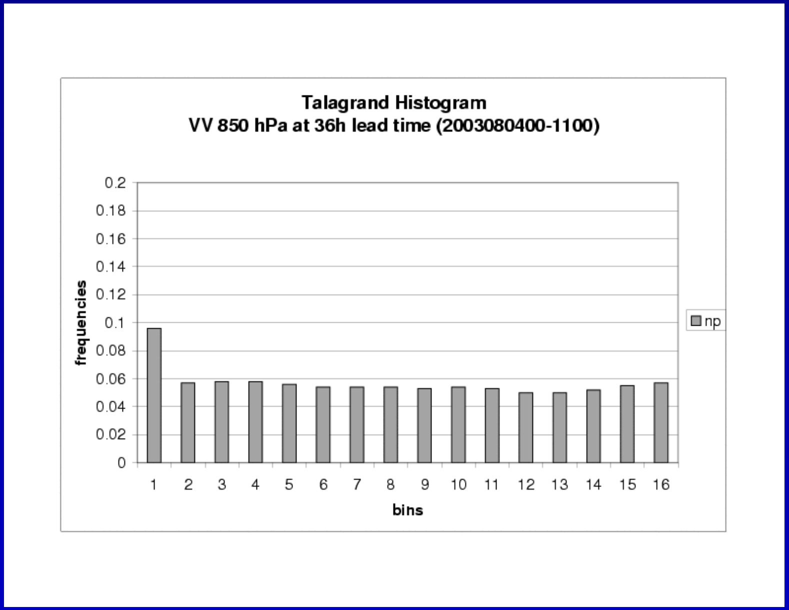


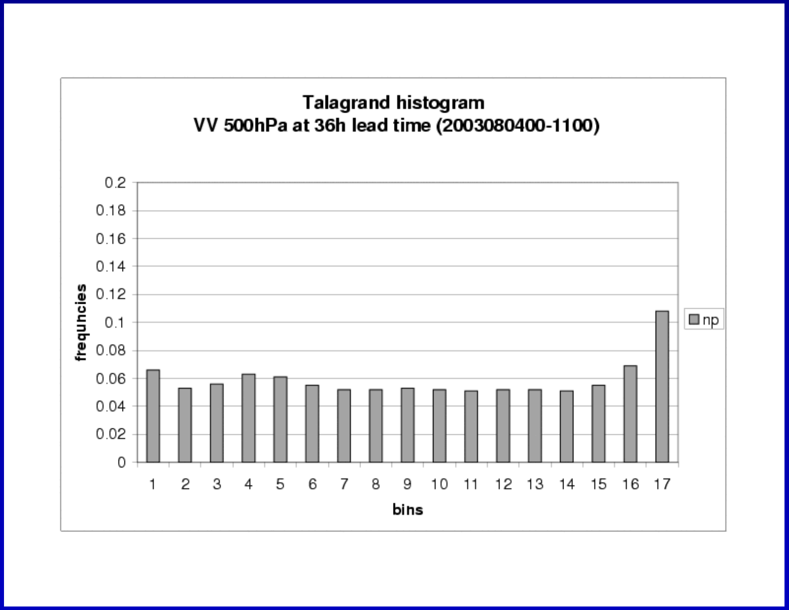
Blue: Unperturbed

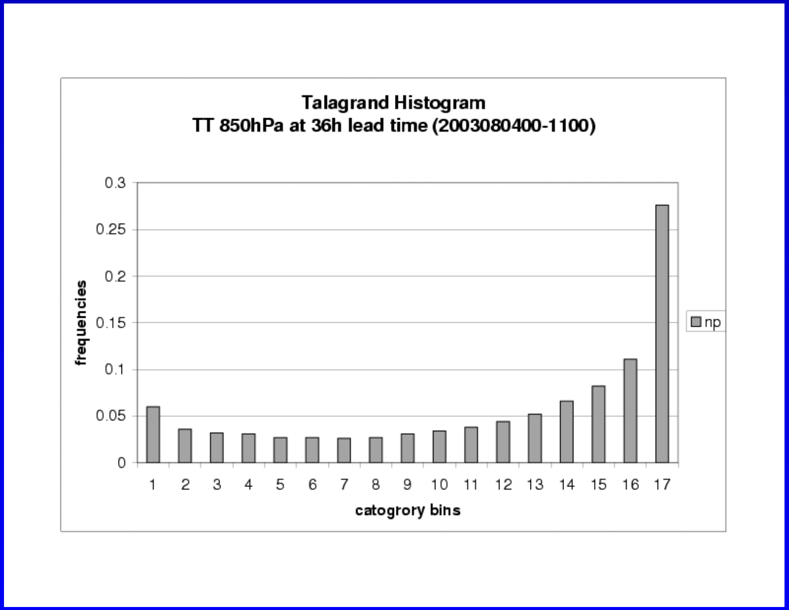
Black: Ens. Mean

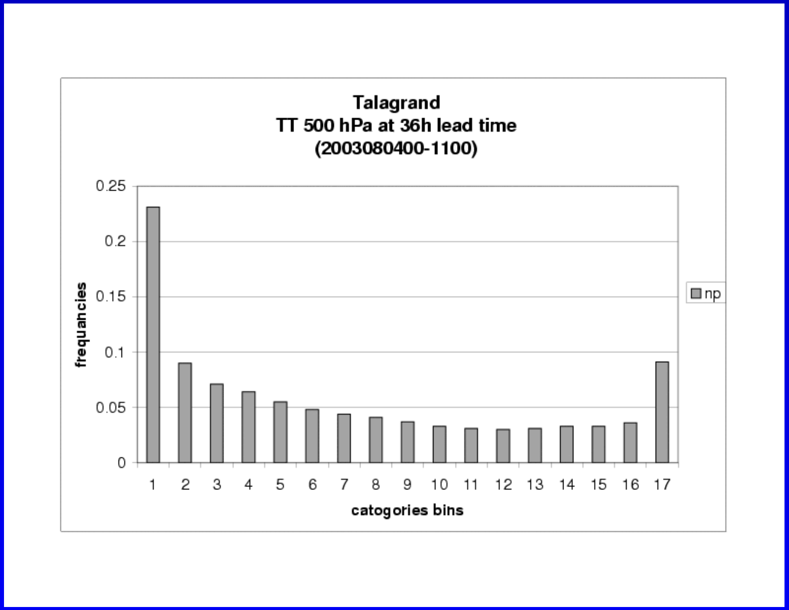
Red: Ens. Std Dev.

Talagrand Diagrams (August 4 to 11, 2003)

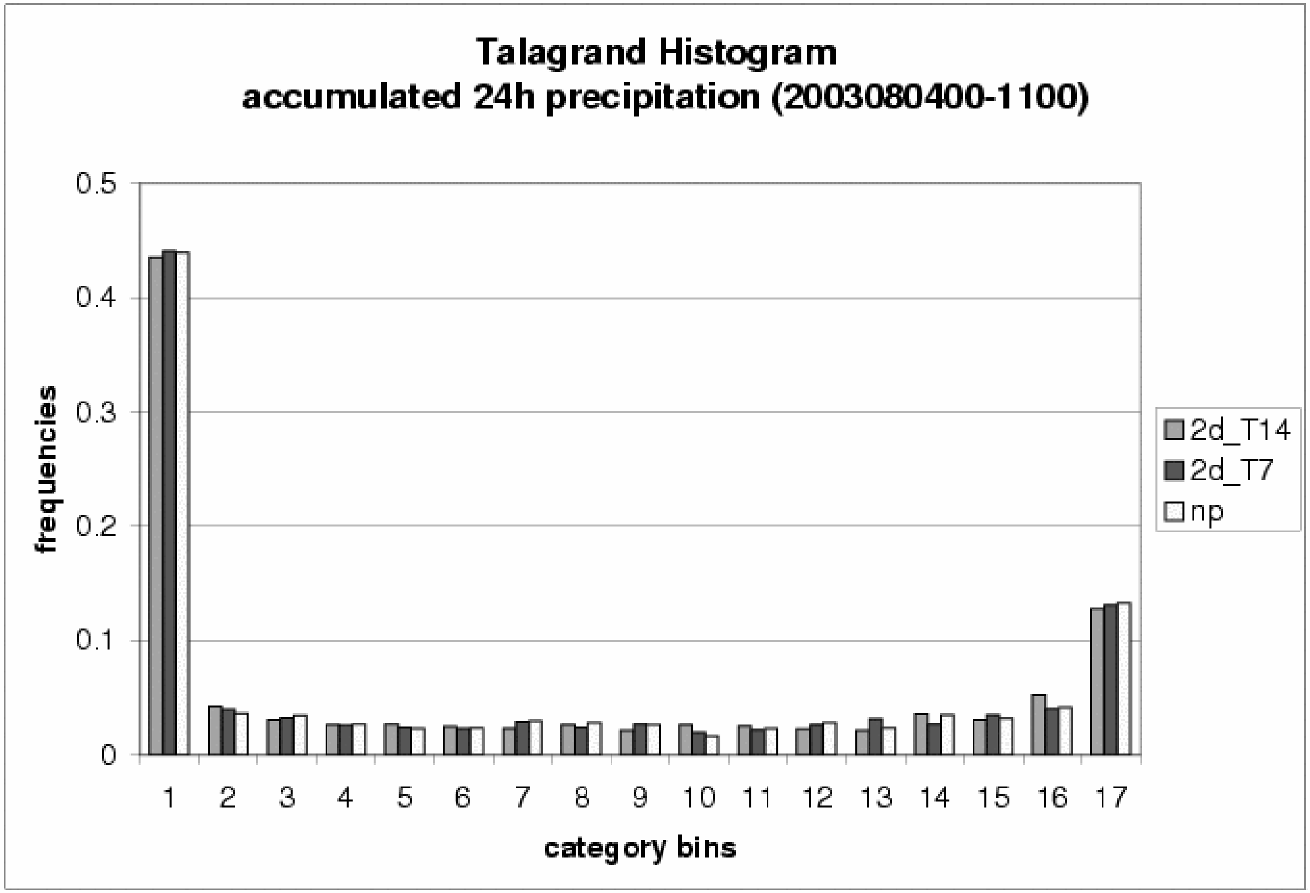




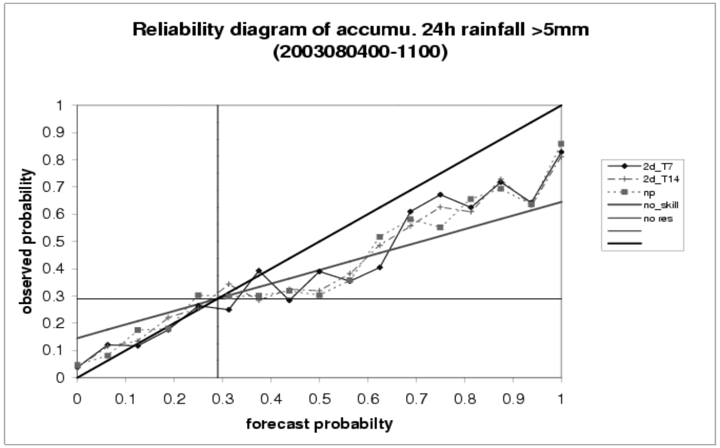


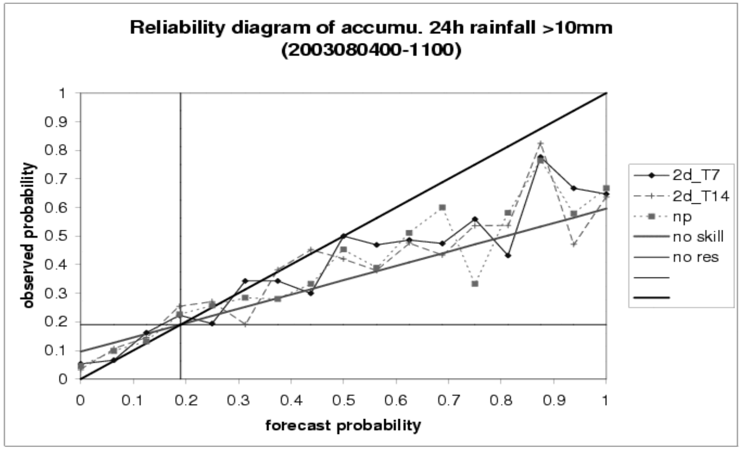


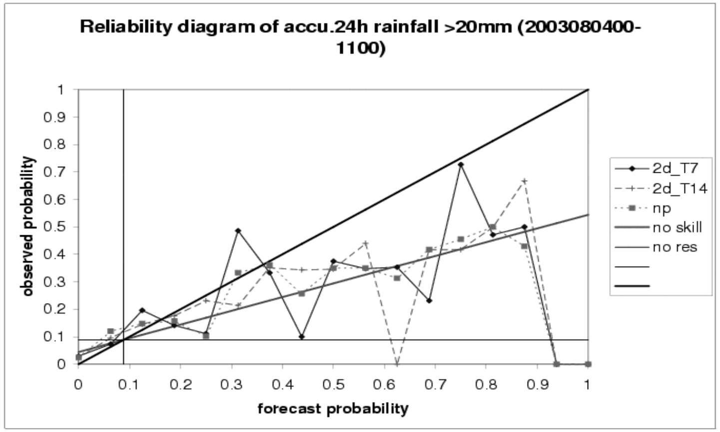
Talagrand Diagrams (August 4 to 11, 2003)

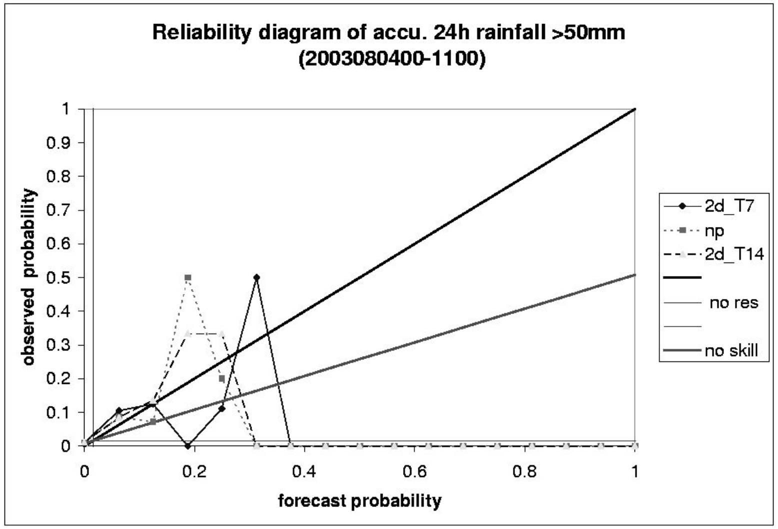


Reliability Diagrams for Precipitation









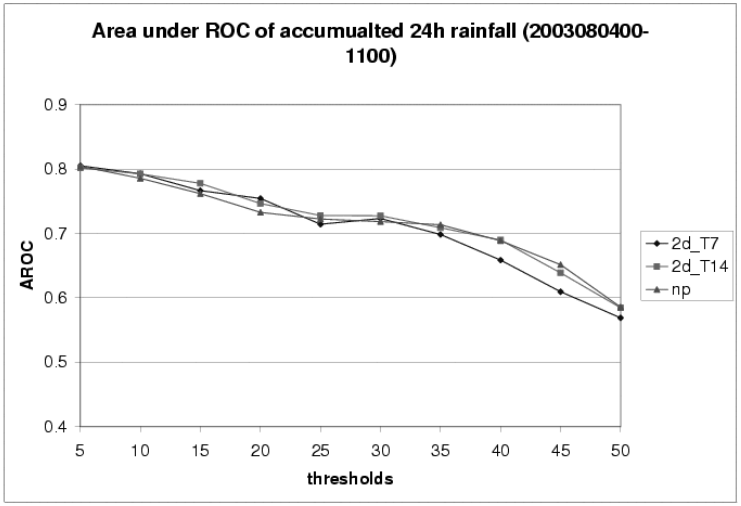
Relative Operating Characteristics (ROC) and Brier Skill Scores (BSS)

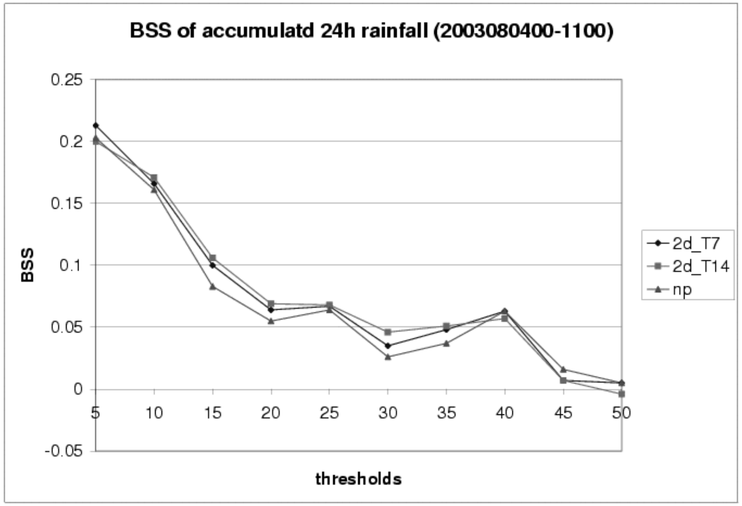
- *For precipitation:*
  - *Binary event: 24-hour accumulation (lead times from 12 to 36 hours) is greater than some specified thresholds*
  - *Dataset: 300 rain gauges over Québec*
- *For temperature:*
  - *Binary event: temperature is lower than mean temperature minus one standard deviation (different at each grid point and calculated over the studied period)*
  - *Dataset: 3D-VAR analyses at 15 km resolution*

Area under ROC

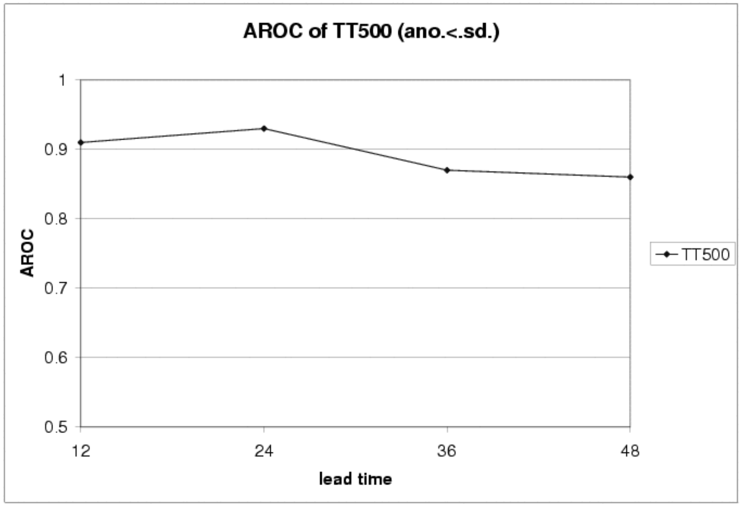
Brier Skill Score

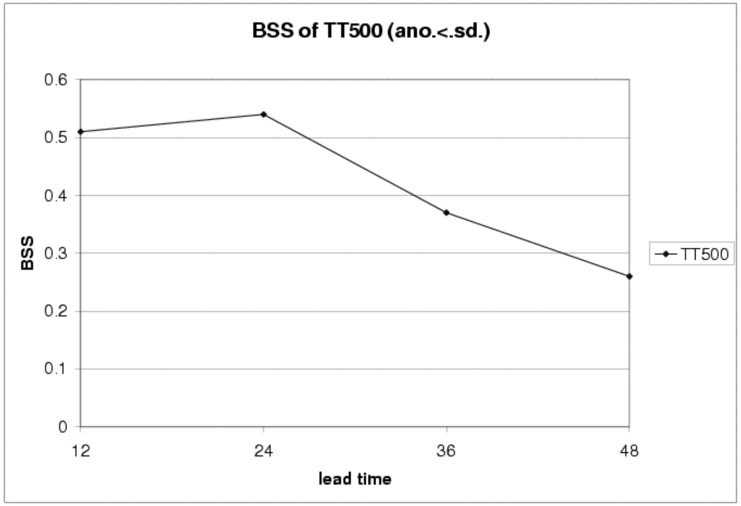
Precip. 24-h accum





Temperature





Long term objectives

- *Find a better way to account for uncertainties of the model, perhaps by introducing parameterizations that are inherently stochastic*
- *Develop a regional ensemble Kalman filter (stretched grid or limited area model)*
- *Compare the singular vector approach and a (still to be built) regional EnKF for regional ensemble predictions*