Comparison of the Ensemble Transform and the Ensemble Data Assimilation techniques for background error simulation in ALADIN

Gergely Bölöni and Edit Adamcsek

5-8 April, 2011
Norrköping

HIRLAM-ALADIN
workshop

Outline

- Introduction
- The simulation techniques in play (Ensemble Data Assimilation and Ensemble Transform)
- The LAM experiments
- Diagnostic comparisons
- Impact studies

Introduction

Aim: simulation of background errors ($\boldsymbol{\varepsilon}_b$) in order to generate a statistical sample for the computation of the background error covariance matrix ($\mathbf{B}$) in the variational analysis:

$$\mathbf{B} = \mathrm{E}\left(\boldsymbol{\varepsilon}_b \boldsymbol{\varepsilon}_b^T\right) \qquad \mathrm{J}_b(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b)$$

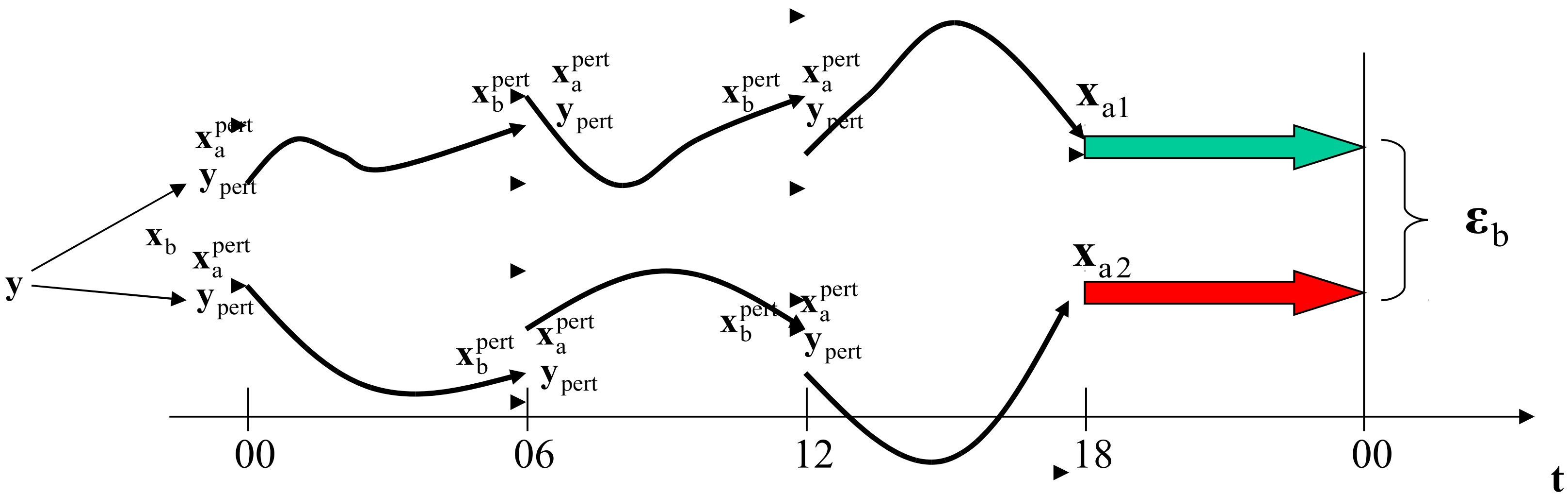
The simulation techniques in play

Background error simulation with EDA

$$\mathbf{x}_{b1} = M\,\mathbf{x}_{a1}$$
$$\mathbf{x}_{b2} = M\,\mathbf{x}_{a2}$$

$$\boldsymbol{\varepsilon}_{b} \approx \mathbf{x}_{b1} - \mathbf{x}_{b2}$$

(EDA: Ensemble Data Assimilation)



5-8 April, 2011
Norrköping

HIRLAM-ALADIN
workshop

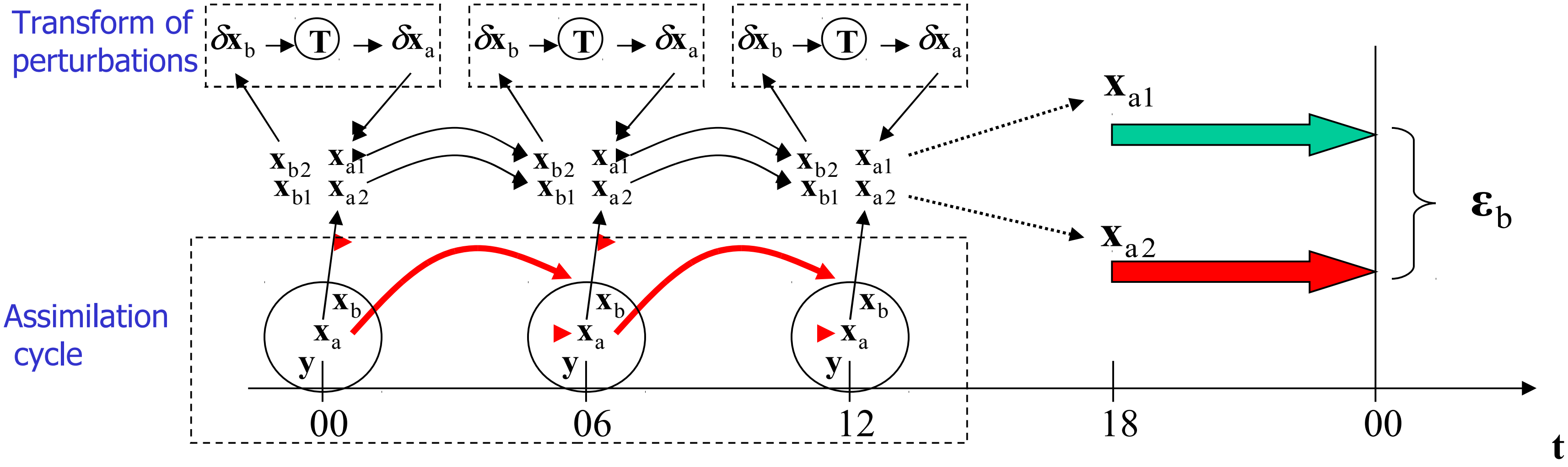
The simulation techniques in play

Background error simulation with ET

$$\mathbf{x}_{b1} = M\,\mathbf{x}_{a1}$$
$$\mathbf{x}_{b2} = M\,\mathbf{x}_{a2}$$

$$\boldsymbol{\varepsilon}_b \approx \mathbf{x}_{b1} - \mathbf{x}_{b2}$$

(ET: Ensemble Transform)



5-8 April, 2011
Norrköping

HIRLAM-ALADIN workshop

The LAM experiments

In a LAM:

- we can take the benefit of global error simulations (in the form of LBCs)
- we would like that the $\boldsymbol{\varepsilon}_b$ sample is suitable to represent background errors on the (smaller) spatial scales of the LAM model
- so we go for global (LBC) + local (initial) perturbations

The LAM experiments

LBC petrurbation (coupling) for all experiments:
- IFS EDA (Experiment by Isaksen et al., 07/2007, 4DVAR T255/L91)

Initial perturbation experiments (period 01-31/07/2007):

- **DSC-EDA**: downscaling of the IFS EDA

$$\boldsymbol{\varepsilon}_{\mathrm{b}} \approx M\,\mathrm{P}\mathbf{x}_{\mathrm{a1}}^{\mathrm{IFS-EDA}} - M\,\mathrm{P}\mathbf{x}_{\mathrm{a2}}^{\mathrm{IFS-EDA}}$$

$\mathrm{P}\mathbf{x}_{\mathrm{a1,2}}^{\mathrm{IFS-EDA}}$: global EDA analyses interpolated to the ALADIN domain

- **LAM-EDA**: local EDA initial perturbations

$$\boldsymbol{\varepsilon}_{\mathrm{b}} \approx M\,\mathbf{x}_{\mathrm{a1}}^{\mathrm{LAM-EDA}} - M\,\mathbf{x}_{\mathrm{a2}}^{\mathrm{LAM-EDA}}$$

$\mathbf{x}_{\mathrm{a1,2}}^{\mathrm{LAM-EDA}}$: local analyses with perturbed observations

- **LAM-ET**: local ET initial perturbations

$$\boldsymbol{\varepsilon}_{\mathrm{b}} \approx M\,\mathbf{x}_{\mathrm{a1}}^{\mathrm{LAM-ET}} - M\,\mathbf{x}_{\mathrm{a2}}^{\mathrm{LAM-ET}}$$

$\mathbf{x}_{\mathrm{a1,2}}^{\mathrm{LAM-EDA}}$: local analyses with ET perturbations

Diagnostics

- **Spectral variance**: variance of the simulated error → diagnoses how the variance is distributed according to spatial scales

- **Spectral spread-skill**: spread and rmse of the ensemble → measures if the error simulation is over or underdispersive (and on which spatial scales)

- **Spectral PECA** (Perturbation vs. Error Correlation Analysis): $\mathrm{corr}\left(|\varepsilon_b|,|\varepsilon_b^{ref}|\right)$

$\varepsilon_b = \overline{x_b} - x_{b,j}$ simulated background error

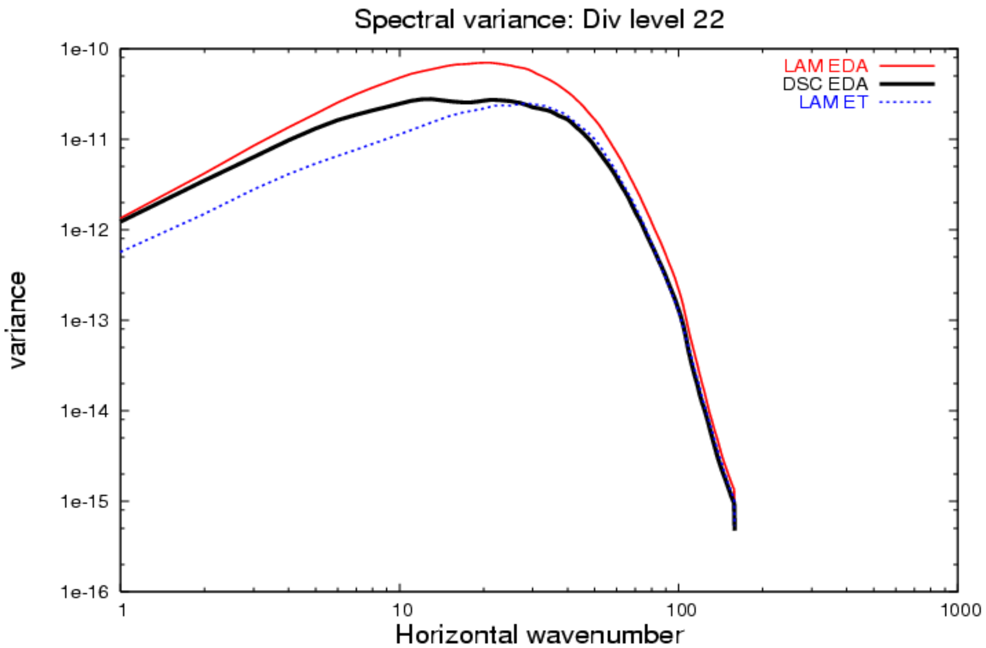
$\varepsilon_b^{ref} = x_a^{verif} - x_{b,j}$ „real” background error ($x_a^{verif} = x_a^{Varpack} \approx x_t$)

→ measures how much the „size” of the simulated error is similar to the size of the „real” (!) background error (and on which spatial scales)

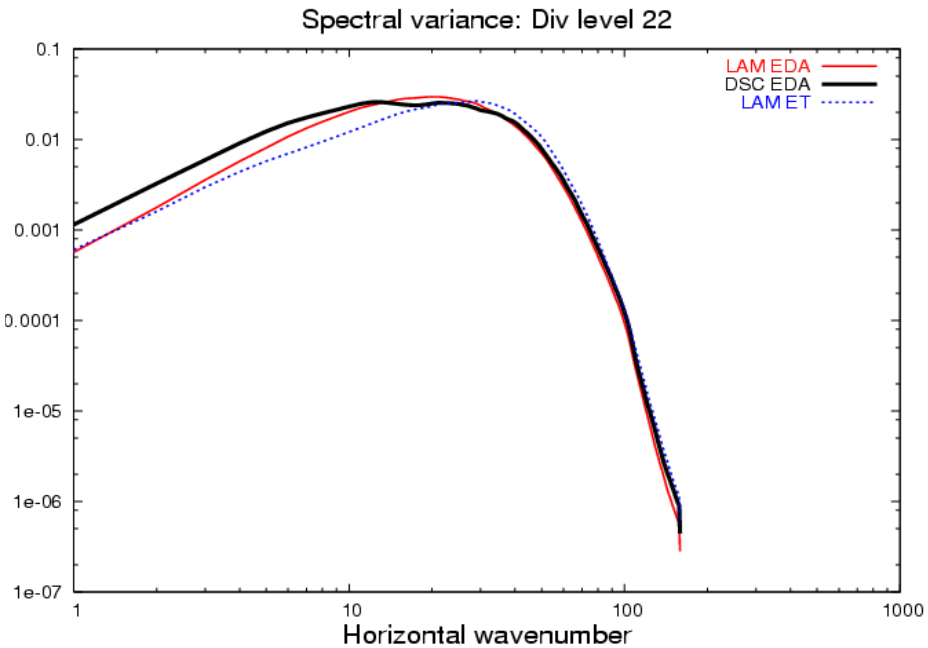
Diagnostic comparisons

Divergence at ~500hPa

Spectral error variance



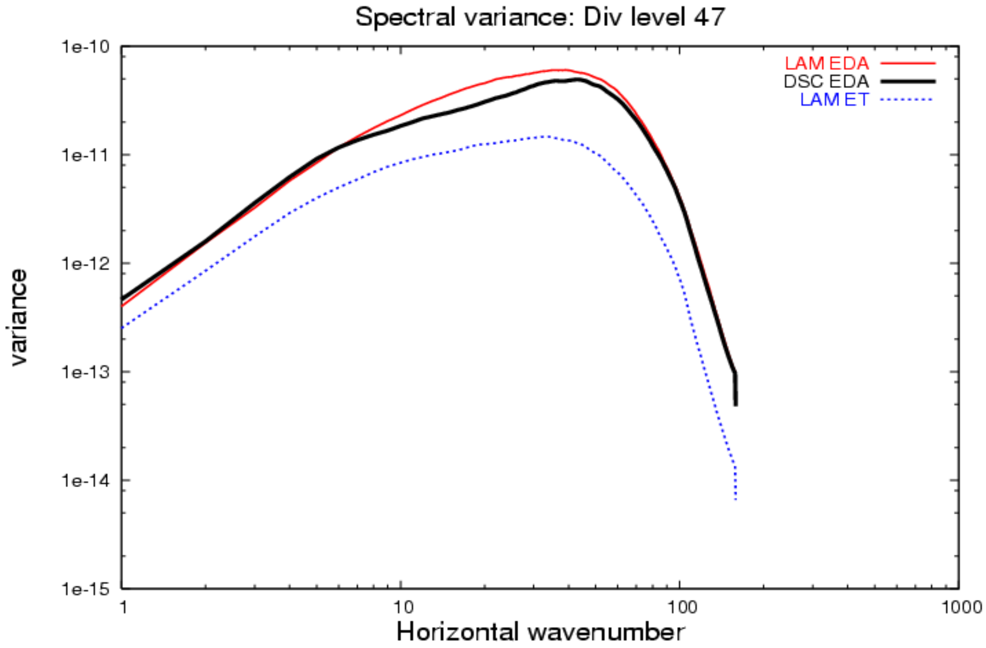
Normalized spectral error variance



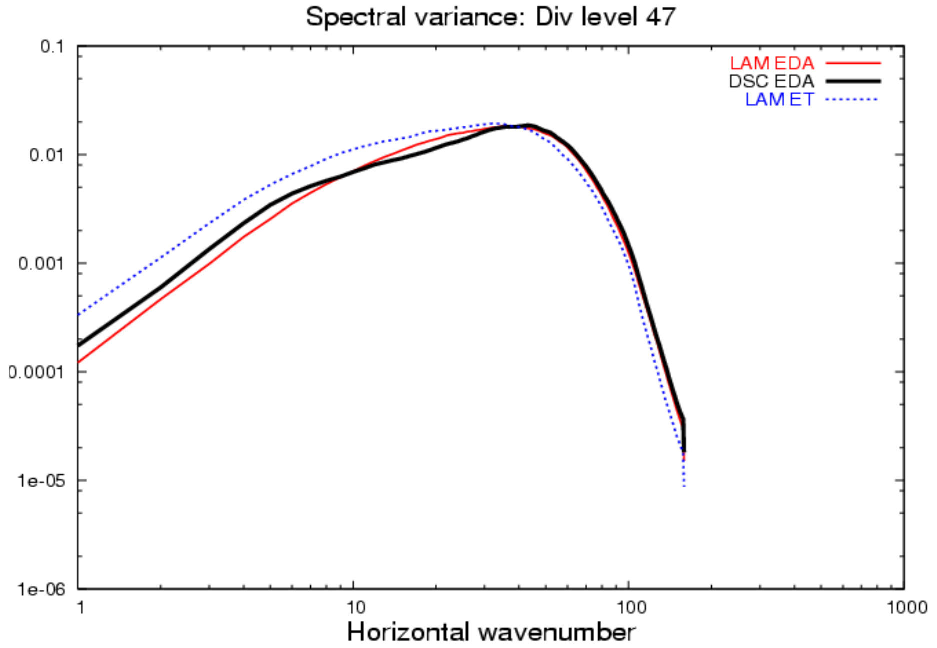
Diagnostic comparisons

Divergence at ~1000hPa

Spectral error variance



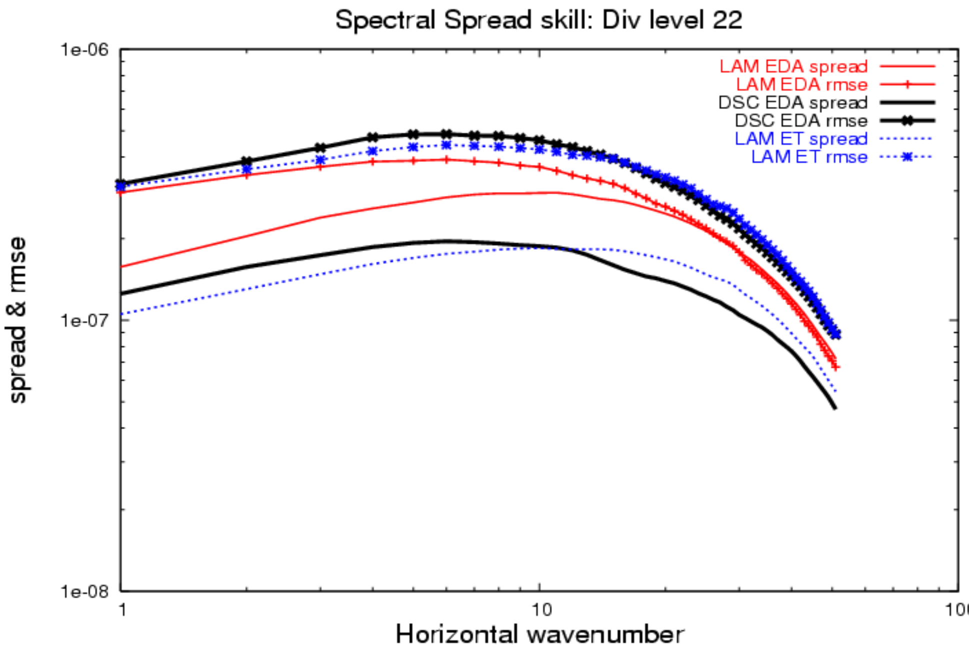
Normalized spectral error variance



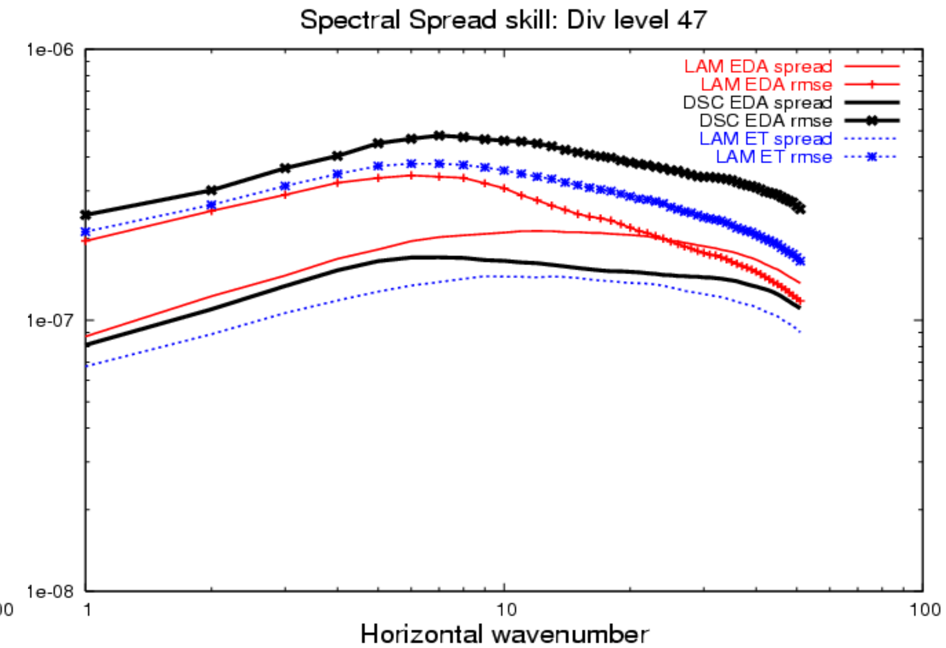
Diagnostic comparisons

Spread-skill (spread-rmse relationship for +6h)

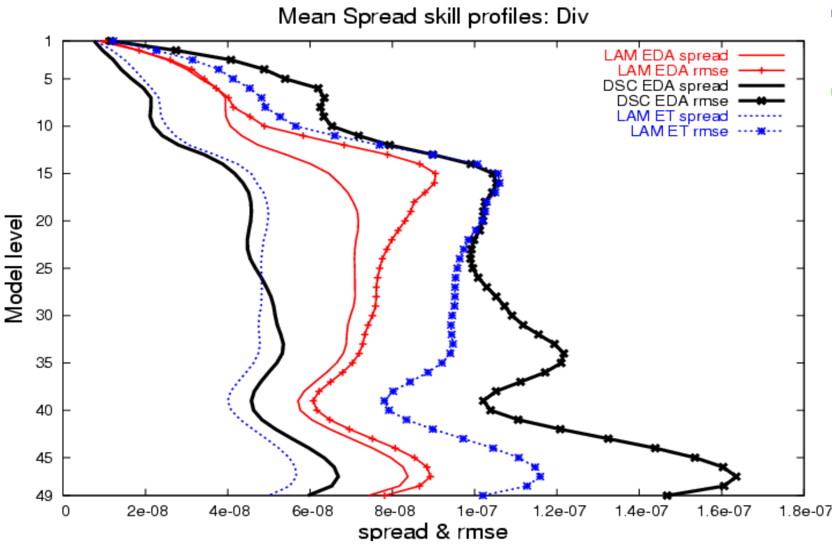
Divergence at ~500hPa



Divergence at ~1000hPa

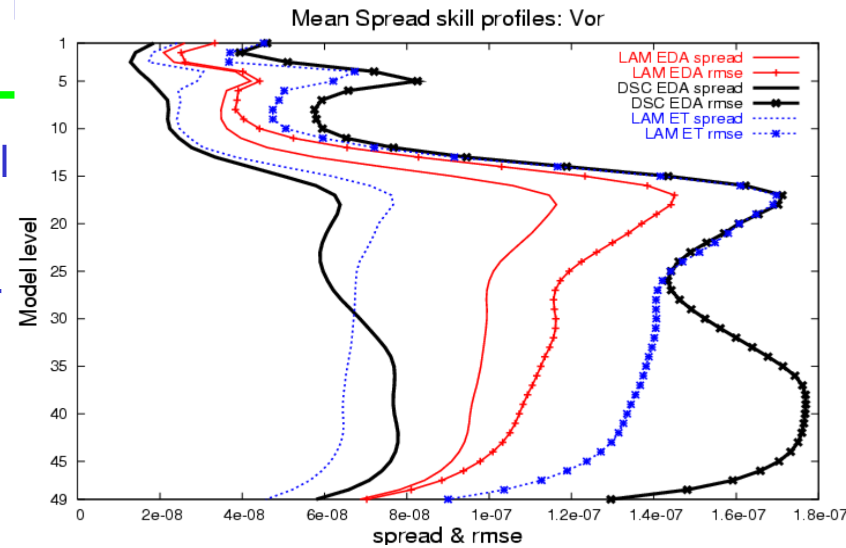


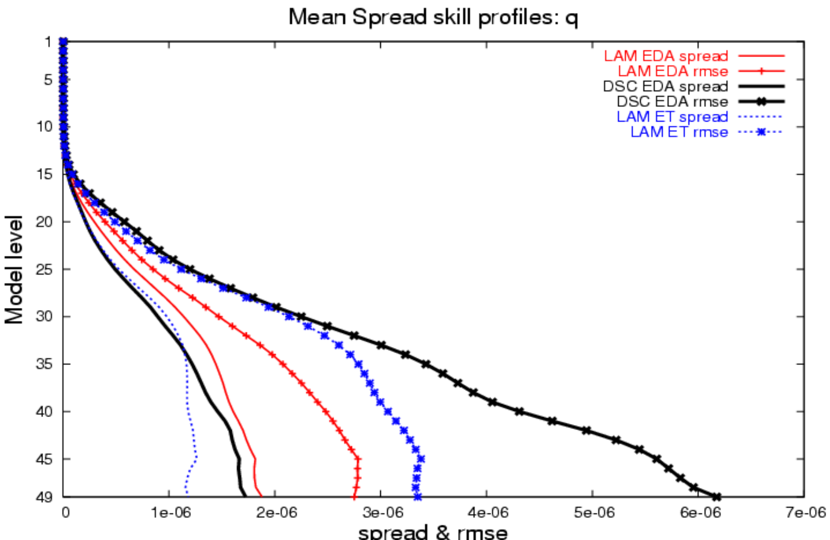
Diagnostic comparisons



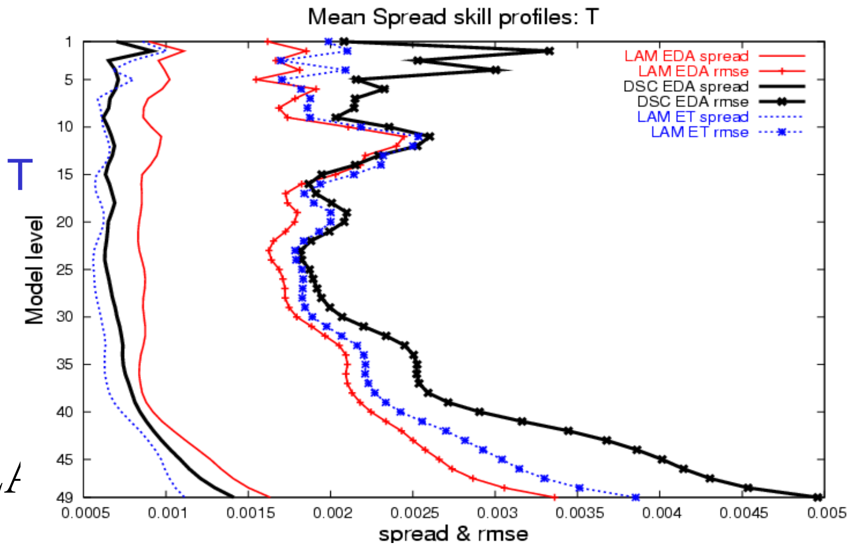
Spread-skill profiles

Div Vor





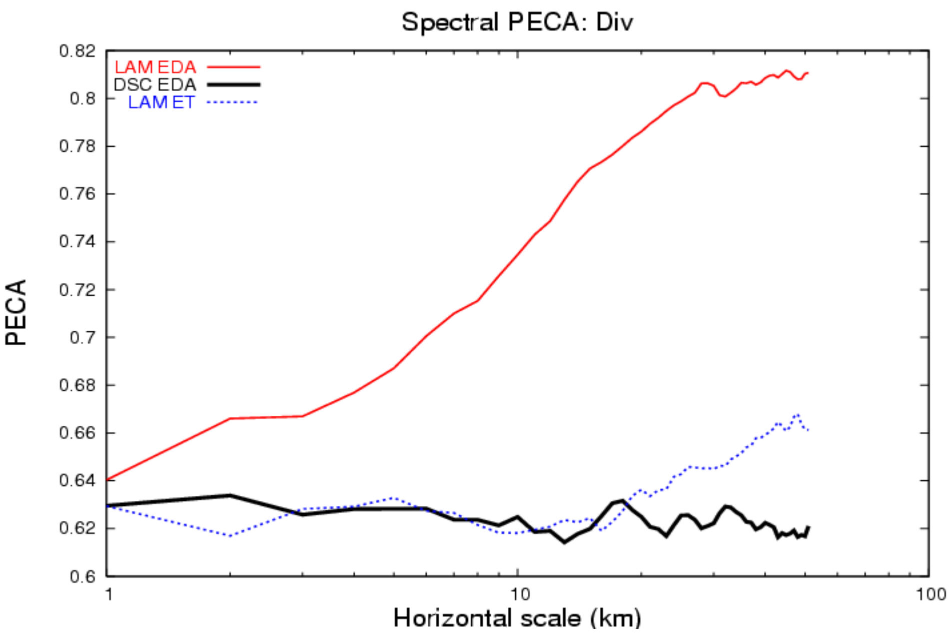
q T



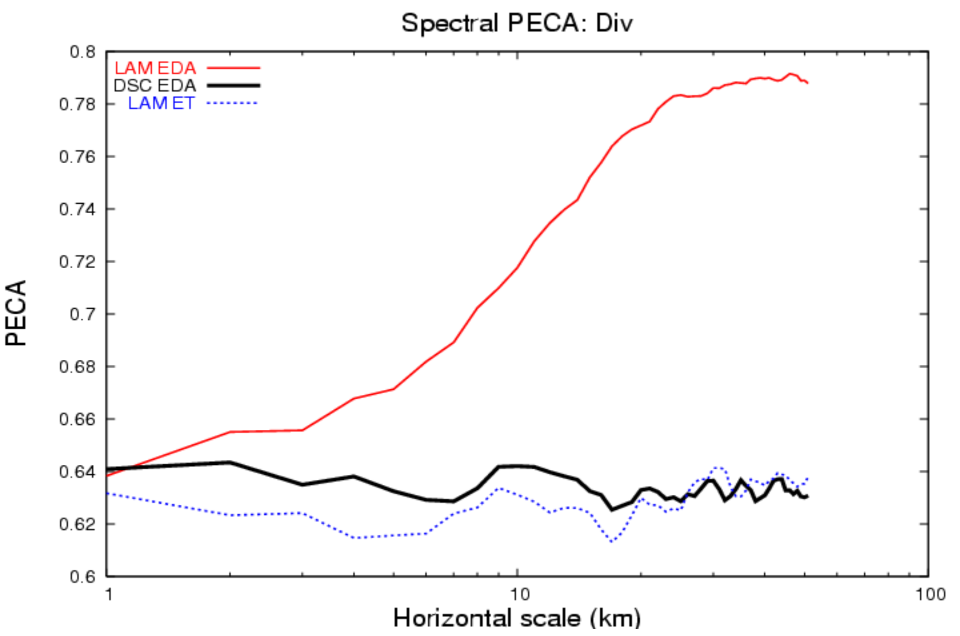
Diagnostic comparisons

PECA: Perturbation versus Error Correlation Analysis

Divergence at ~500hPa



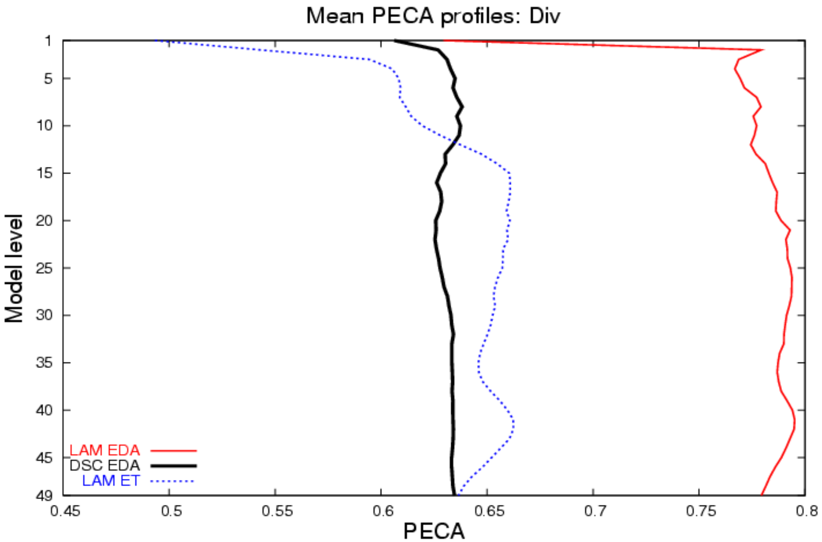
Divergence at ~1000hPa



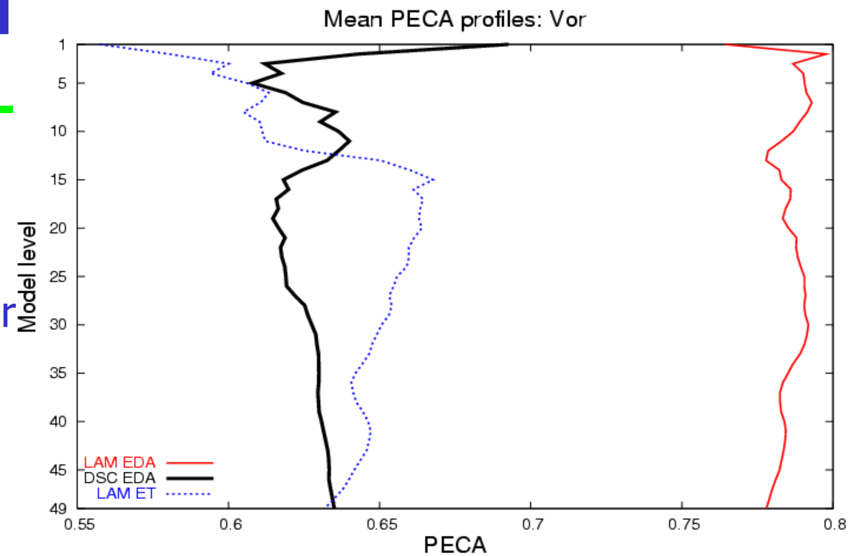
Diagnostic comparisons

PECA profiles

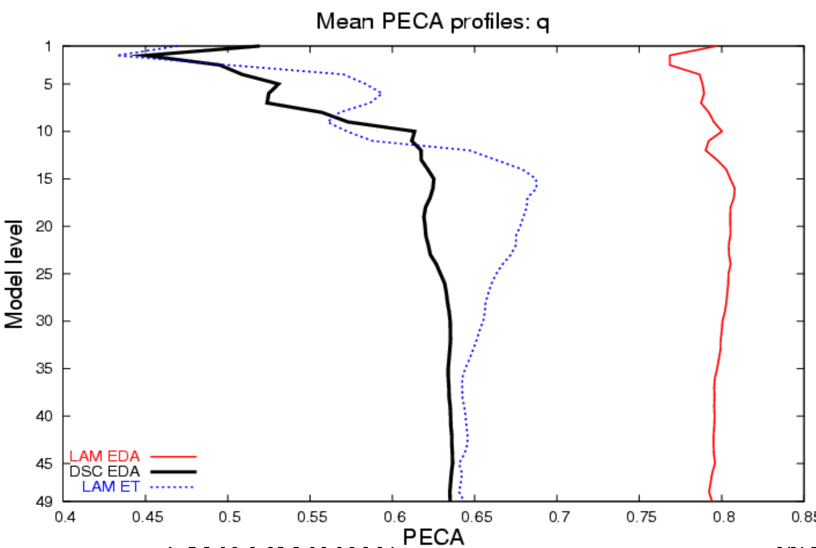
Div



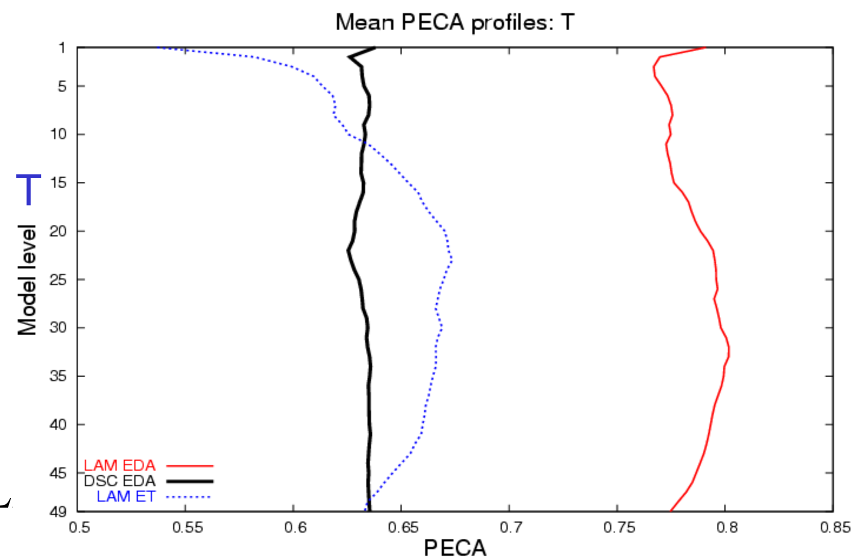
Vor



q



T



Impact studies

Aim: test the impact of the different error simulation techniques on the analysis/forecast ➔ computation of **B** matrices based on the different error simulations ➔ reinject them into real assimilation experiments and verify the analyses and forecasts

Period: 01-31/07/2007 ➔ idealized experiments (the period is the same as used for the error simulation)

2 data assimilation/forecast **experiments**:

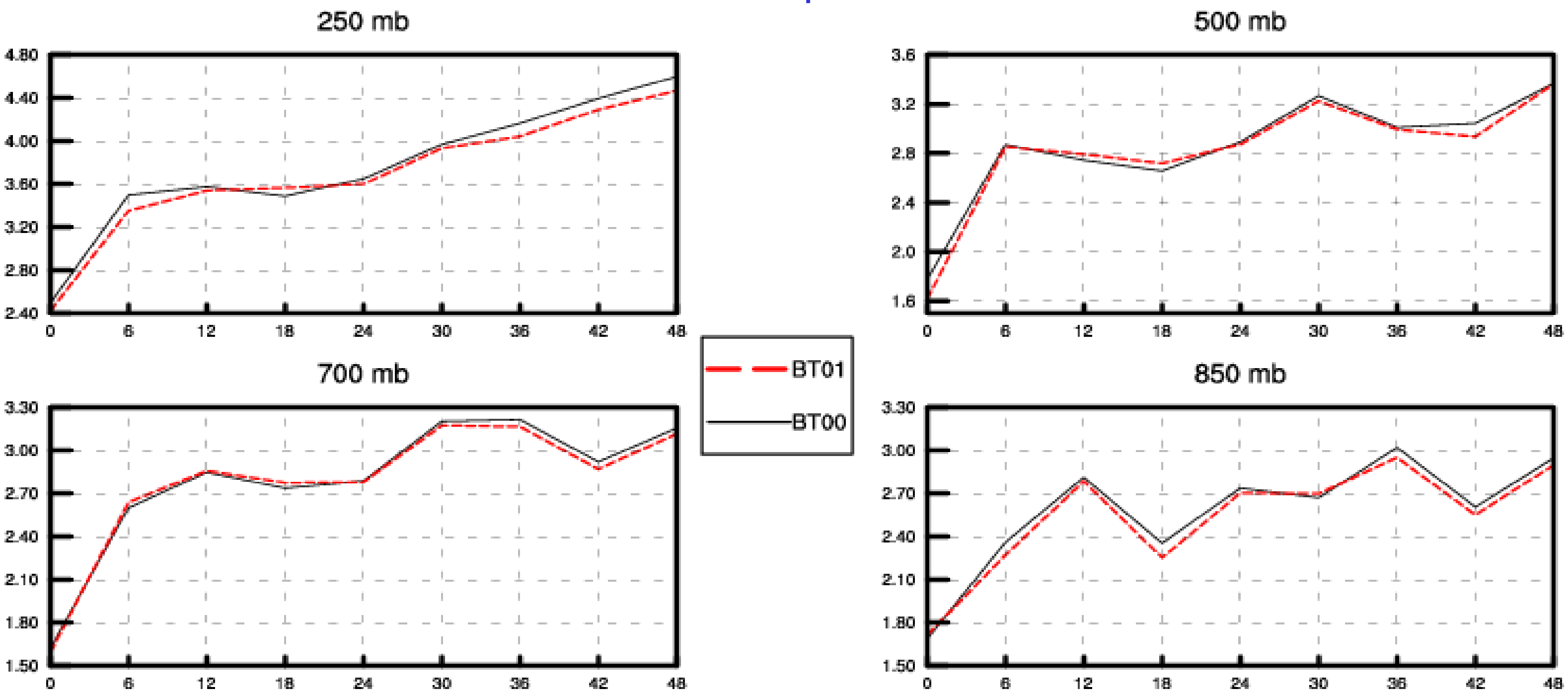
BT00: assimilation cycle using **B** based on the **DSC-EDA** error simulation

BT01: assimilation cycle using **B** based on the **LAM-EDA** error simulation

Impact studies

RMSE against TEMPs and SYNOPs

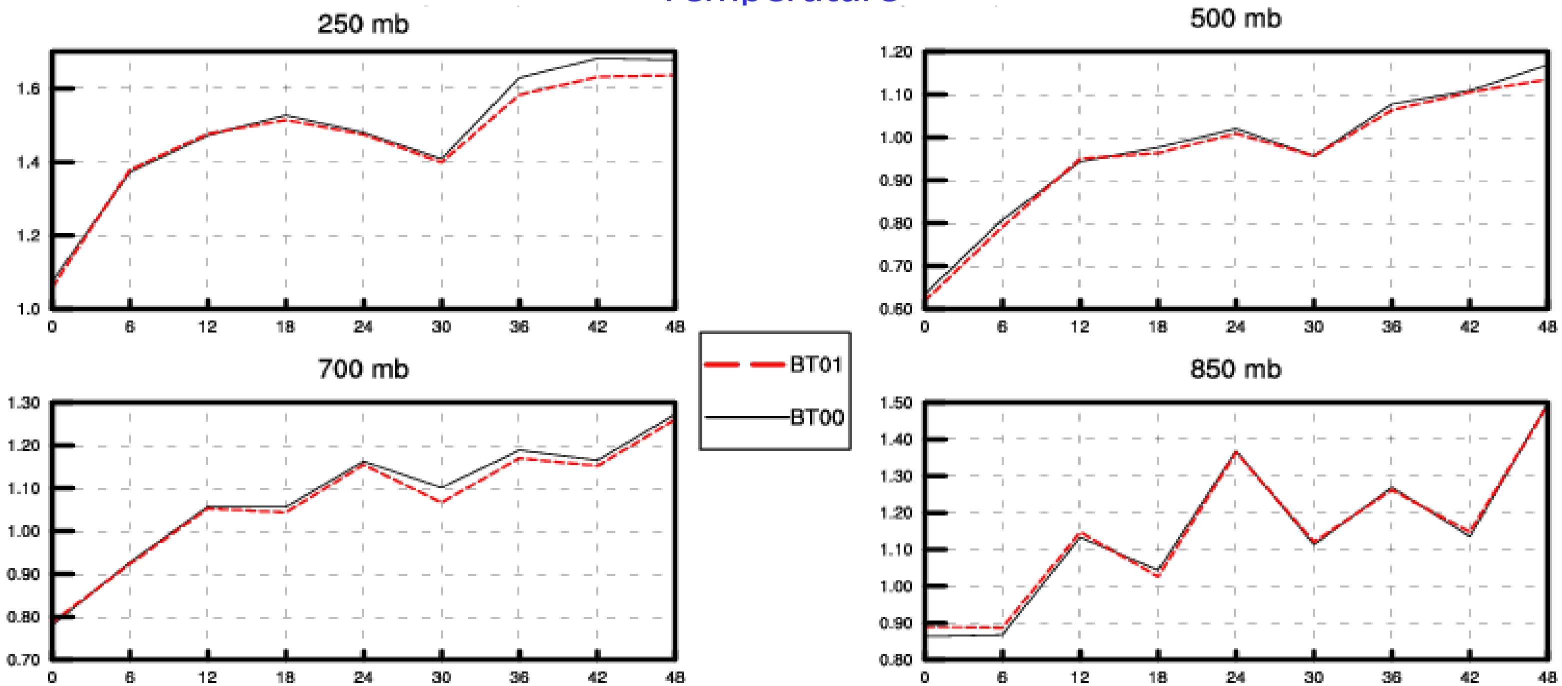
Wind speed



Impact studies

RMSE against TEMPs and SYNOPs

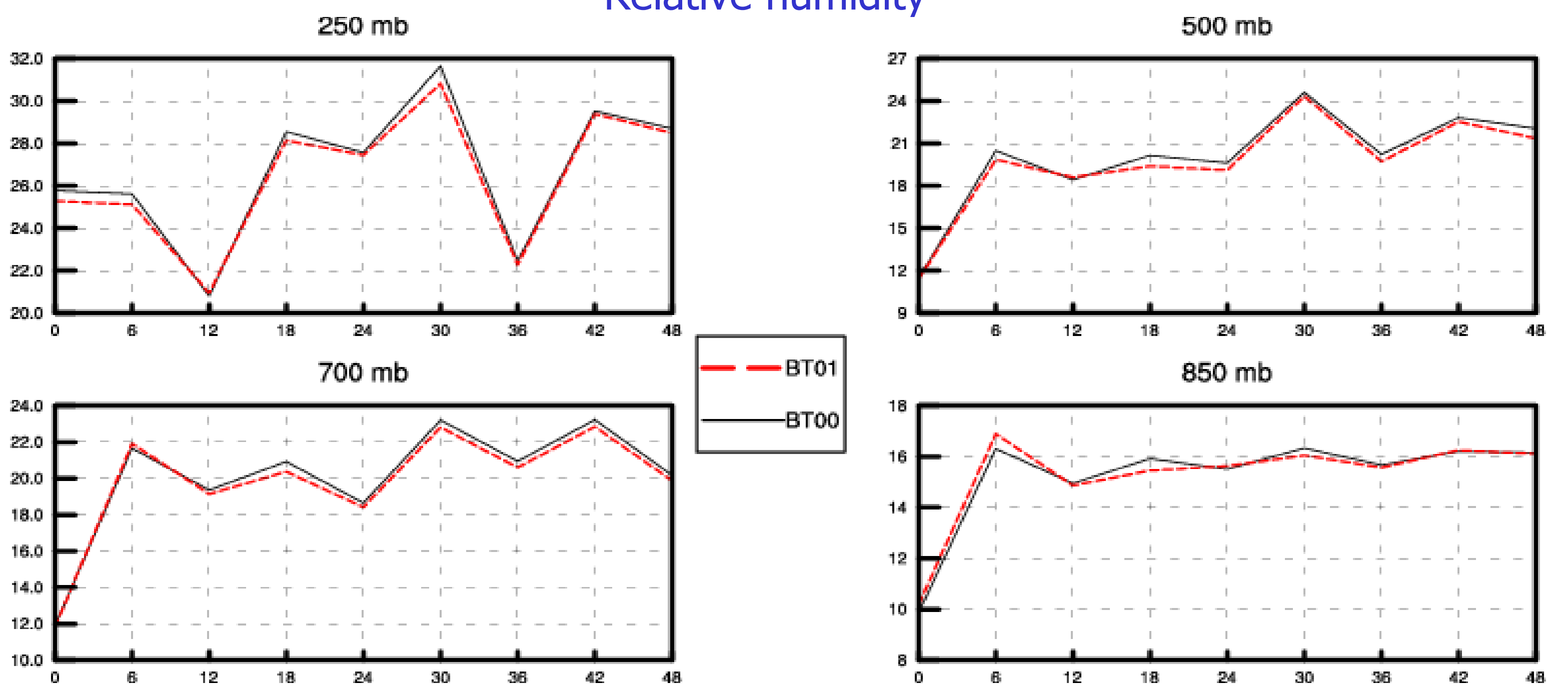
Temperature



Impact studies

RMSE against TEMPs and SYNOPs

Relative humidity

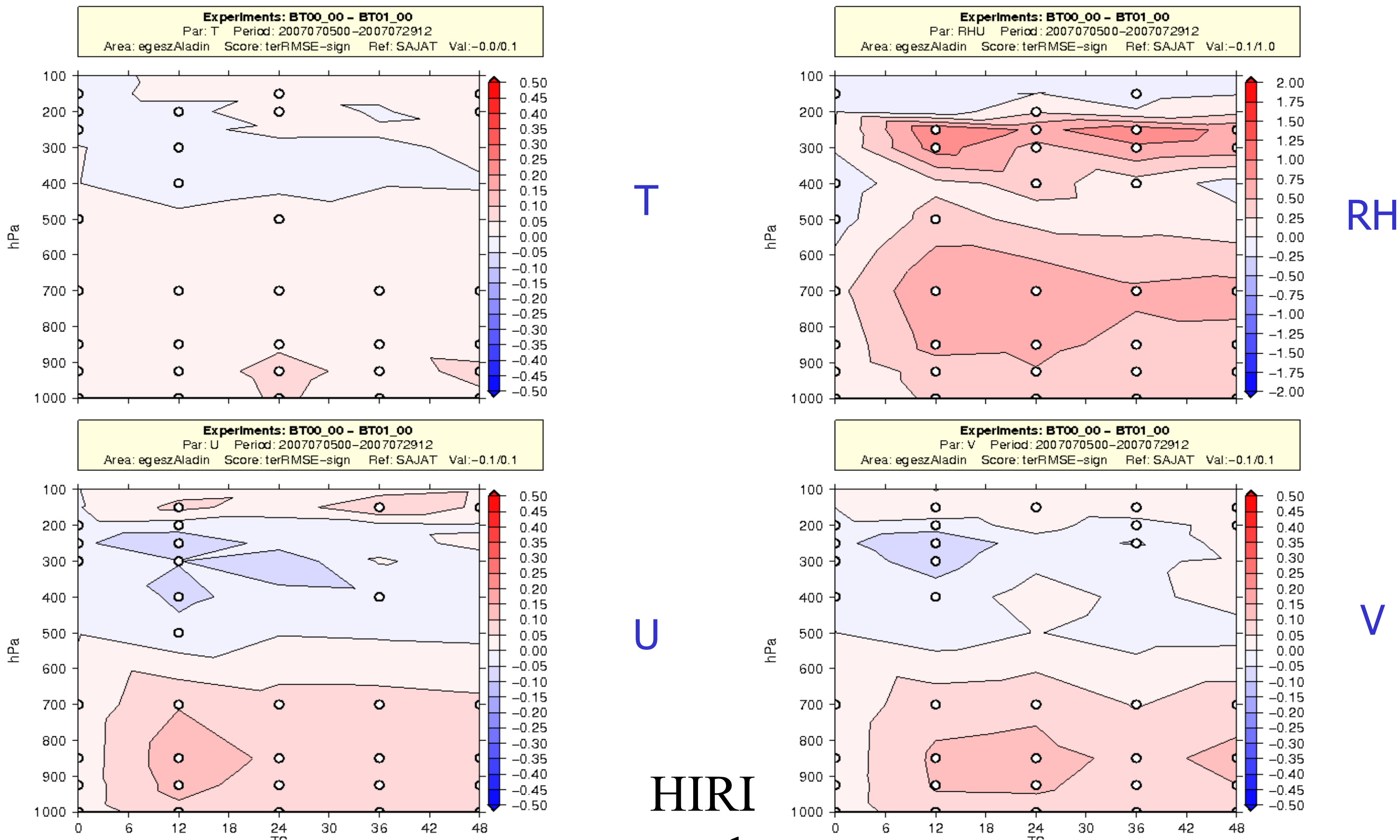


5-8 April, 2011
Norrköping

HIRLAM-ALADIN
workshop

Impact studies

RMSE against analysis (each experiment against its „own” analyses)



Preliminary conclusions

- **LAM-EDA** adds more variance to **DSC-EDA** and often not only in the small scales (scales of observing network?). **LAM-ET** (inspite of the low variance) puts the most variance to the small scales.

- The tried error simulations are mostly underdispersive (except **LAM-EDA** on the small scales). The Spread-skill relationship is the best for **LAM-EDA** (then **LAM-ET** then **DSC-EDA**). The spread of **LAM-ET** is too low. The rmse is decreased by the LAM experiments compared to **DSC-EDA**.

- PECA correlations are the best for **LAM-EDA** (then for **LAM-ET** then for **DSC-EDA**)

- The 3 diagnostics (Spectral variance, Spread-skill, PECA) are in good correspondance with each-other

- Overall **LAM-EDA** seems to be the best of the 3 simulation techniques, however there is a potential in the **LAM-ET** technique in case of proper inflation (increased spread).

- Assimilation/forecast experiments show an improvement using a **B** matrix based on the **LAM-EDA** simulation compared to the use of **DSC-EDA**

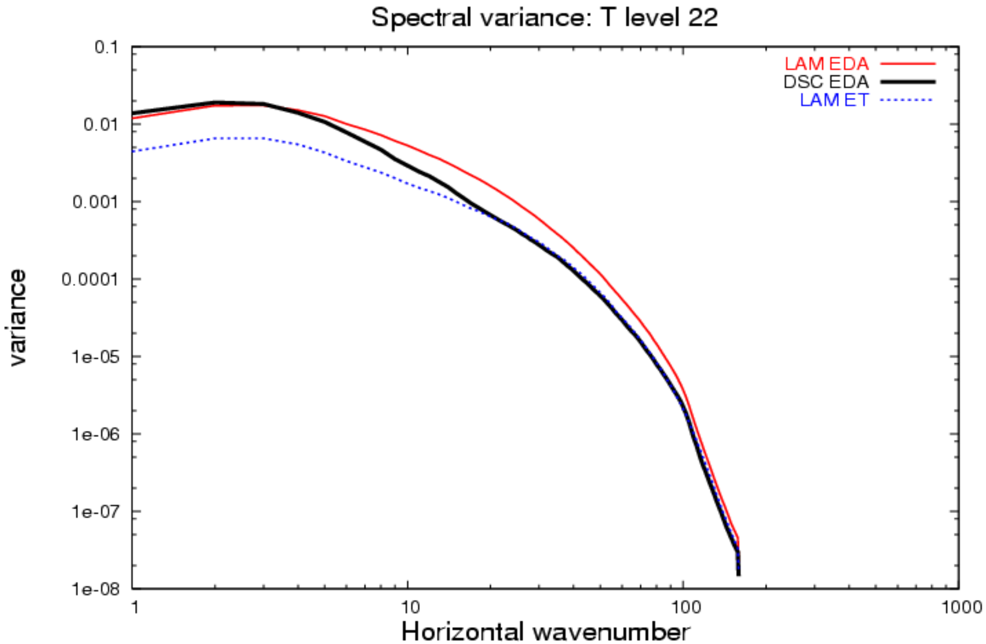
Thank you for your attention!

?

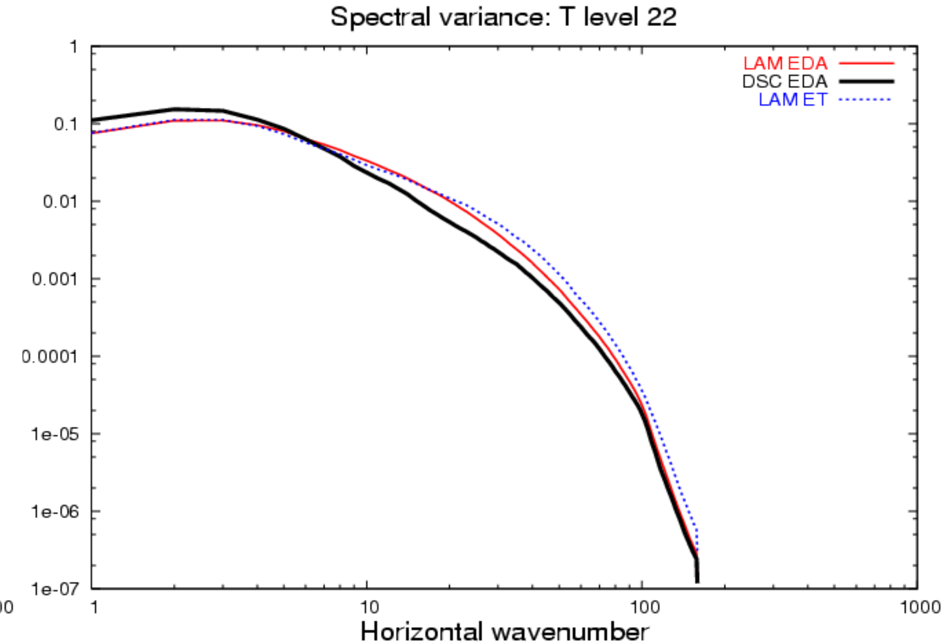
Diagnostic compatrisons

Temperature at ~500hPa

Spectral error variance



Normalized spectral error variance

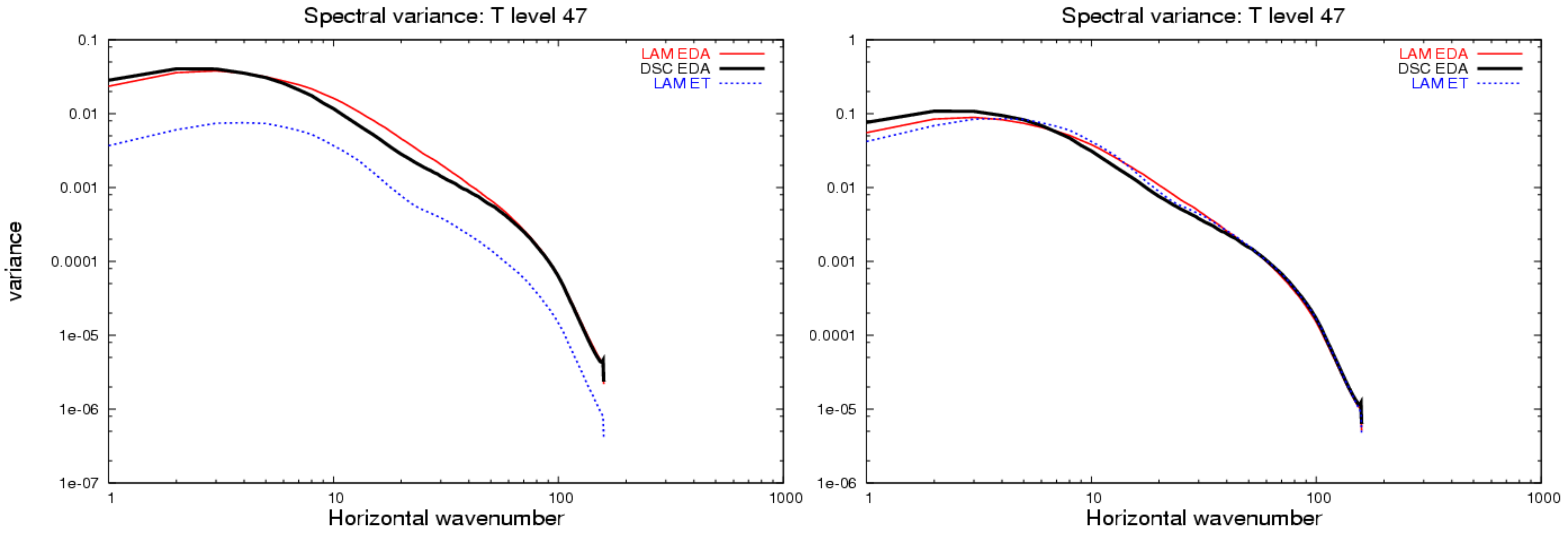


Diagnostic compatrisons

Temperature at ~1000hPa

Spectral error variance

Normalized spectral error variance



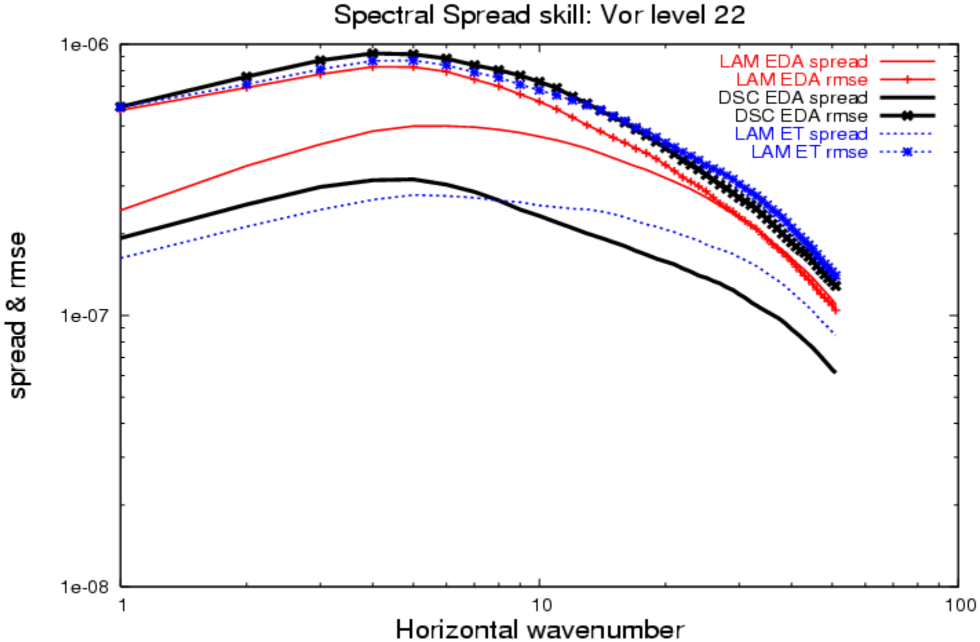
5-8 April, 2011
Norrköping

HIRLAM-ALADIN
workshop

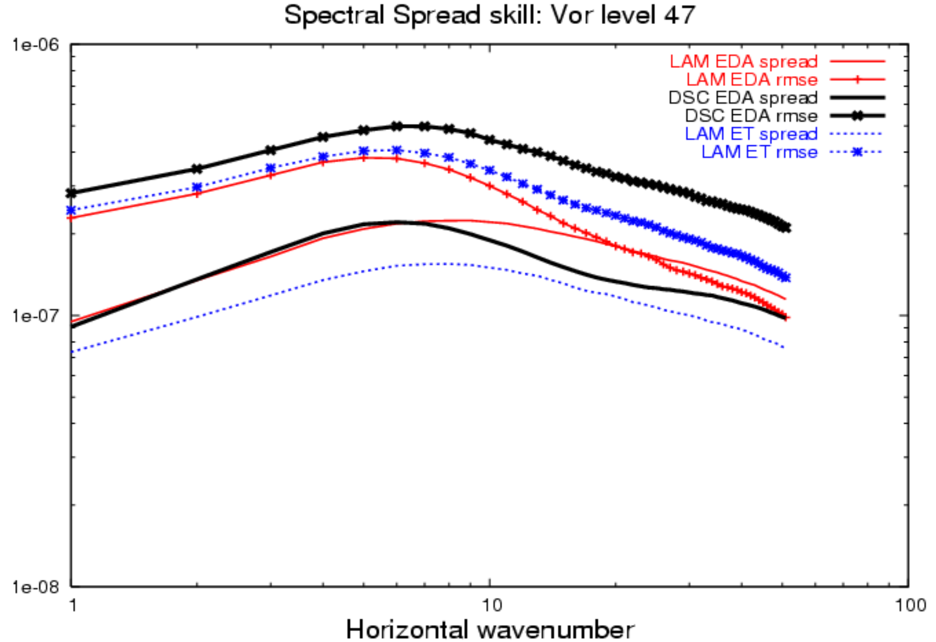
Diagnostic compatrisons

Spread-skill (spread-rmse relationship)

Vorticity at ~500hPa



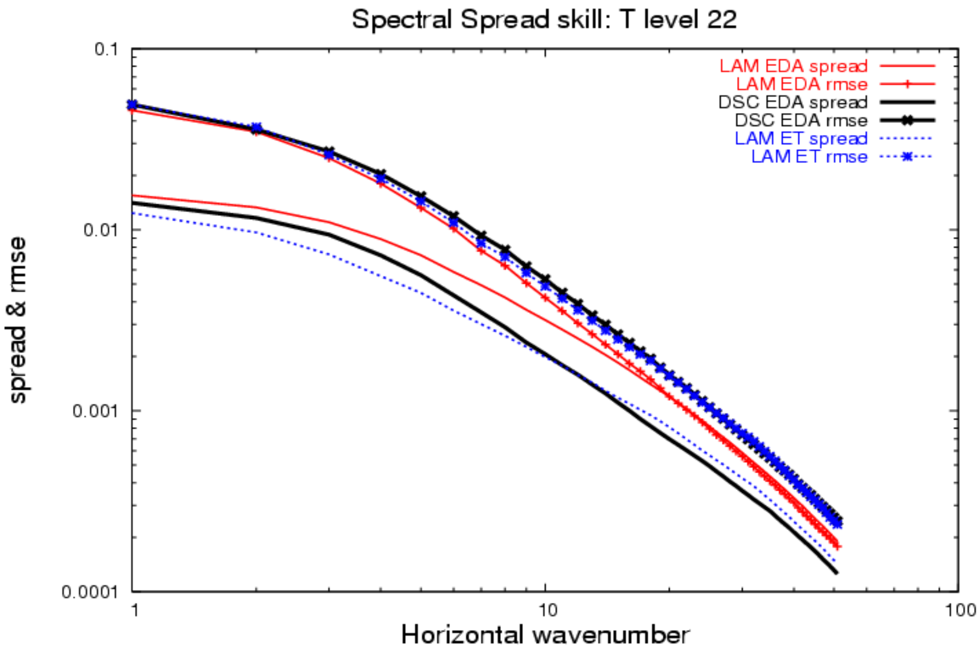
Vorticity at ~1000hPa



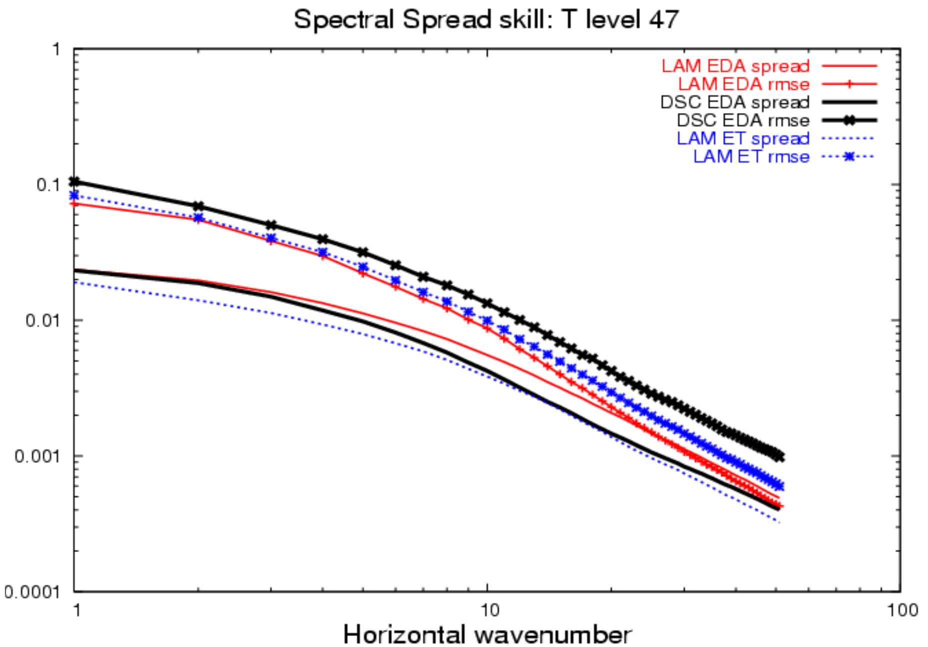
Diagnostic compatrisons

Spread-skill

Temperature at ~500hPa



Temperature at ~1000hPa

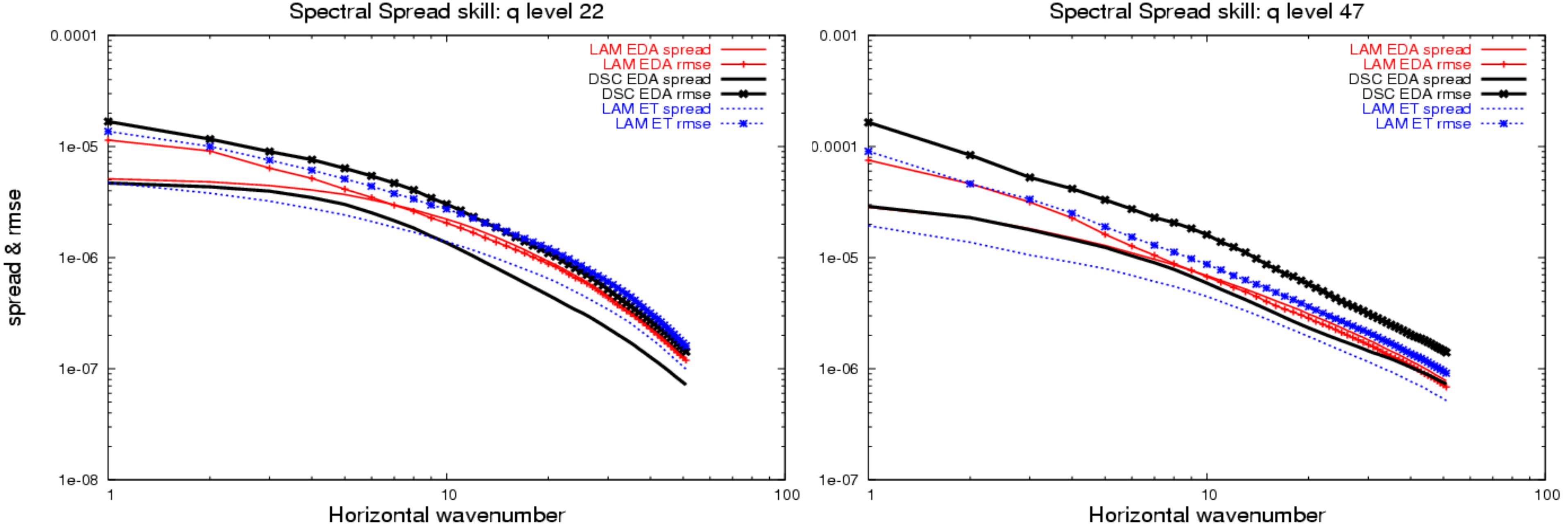


Diagnostic compatrisons

Spread-skill

Specific humidity at ~500hPa

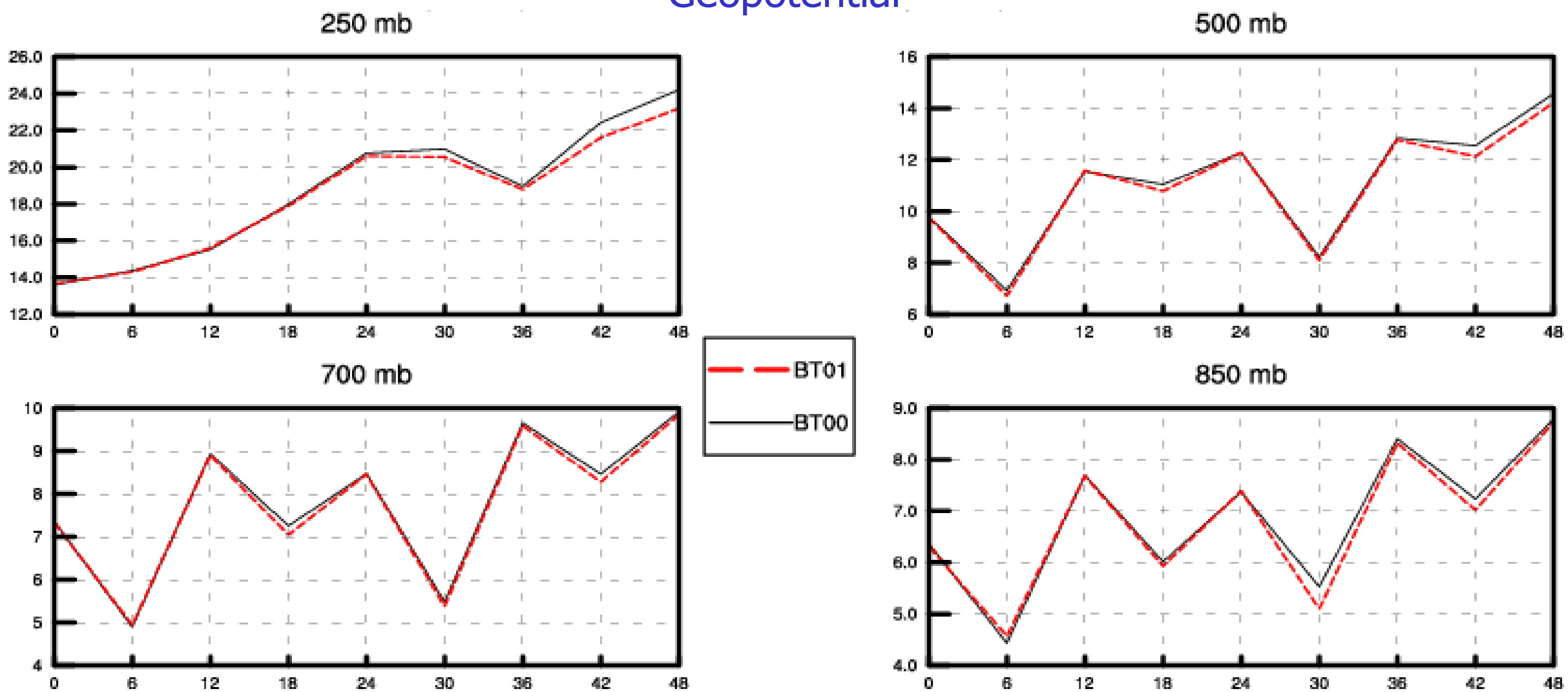
Specific humidity at ~1000hPa



Impact studies

RMSE against TEMPs and SYNOPs

Geopotential



Diagnostics

Error variance:

$$Var(\varepsilon_b) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\varepsilon_b^i - \overline{\varepsilon_b}\right)^2}$$

N= member size + time realizations

Spread:

$$Sp = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(x_b^{t,j} - \sum_{j=1}^{M} x_b^{t,j}\right)^2}$$

T= time realizations M= member size

RMSE:

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(x_a^{verif,t} - \sum_{j=1}^{M} x_b^{t,j}\right)^2}$$

T= time realizations M= member size

Diagnostics

PECA (Perturbation vs. Error Correlation Analysis):

$$\mathrm{Corr}\left(|\varepsilon_b|,|\varepsilon_b^{ref}|\right)=\frac{\mathrm{Cov}\left(|\varepsilon_b|,|\varepsilon_b^{ref}|\right)}{\sigma(\varepsilon_b)\sigma(\varepsilon_b^{ref})}=\frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(|\varepsilon_b^{i}-\overline{\varepsilon_b}|\right)\left(|\varepsilon_b^{ref,i}-\overline{\varepsilon_b^{ref}}|\right)}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\varepsilon_b^{i}-\overline{\varepsilon_b}\right)^2}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\varepsilon_b^{ref,i}-\overline{\varepsilon_b^{ref}}\right)^2}}$$

$\varepsilon_b=\overline{x_b}-x_{b,j}$ simulated background error

$\varepsilon_b^{ref}=x_a^{verif}-x_{b,j}$ „real" background error ($x_a^{verif}\approx x_t$)

N= member size + time realizations