The correction of forecast errors based on MOS: The impact of initial condition and model errors

S. Vannitsem

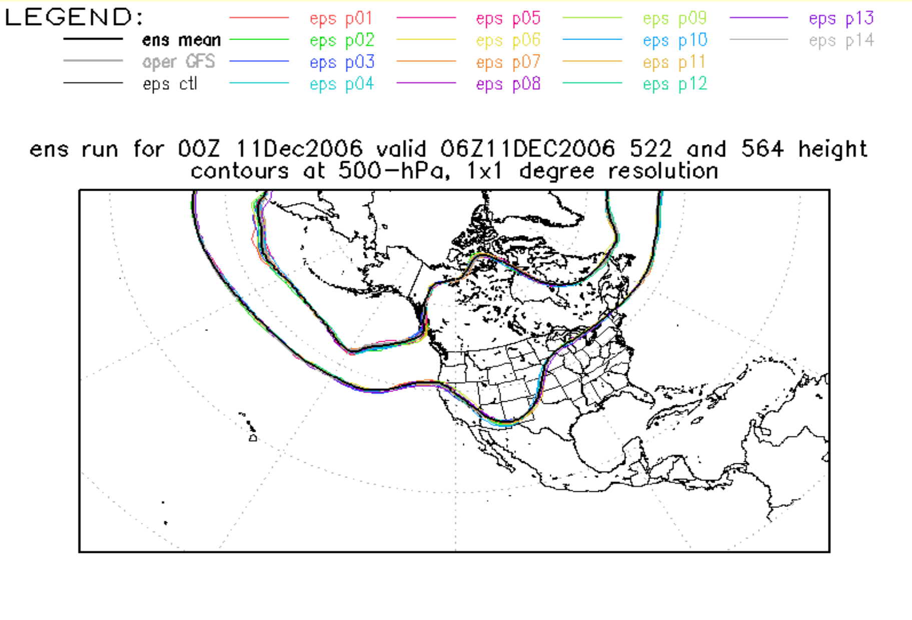
Institut Royal Météorologique de Belgique

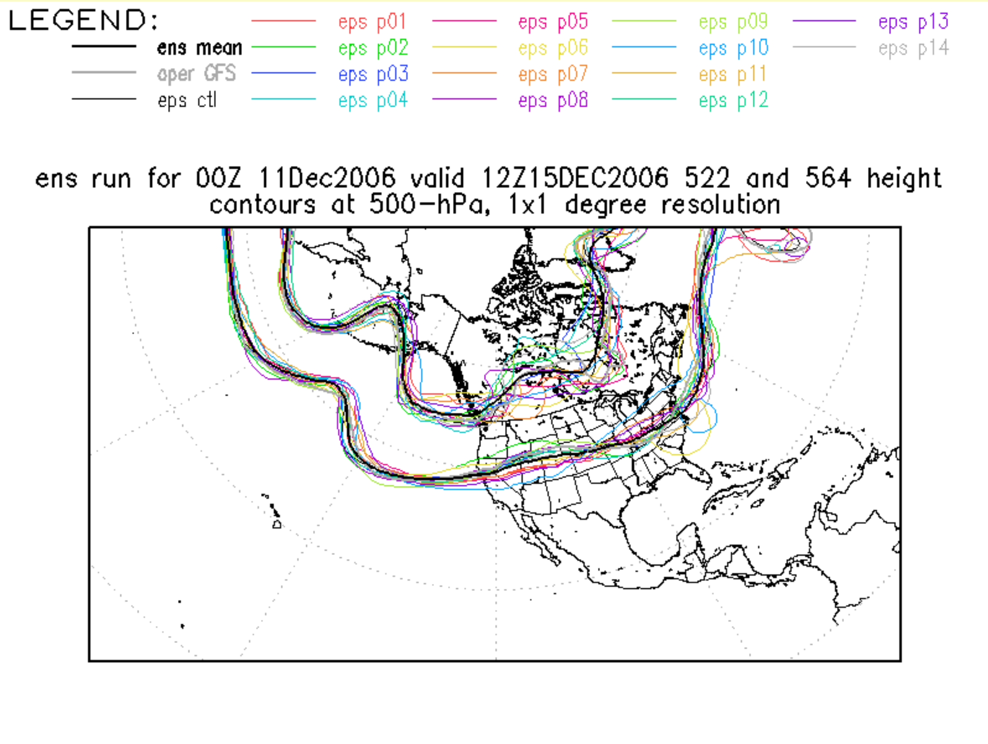
Dresden, July 30, 2009

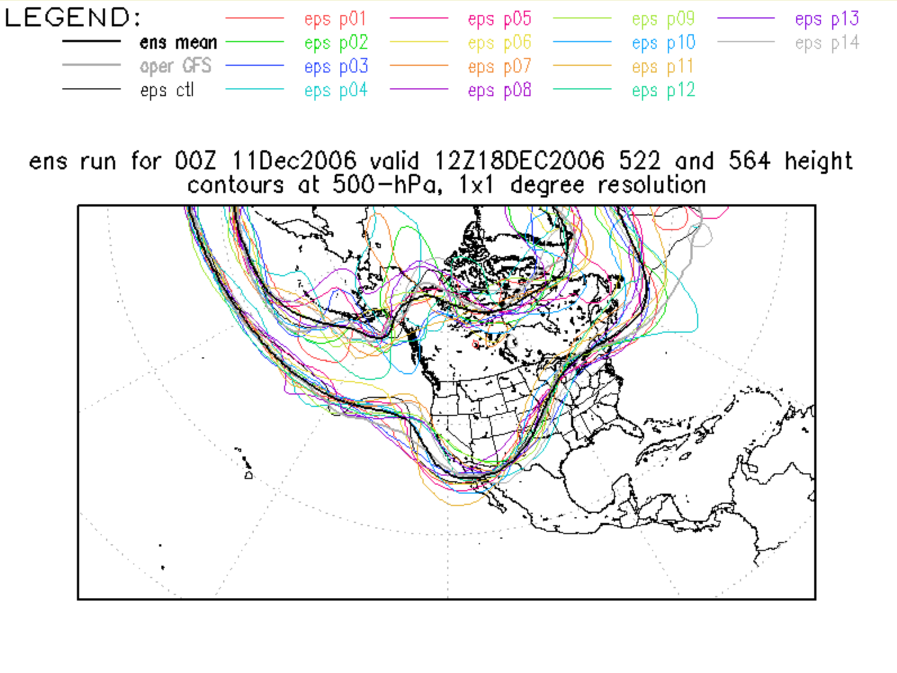
Outline

1. Introduction
2. The MOS technique
3. The role of initial and model errors: a dynamical view
4. Analysis of operational forecasts
5. A unified MOS scheme for deterministic AND ensemble forecasts: EVMOS
6. Conclusions and perspectives

The quality of atmospheric forecasts decreases as a function of time



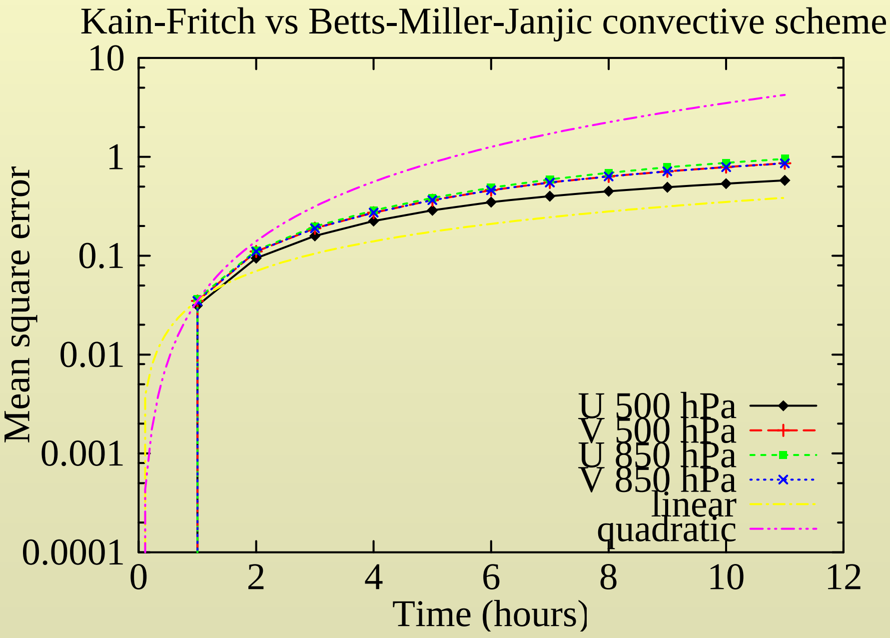


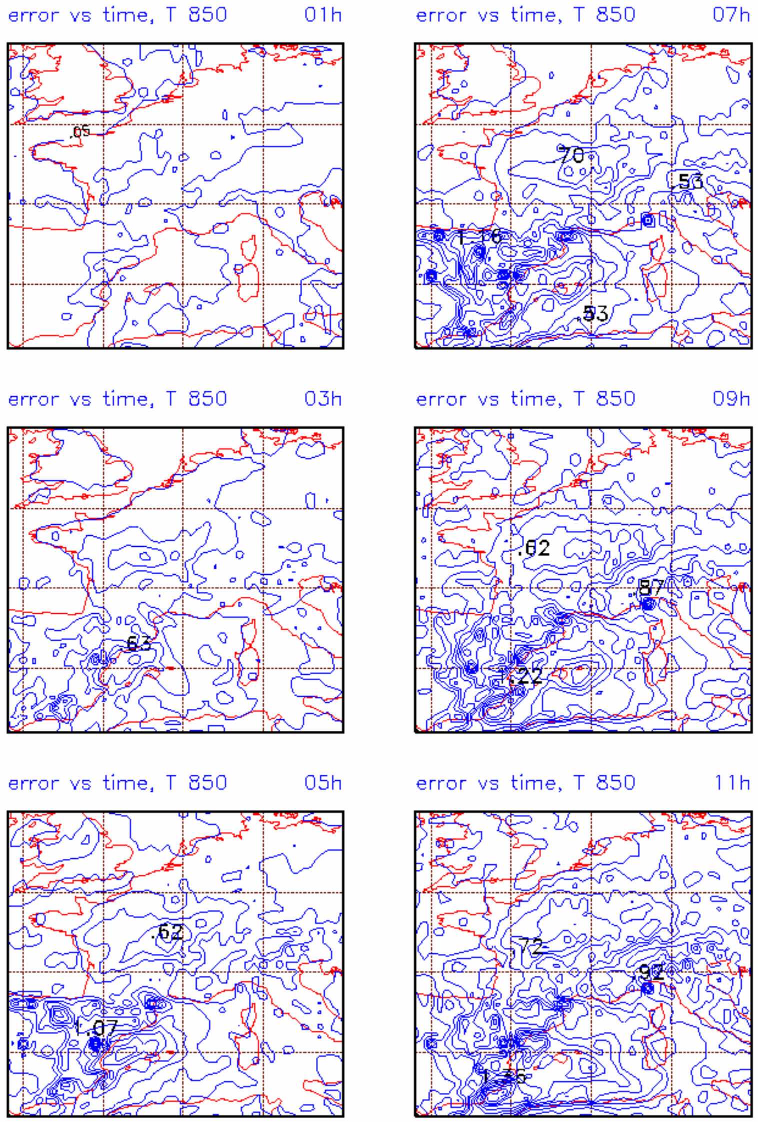


NCEP ensemble forecasts

Model error dynamics

Eta model runs with
Two different convective
schemes





Two sources of errors affecting the forecasts and the data assimilation process:

-initial condition errors
-model errors

See *C. Nicolis, Rui A. P. Perdigao, and S. Vannitsem* Dynamics of Prediction Errors under the Combined Effect of Initial Condition and Model Errors, *Journal of the Atmospheric Sciences* , 66, 766–778, 2009.

How can we improve the forecast quality (Besides the improvement of the model and the observational system)?

Use of techniques to post-process the forecasts in order to improve their quality.

What is corrected by the post-processing method?

The Model Output Statistics technique

Post-processing of forecasts in order to improve their quality based on information gathered from past forecasts

- 'Deterministic' approach: Linear techniques (Classical MOS, perfect prog), adaptative Kalman filtering. Ref: Glahn and Lowry, 1972, JAM, 11, 1203; Wilks, 2006, Academic Press;….

- 'Deterministic' approach: Nonlinear techniques (Neural networks, nonlinear fits). Ref: Marzban, 2003, MWR, 131, 1103; Casaioli et al, 2003, NPG, 10, 373;….

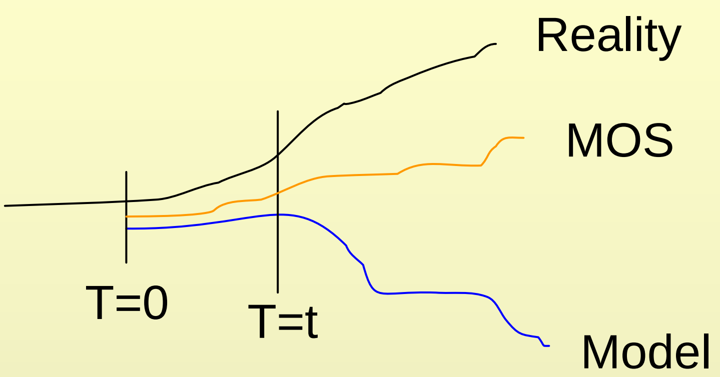
- Probabilistic approaches: correction of the (deterministic) precipitation forecasts. Ref: Lemcke and Kruizinga, 1988, MWR, 116, 1077;….

- (True) Probabilistic approaches: Calibration of the ensemble forecasts. Ref: Roulston and Smith, 2003, Tellus, 55A, 16; Gneiting et al, 2005, MWR, 133, 1098; Wilks, 2006, MA, 13, 243; Hamill and Wilks, 2007, MWR, 135, 2379;…….

The Linear MOS technique

Linear regression between a set of observables of forecasts and observations

$$X_c(t) = \alpha(t) + \sum_{i=1}^{n} \beta_i(t) V_i(t)$$



$$J(t) = \sum_{k=1}^{M} (X_{c,k}(t) - X_k(t))^2$$

M past forecasts

Minimization of J(t)

n=1

$$\alpha(t) = \langle X(t) \rangle - \beta(t) \langle V(t) \rangle$$

$$\beta(t) = \frac{\langle X(t)V(t) \rangle - \langle X(t) \rangle \langle V(t) \rangle}{\langle V(t)^2 \rangle - \langle V(t) \rangle^2}$$

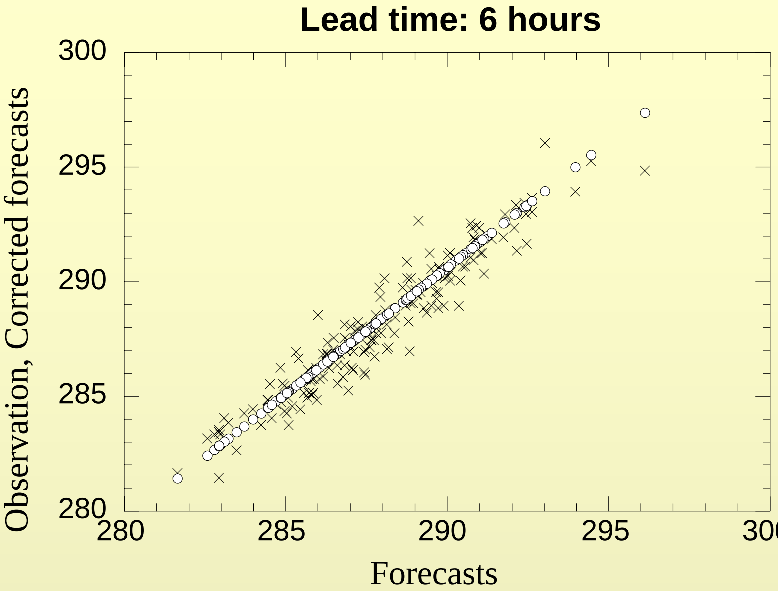
Example of fit: 2-m Temperature, Uccle

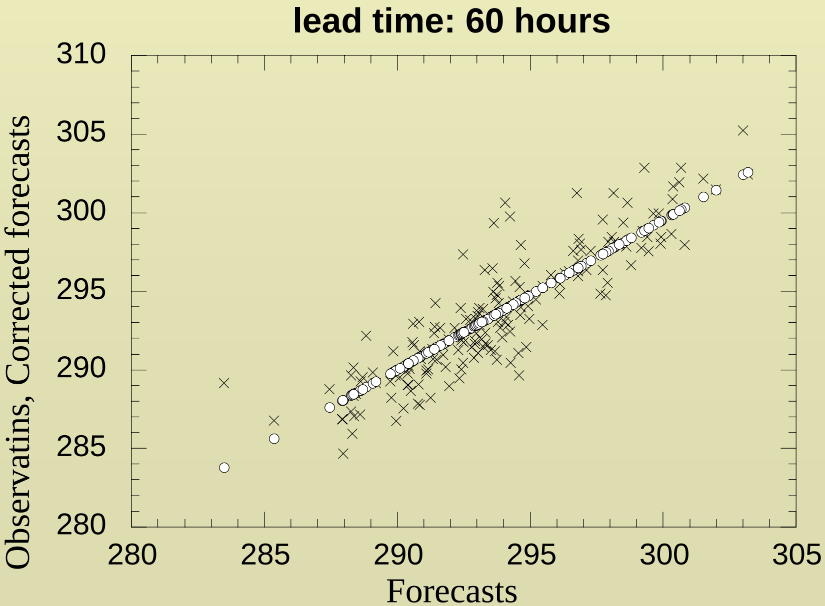
Training:

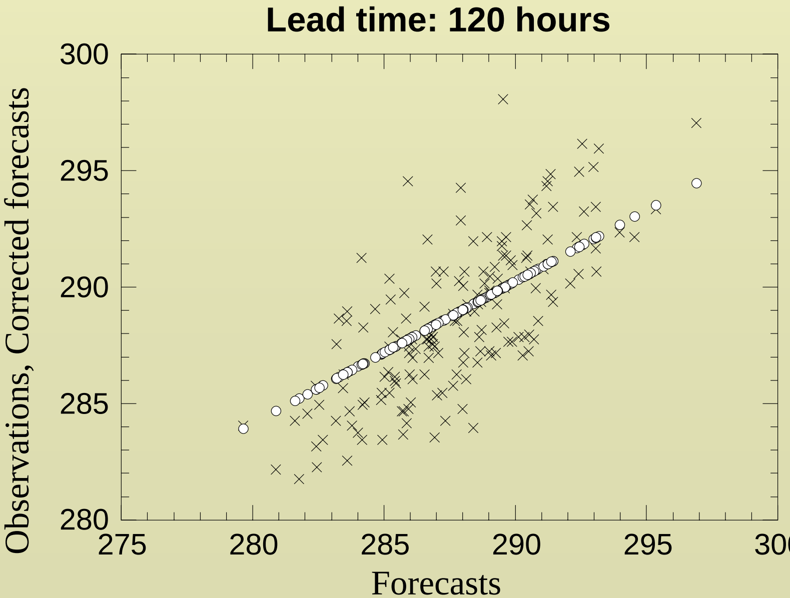
2 Summer seasons
2002-2003

Verification:

2004-2005







Properties of the MOS forecast

$$< X_c(t) >=< X(t) >$$

$$\sigma_c^2(t) =< (X_c(t) - < X(t) >)^2 >= \frac{C(X(t),V(t))^2}{\sigma_{V(t)}^2 \sigma_X^2} \sigma_X^2 = \beta(t)^2 \sigma_{V(t)}^2$$

Progressive decrease

The mean square error

$$< (X_c(t) - X(t))^2 >=< (V(t) - X(t))^2 > -(\sigma_C(t) - \sigma_V(t))^2 - (< X(t) > - < V(t) >)^2$$

Drift correction

Variability correction

Short term dynamics of MOS forecasts (Formalism based on Nicolis, 2004, JAS)

The real system

$$\frac{d\vec{X}}{dt} = \vec{F}(\vec{X}, \vec{Y}, \{\mu\})$$

$$\frac{d\vec{Y}}{dt} = \vec{S}(\vec{X}, \vec{Y}, \{\lambda\})$$

→ Variables not described by the model

$\vec{X}$ and $\vec{\mathrm{V}}$ span the same phase space

The model

$$\frac{\mathrm{d}\vec{\mathrm{V}}}{\mathrm{dt}} = \vec{\mathrm{G}}(\vec{\mathrm{V}}, \{\mu'\})$$

Formal sols

$$\vec{V}(t) = \vec{V}(0) + \int_0^t d\tau \vec{G}(\vec{V}(\tau), \{\mu'\})$$

$$\vec{X}(t) = \vec{X}(0) + \int_0^t d\tau \vec{F}(\vec{X}(\tau), \vec{Y}(\tau), \{\mu\})$$

Typical initial errors

$$\vec{V}(0) = \vec{X}(0) + \vec{\varepsilon}(0) \longrightarrow$$

Gaus random noise

+

Systematic error

$$\vec{\varepsilon} = <\vec{\varepsilon}> + \sigma_{\varepsilon} N(0,1)$$

$$(\sigma_C(t) - \sigma_V(t))^2 = S_0 + S_1 t + \frac{1}{2} S_2 t^2 + O(t^3)$$

$$S_0 = \frac{\sigma^4_{\varepsilon(0)}}{\sigma^2_{\varepsilon(0)} + \sigma^2_{X(0)}}$$

$$S_1 \propto 2\sigma^2_{\varepsilon(0)} C(V(0), G(\vec{X}(0)) - F(\vec{X}(0)))$$

$$S_2 \propto 2C(V(0), G(\vec{X}(0)) - F(\vec{X}(0))^2$$

$$C(V(0), G(\vec{X}(0)) - F(\vec{X}(0))) = <V(0)(G-F)> - <V(0)><G-F>$$

FOR 2 PREDICTORS: C(V1,G-F), C(V2,G-F)

$$(<X(t)> - <V(t)>)^2 = S'_0 + S'_1 t + \frac{1}{2} S'_2 t^2 + O(t^3)$$

Systematic initial error

$$S'_0 = (<X(0)> - <V(0)>)^2 = <\varepsilon>^2$$

Systematic model error

$$S'_1 = 2(<X(0)> - <V(0)>)(<F(\vec{X}(0)) - G(\vec{V}(0))>)$$

$$S'_2 = 2(<F(\vec{X}(0)) - G(\vec{V}(0))>)^2 + 2(<X(0)> - <V(0)>)(<\frac{d}{dt}(F(\vec{X}(0)) - G(\vec{V}(0)))\Big]_0 >)$$

Summary

A. Correction of initial condition errors

- Systematic initial error is corrected.

- Random initial errors are only partially corrected.

$$S_0 = \frac{\sigma^4_{\varepsilon(0)}}{\sigma^2_{\varepsilon(0)} + \sigma^2_{X(0)}}$$

B. Correction of model errors

- The systematic part is corrected.

- The time-dependent part can be corrected provided that the covariance of the model error with the predictors is high.

The Lorenz system

$$\frac{dx}{dt} = -y^2 - z^2 - ax + aF$$

$$\frac{dy}{dt} = xy - bxz - y + G$$

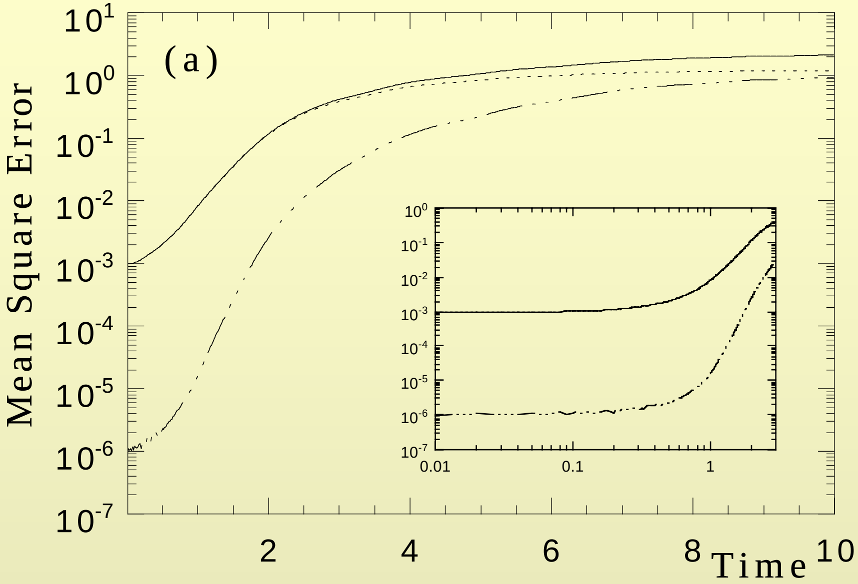
$$\frac{dz}{dt} = -bxy + xz - z$$

Chaotic for a=0.25, b=6, F=16, G=3

- Initial condition errors: N(0,s), s small
- Model errors (parametric error on a, b, F or G)

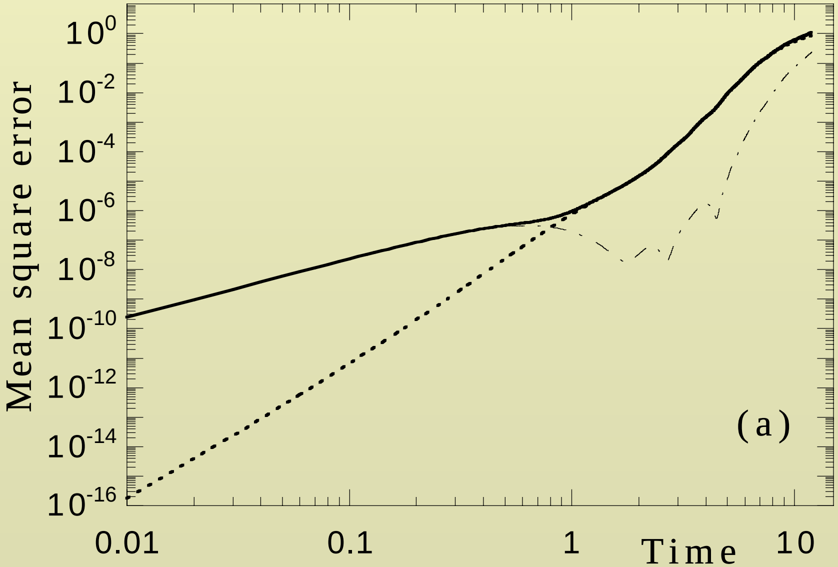
Results obtained with 100,000 realizations starting from different initial conditions

The Lorenz system (continued)



Correction of variable x

Initial condition error only



Error in parameter a

Model error only

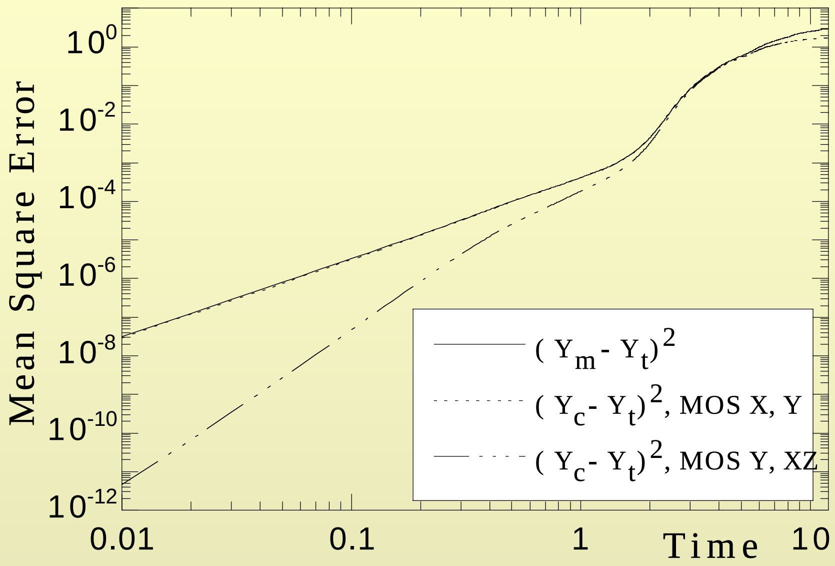
The Lorenz system (continued)

Model errors only in the parameter b

Correction of variable y

Different MOS schemes

- Predictors: X, Y
- Predictors: Y, XZ



Partial conclusions

- IC not well corrected by MOS
- MOS can correct model errors when model predictors well chosen

Application to the ECMWF forecasting system: temperature

Data: temperature for the period December 1 2001 to November 30 2005

- Training period: December 1 2001 to November 30 2003

- Independent evaluation period: December 1 2003 to November 30 2005

-Temperature 500 hPa, 850 hPa
-Temperature at 2 meters

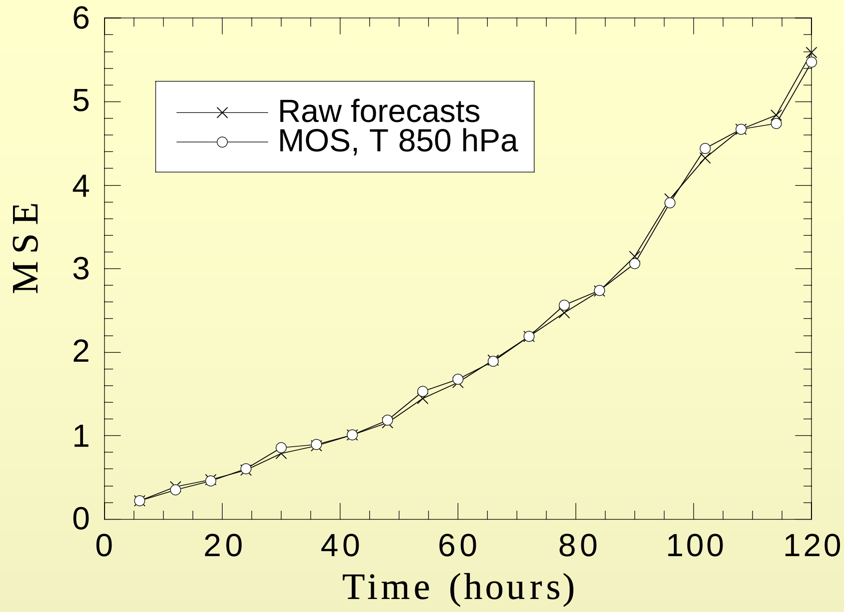
→Evaluation on a grid covering Belgium (verification using the set of analyses)

(52 N, 2 E) to (49 N, 7 E)

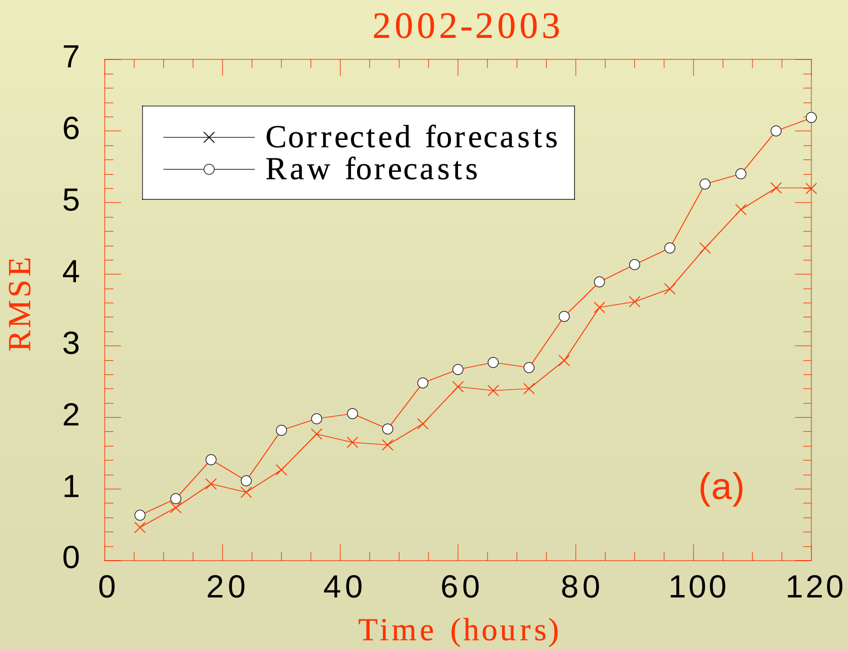
→Evaluation at some synoptic stations: model forecast given by the closest gridpoint

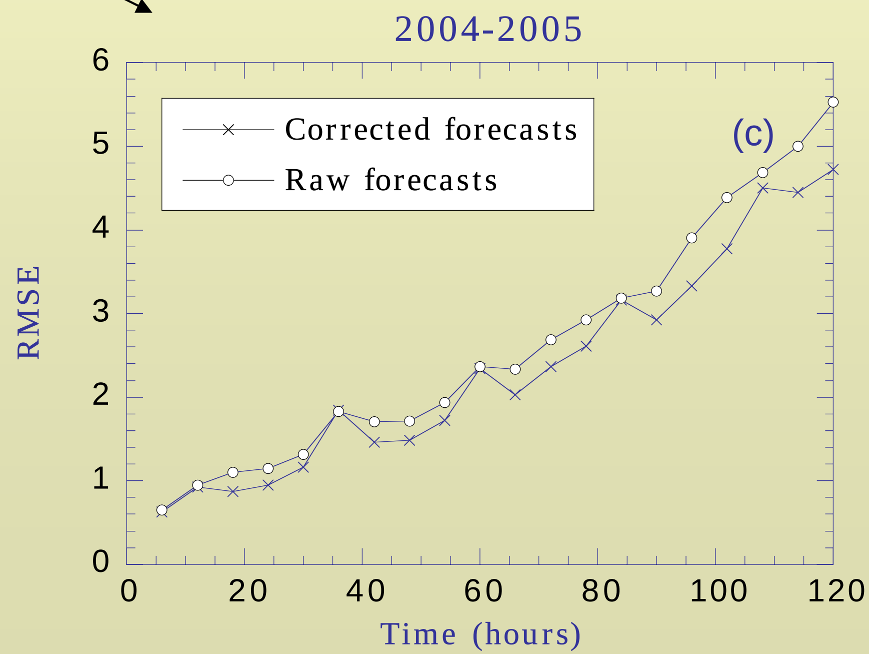
Temperature at 850 hPa

Results averaged over the whole grid

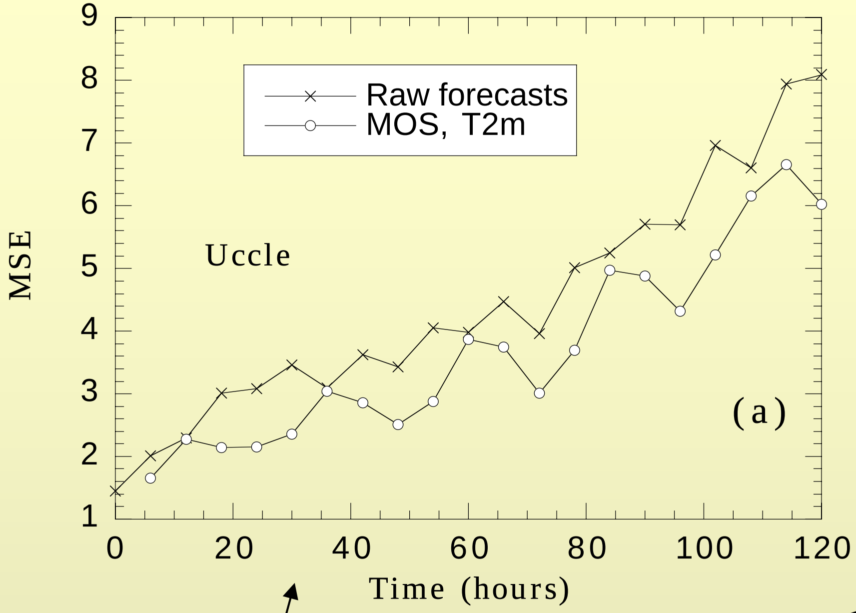


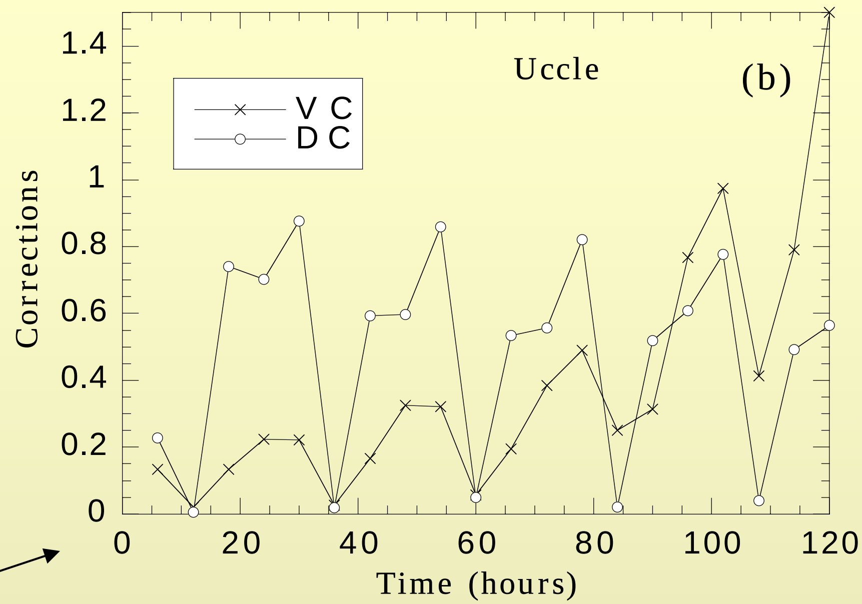
Temperature at 2 meters



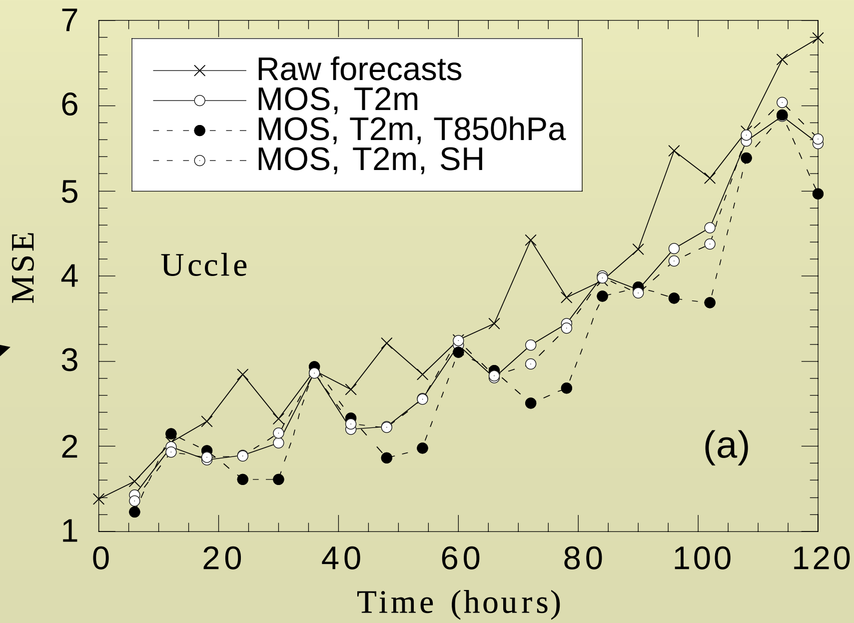


Uccle-Ukkel



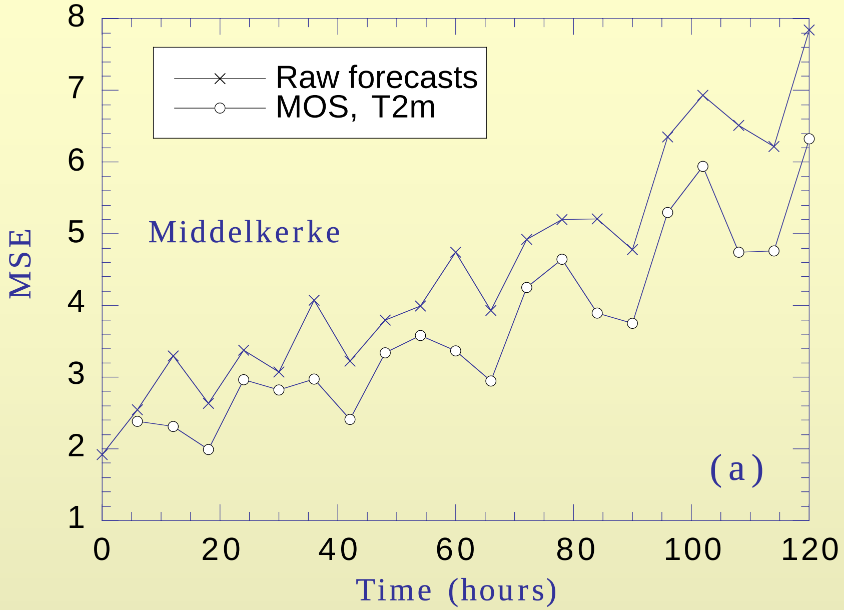


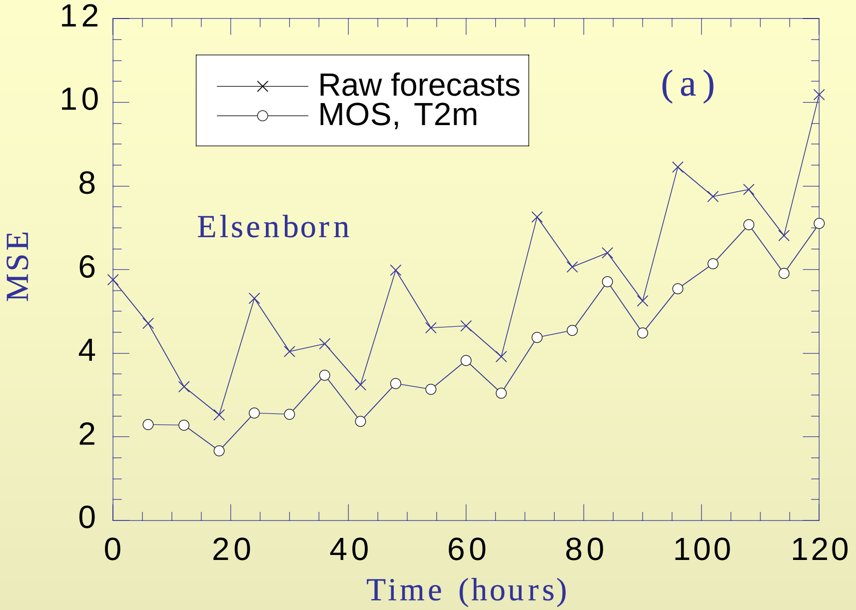
Training + verif on 2002-2003



Training: 2002-2003
Analysis on 2004-2005

Other stations





Training + verif on 2002-2003

A second predictor?

Sensible heat flux

A second predictor?

TMP 850 hPa

New Linear MOS

$$X = \alpha + \beta\varsigma + \kappa \qquad\qquad V = \varsigma + \delta$$

$$J(t) = \sum_{k=1}^{M} \frac{(V_k(t) - \varsigma_k(t))^2}{\sigma_\delta^2(t)} + \sum_{k=1}^{M} \frac{(X_k(t) - (\alpha + \beta\varsigma_k(t)))^2}{\sigma_\kappa^2(t)}$$

Intermediate cost function:

$$J(t) = \sum_{k=1}^{M} \frac{(\overbrace{(\alpha(t) + \beta(t)V_k(t))}^{X_{c,k}(t)} - X_k(t))^2}{\sigma_\kappa^2(t) + \beta^2(t)\sigma_\delta^2(t)}$$

One ‘free’ parameter:

$$\lambda = \frac{\sigma_\delta^2}{\sigma_\kappa^2} \text{ fixed to } \frac{\sigma_V^2}{\sigma_X^2}$$

Needs some knowledge about the sources of errors

Minimization

$$\alpha(t) = < X(t) > - \beta(t) < V(t) >$$

$$\beta(t) = \sqrt{\frac{< X(t)^2 > - < X(t) >^2}{< V(t)^2 > - < V(t) >^2}} = \sqrt{\frac{\sigma_X^2(t)}{\sigma_V^2(t)}}$$

Statistical Properties

$$< X_C(t) > = < X(t) >$$

$$\sigma_C^2(t) = < (X_C(t) - < X(t) >)^2 > = \beta(t)^2 \sigma_{V(t)}^2 = \sigma_{X(t)}^2$$

And the third moment,

$$< (X_C(t) - < X_C(t) >)^3 > = \beta^3(t) < (V(t) - < V(t) >)^3 >$$

This development can be extended to two predictors.

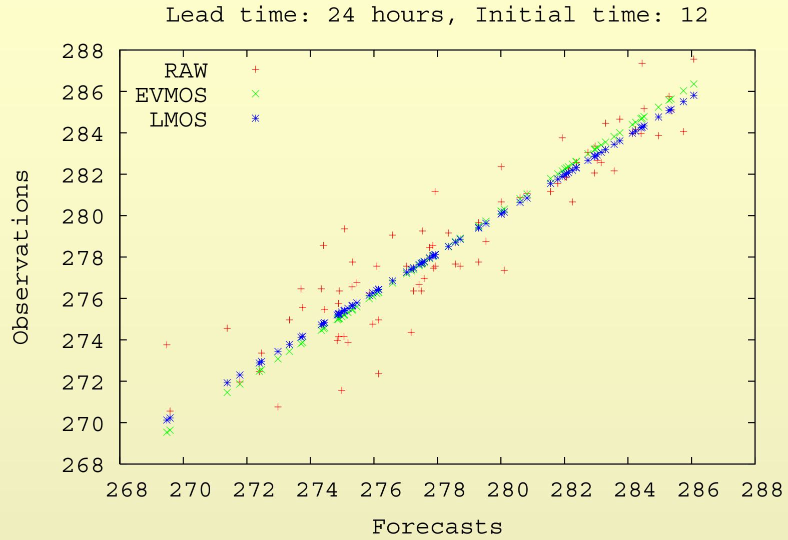
Vannitsem, S., A unified Linear Model Output Statistics scheme for both deterministic and ensemble forecasts. Accepted in *Quart. J. Roy. Met. Soc.*, 2009

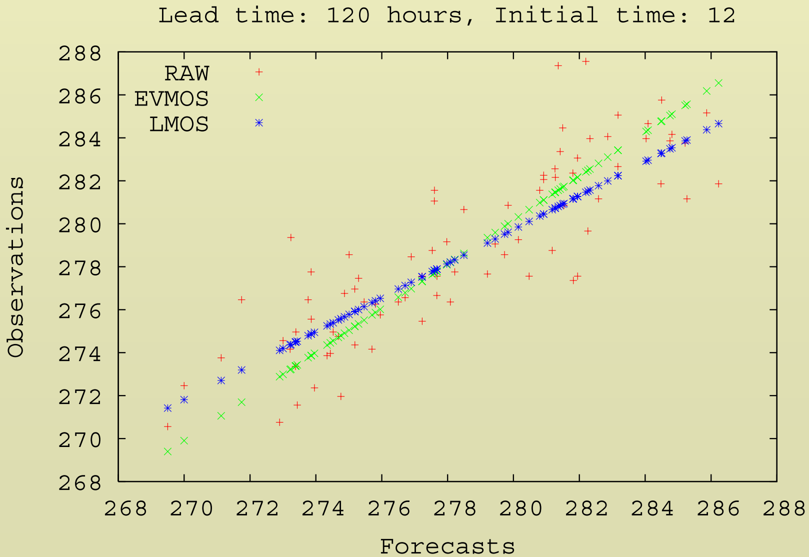
MOS at Station ‘Uccle’.

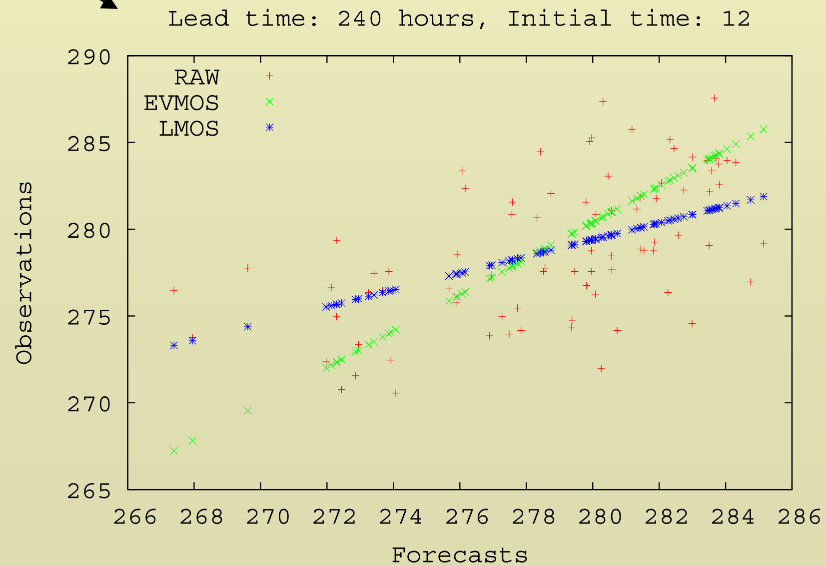
Control Forecast is used.

Training: DJF 2002-2004

Verif: DJF 2002

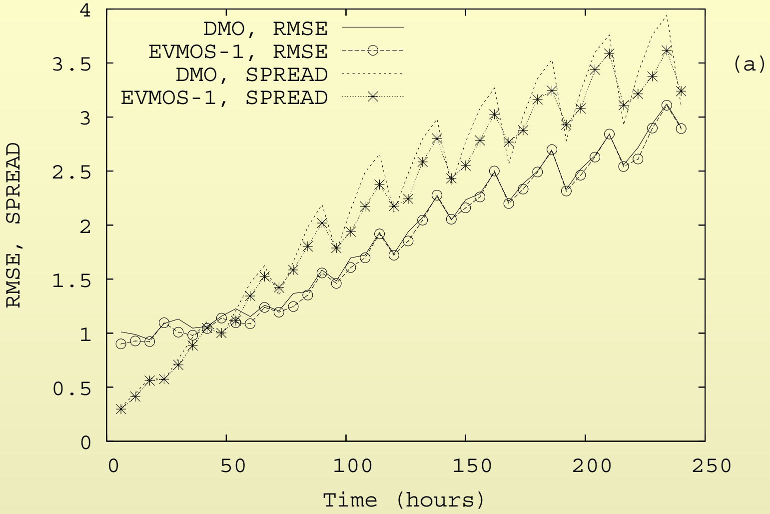


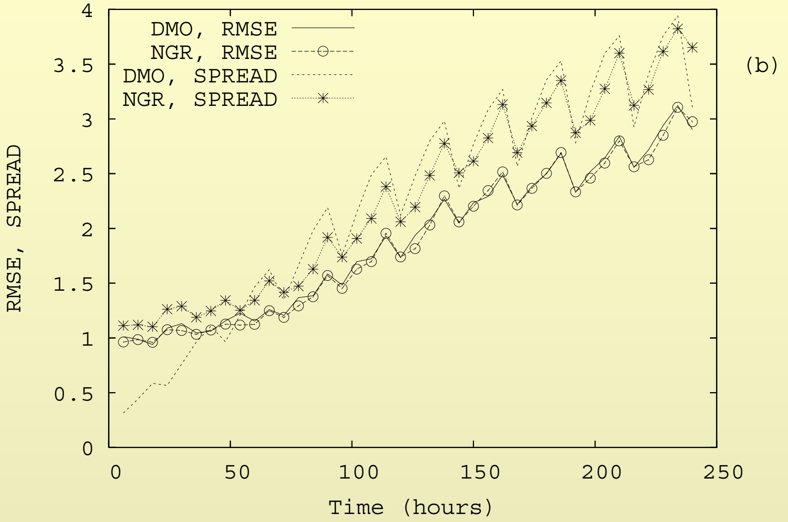


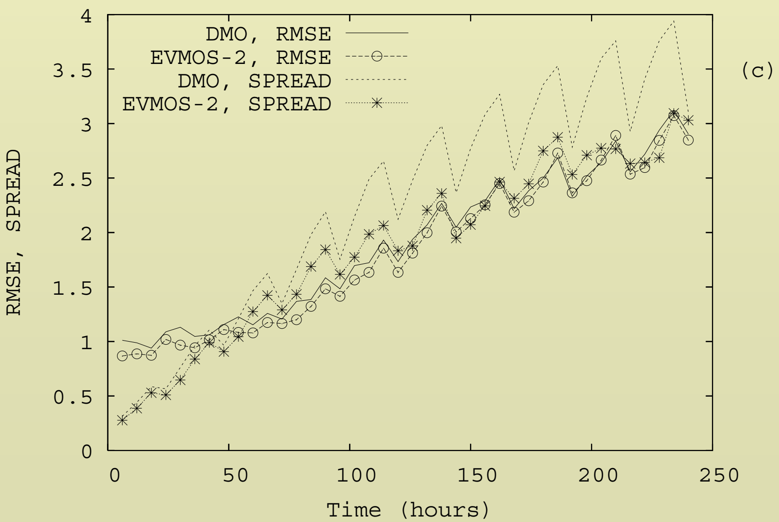


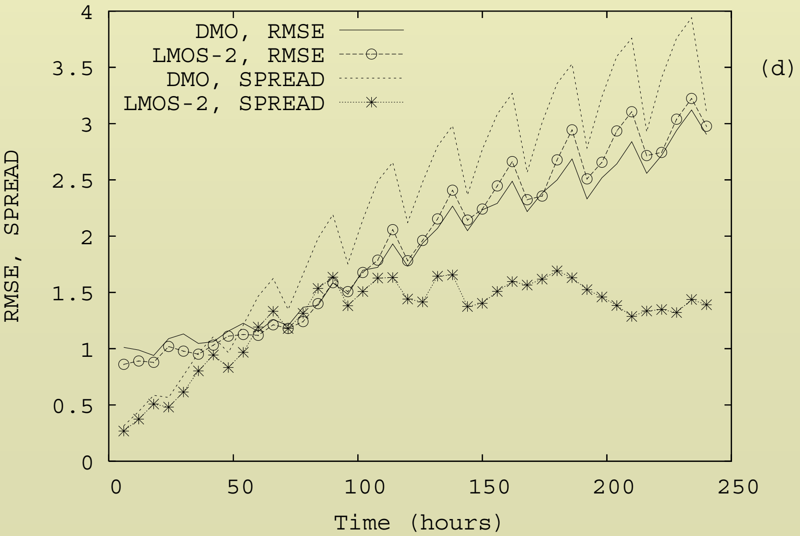
Application to an operational ensemble: ECMWF, Winter, Uccle

2 meter Temperature









Conclusions on MOS

A dynamical analysis of the MOS correction has been undertaken with emphasis on the correction of

- initial condition errors
- Model errors

Both partly corrected but correction more sensitive to model errors.

One central quantity

$$C(V(0), G(\vec{X}(0)) - F(\vec{X}(0)))$$

An additional model observable can improve considerably the correction if proportional to G-F

For the ECMWF deterministic forecast:

TMP 2m

- Substantial corrections with MOS with 1 observable
- MOS with 2 observables:
    - In the central part of Belgium: Tmp 850 hPa
    - On the coast: Sensible Heat Flux

TMP 850 hPa

- No correction. ➔ Mainly initial condition errors

More infos:

Vannitsem S. and C. Nicolis, Dynamical properties of Model Output Statistics forecasts. Mon. Wea. Rev., 136, 405-419, 2008.

Vannitsem S., Dynamical properties of MOS forecasts. Analysis of the ECMWF operational forecasting system. Weather and Forecasting, 23, 1032-1043, 2008.

Correction of Ensemble forecasts based on EVMOS

A unified Scheme for both deterministic and ensemble forecasts is proposed, based on the assumption that the forecast displays errors.

-Do not damp the ensemble spread anymore.

-The correction provides a mean and a variability in agreement with the observations

-The scheme can be trained on the control forecast only.

References:

Vannitsem, S., A unified Linear Model Output Statistics scheme for both deterministic and ensemble forecasts. Accepted in *Quart. J. Roy. Met. Soc.*, 2009