Background model error localization through an ensemble Kalman filter

F. Di Giuseppe
(M.Elementi, C.Marsigli, T.Paccagnella)

Servizio IdroMeteoClima -Bologna,IT

COSMO-GM 2009

F. Di Giuseppe, C Marsigli and T. Paccagnella: Background model error localization through an ensemble Kalman filter: QJRM (Submitted 2009)

Outline

Problem

Assimilation systems based on optimal interpolation techniques require the specification of model errors (**covariance matrix B):**

- **B depends on weather regimes**
- **B is scale dependent**

In regional models there are mainly two sources of errors:

Large scale errors (at driving model resolution) due to uncertainties in initial and boundary conditions;
Small scale errors (at model resolution) due to the low predictability of events at the meso-gamma scale.

Problem

Assimilation systems based on optimal interpolation techniques require the specification of model errors (**covariance matrix B):**

- **B depends on weather regimes**
- **B is scale dependent**

In regional models there are mainly two sources of errors:

Large scale errors (at driving model resolution) due to uncertainties in initial and boundary conditions;
Small scale errors (at model resolution) due to the low predictability of events at the meso-gamma scale.

Problem

Assimilation systems based on optimal interpolation techniques require the specification of model errors (**covariance matrix B):**

- **B depends on weather regimes**
- **B is scale dependent**

In regional models there are mainly two sources of errors:

Large scale errors (at driving model resolution) due to uncertainties in initial and boundary conditions;
Small scale errors (at model resolution) due to the low predictability of events at the meso-gamma scale.

B matrix calculations

Traditional approaches for estimating the background error covariance are built on background climatology (flow-independent)

- The Hollingsworth and Lönnberg (1986) method. Long time-series of comparison model observations
- The NMC method (Parrish and Derber,1992). Long time-series of differences between 48h and 24h forecasts verifying at the same time.
- The analysis–ensemble method.

  Calculates statistics from pairs of backgrounds generated from perturbed analysis after few days from the initialisation.

Road-map

WHAT: Develop a simple technique to account both for the **time-dependent** and **space-dependent** components of the model error.

- The **time-dependency** is accounted for by calculating **B** at the beginning of each assimilation cycle
- The **space-dependency** is accounted for by a localization procedure

HOW:

- Use a multi-analysis model-perturbed ensemble system to identify areas of homogeneous spread ("island") where model errors are likely to be homogeneous.
- Calculate flow-dependent **B** matrices inside the selected islands.

WHERE:

The new estimation of error covariance matrices are used to improve 1D variational retrieval of temperature and humidi profiles from MSG-SEVIRI radiances which are then assimilated into the nudging scheme of the regional model COSMO

Road-map

WHAT: Develop a simple technique to account both for the **time-dependent** and **space-dependent** components of the model error.

- The **time-dependency** is accounted for by calculating **B** at the beginning of each assimilation cycle
- The **space-dependency** is accounted for by a localization procedure

HOW:

- Use a multi-analysis model-perturbed ensemble system to identify areas of homogeneous spread ("island") where model errors are likely to be homogeneous.
- Calculate flow-dependent **B** matrices inside the selected islands.

WHERE:

The new estimation of error covariance matrices are used to improve 1D variational retrieval of temperature and humidi profiles from MSG-SEVIRI radiances which are then assimilated into the nudging scheme of the regional model COSMO

Road-map

WHAT: Develop a simple technique to account both for the **time-dependent** and **space-dependent** components of the model error.

- The **time-dependency** is accounted for by calculating **B** at the beginning of each assimilation cycle
- The **space-dependency** is accounted for by a localization procedure

HOW:

- Use a multi-analysis model-perturbed ensemble system to identify areas of homogeneous spread ("island") where model errors are likely to be homogeneous.
- Calculate flow-dependent **B** matrices inside the selected islands.

WHERE:

The new estimation of error covariance matrices are used to improve 1D variational retrieval of temperature and humidi profiles from MSG-SEVIRI radiances which are then assimilated into the nudging scheme of the regional model COSMO

Road-map

WHAT: Develop a simple technique to account both for the **time-dependent** and **space-dependent** components of the model error.

- The **time-dependency** is accounted for by calculating **B** at the beginning of each assimilation cycle
- The **space-dependency** is accounted for by a localization procedure

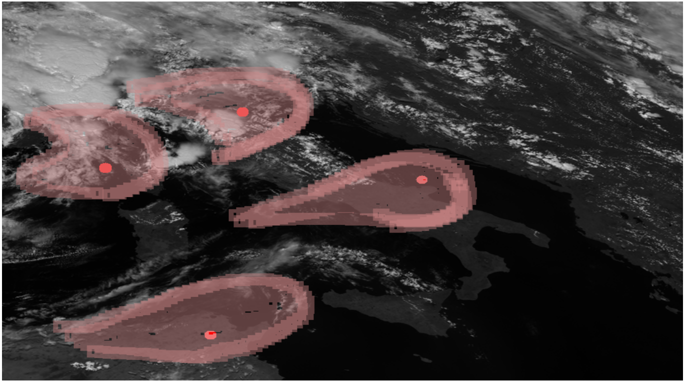
HOW:

- Use a multi-analysis model-perturbed ensemble system to identify areas of homogeneous spread ("island") where model errors are likely to be homogeneous.
- Calculate flow-dependent **B** matrices inside the selected islands.

WHERE:

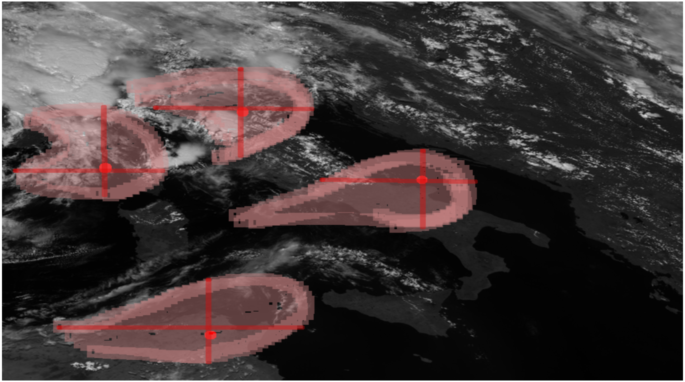
The new estimation of error covariance matrices are used to improve 1D variational retrieval of temperature and humidi profiles from MSG-SEVIRI radiances which are then assimilated into the nudging scheme of the regional model COSMO

Sketch of the methodology



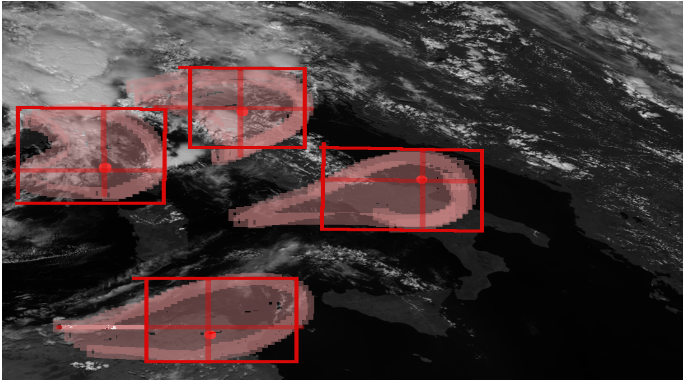
A two dimensional input spread field is computed: points with *spread* $>$ *percmax* are identified as central points for possible islands.

Sketch of the methodology



Then, the search for the island borders proceeds from this point in each of the four cardinal directions until the spread field reaches a value lower than *percmin*

Sketch of the methodology



... the island is completed by filling a rectangular shape

Implementation I

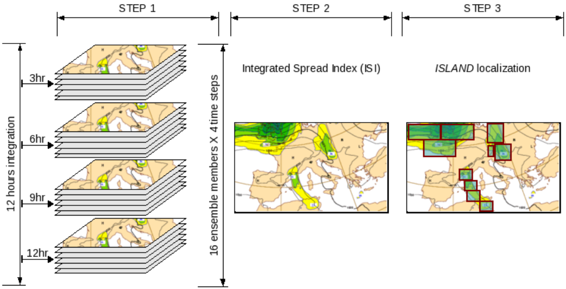
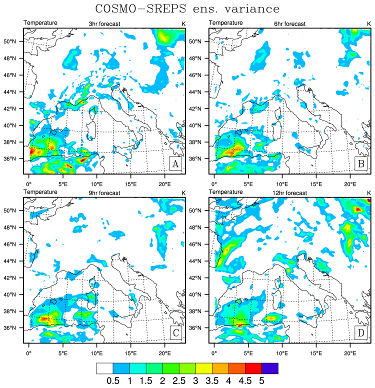
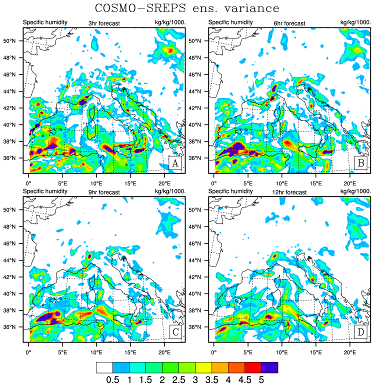


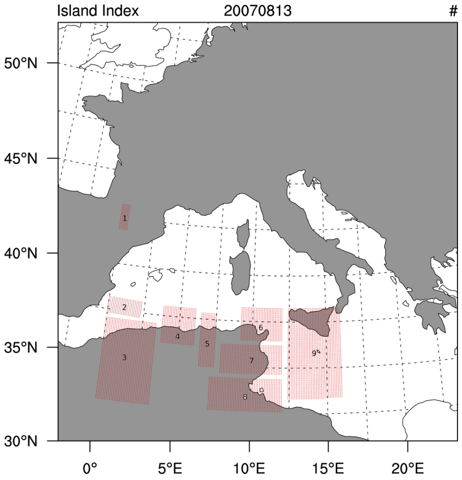
Figure: STEP 1 : generation of ensemble spread fields from the COSMO-SREPS system. STEP 2 : construction of a two-dimensional spread index field. STEP 3 : localization procedure based on a percentiles-based method using the spread index field calculated in STEP 2.

Example


COSMO−SREPS ens. variance
Temperature 3hr forecast K
Temperature 6hr forecast K
Temperature 9hr forecast K
Temperature 12hr forecast K
A
B
C
D
50°N 48°N 46°N 44°N 42°N 40°N 38°N 36°N
0° 5°E 10°E 15°E 20°E
0.5 1 1.5 2 2.5 3 3.5 4 4.5 5


COSMO−SREPS ens. variance
Specific humidity 3hr forecast kg/kg/1000.
Specific humidity 6hr forecast kg/kg/1000.
Specific humidity 9hr forecast kg/kg/1000.
Specific humidity 12hr forecast kg/kg/1000.
A
B
C
D
50°N 48°N 46°N 44°N 42°N 40°N 38°N 36°N
0° 5°E 10°E 15°E 20°E
0.5 1 1.5 2 2.5 3 3.5 4 4.5 5

Example:ISLANDS



Error Profiles

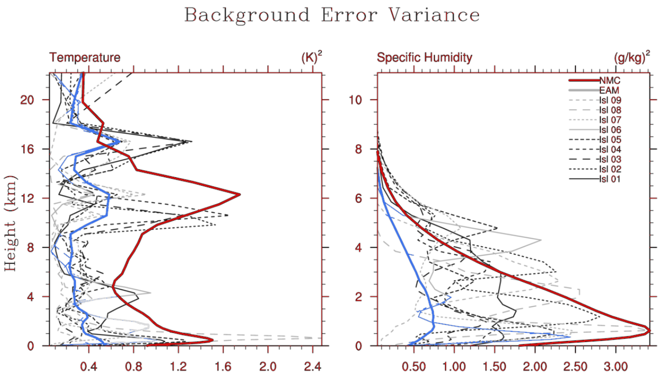


Figure: Temperature and specific humidity variances for the three background error models considered. One day of example is considered, the 13th of August 2007.

Validation Experiments

This new characterization of model error is actually producing better retrieved profiles?

The new approach is compared to state of the art methods

Table: Summary of the configurations and main characteristics of the background error models used in the comparison.

Exp	Method	Time	Space	Comments
ISL	Ensemble	Daily upgrade at the beginning of each assimilation cycle	Spatial localization	Variable number of $\mathcal{B}$ matrices are applied depending on the observation location in the domain
EAM	Ensemble	Daily upgrade at the beginning of each assimilation cycle	Domain spatial average	One $\mathcal{B}$ matrix available for all observations at the beginning of the assimilation cycle
NMC	Statistics based on model departures calculated at +36hrs and +12 hrs forecast times	Seasonal mean (June, July, August)	Domain spatial average	One climatological $\mathcal{B}$ matrix available for all observations and during the whole period.

Validation Experiments

This new characterization of model error is actually producing better retrieved profiles?
The new approach is compared to state of the art methods

Table: Summary of the configurations and main characteristics of the background error models used in the comparison.

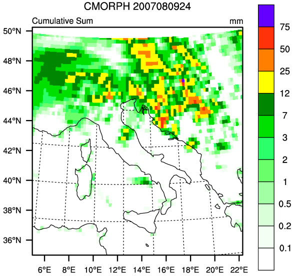
Exp	Method	Time	Space	Comments
ISL	Ensemble	Daily upgrade at the beginning of each assimilation cycle	Spatial localization	Variable number of $\mathcal{B}$ matrices are applied depending on the observation location in the domain
EAM	Ensemble	Daily upgrade at the beginning of each assimilation cycle	Domain spatial average	One $\mathcal{B}$ matrix available for all observations at the beginning of the assimilation cycle
NMC	Statistics based on model departures calculated at +36hrs and +12 hrs forecast times	Seasonal mean (June, July, August)	Domain spatial average	One climatological $\mathcal{B}$ matrix available for all observations and during the whole period.

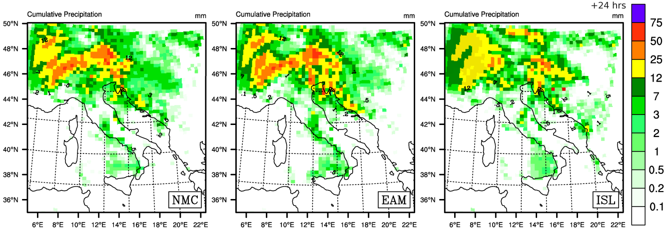
Comparison to observations

Two comparisons are performed:

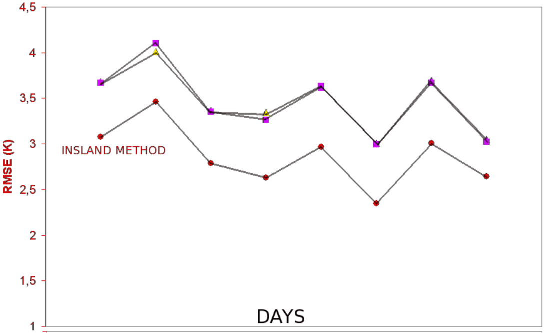
- Forecasted precipitation fields are compared to an independent satellite product,
- 2m surface temperatures are verified against the synop network.

Precipitation





T_{2m} scores



+24 hrs forecast. RMSE relative to the 2m temperature analyses obtained for the eight days considered, at 00 UTC.

Summary

- A simple technique is presented to calculate “flow dependent” error covariance matrices using the outputs of an multi-model perturbed ensemble system
- It has been applied to the retrieval of temperature and humidity profiles from the SEVIRI sensor
- Preliminary tests show that the use of more appropriate model error functions can be beneficial in improving retrieval performance
- A positive impact is also shown in forecast errors (RMS)
- Further impact studies are ongoing and to maximise the effect of the methodology we plan to increase to 3hrs the time step for the **B** calculation

The ensemble spread index (ESI)

The ensemble spread index (ESI) is computed in terms of temperature and specific humidity on model levels (80 fields) by

1. calculating the departure of each ensemble member from the ensemble mean
2. normalising the departures of the different variables
3. convolving the departures with a vertical weighting function to obtain a unique integrated spread field.