The Ensemble Kalman Filter in the CNMCA regional NWP System: Tuning and Tests in a Pre-Operational Configuration

Lucio Torrisi, Francesca Marcucci, Massimo Bonavita[1]

CNMCA, National Meteorological Center,Italy

[1] Current affiliation: ECMWF, Reading (UK)

4th ENKF WS, 6-9 April 2010

Outline

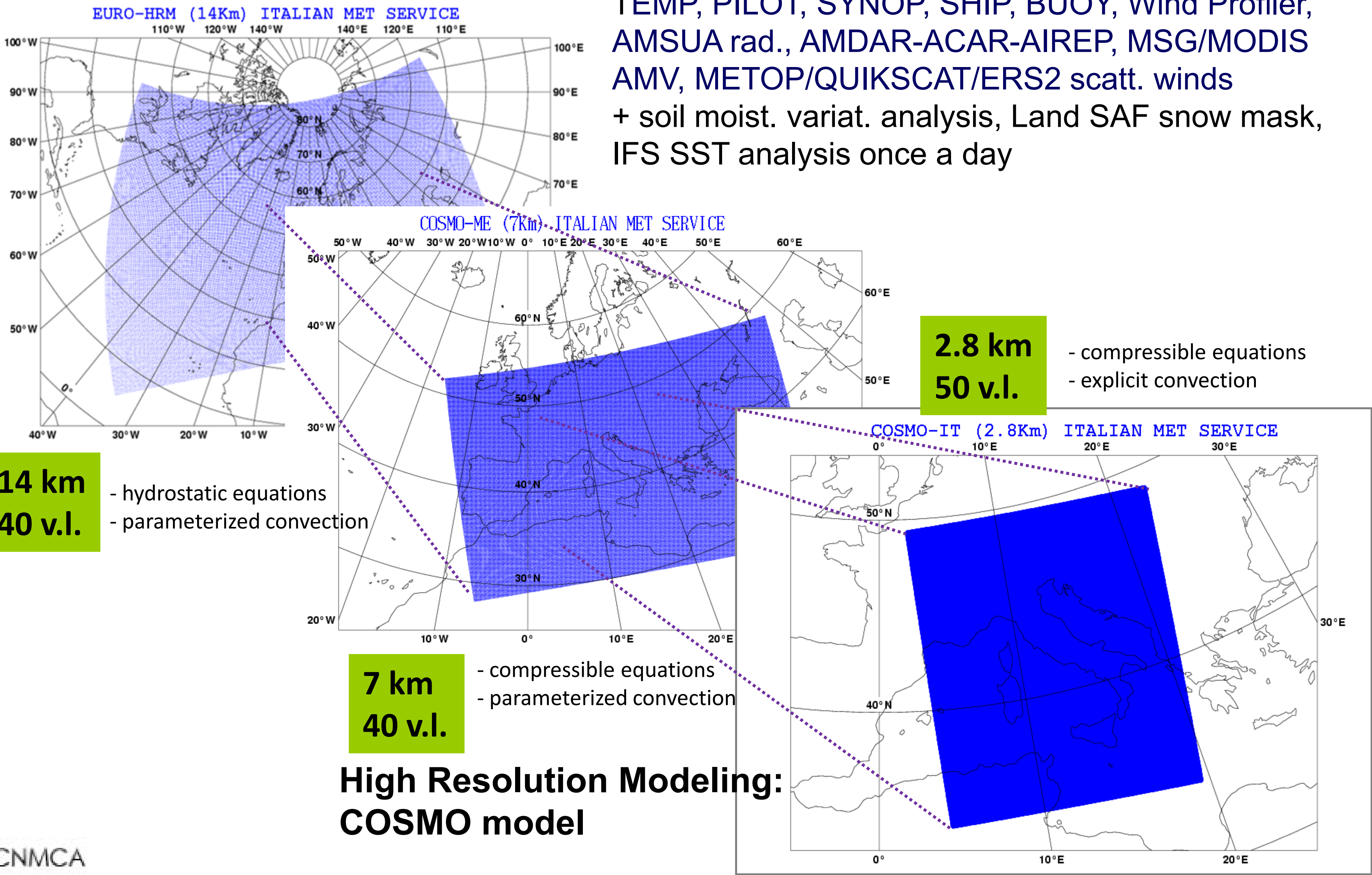
- CNMCA LETKF implementation
  - setup and model error treatment
- Stochastic physics schemes

- version 1: description and results
- version 2: description

- Future developments

Operational NWP System

Data Assimilation: 3DVAR PSAS FGAT (T,u,v,qv,ps) every 3h using TEMP, PILOT, SYNOP, SHIP, BUOY, Wind Profiler, AMSUA rad., AMDAR-ACAR-AIREP, MSG/MODIS AMV, METOP/QUIKSCAT/ERS2 scatt. winds
+ soil moist. variat. analysis, Land SAF snow mask, IFS SST analysis once a day



14 km
40 v.l.

- hydrostatic equations
- parameterized convection

7 km
40 v.l.

- compressible equations
- parameterized convection

2.8 km
50 v.l.

- compressible equations
- explicit convection

High Resolution Modeling: COSMO model

Ensemble Kalman Filter D.A.

Uses an ensemble of N system states to parametrize the distribution

$$P^b = \frac{1}{N-1} X^b X^{b^T} \qquad X^b = x^b - \overline{x}^b$$

It follows the time evolution of the mean and covariance (Gaussian assumption) by propagating the ensemble of states

LETKF FORMULATION *(Hunt et al,2007)*

$$\tilde{H}_n = H(x_n{}^b) - \overline{H}(x^b)$$

$$\tilde{P}^a = \left[(\tilde{H}^T R^{-1} \tilde{H} + (N-1)I\right]^{-1}$$

$$K = X^b \tilde{P}^a \tilde{H}^T R^{-1}$$

$$X^a = X^b W^a$$

$$W^a = \left[(N-1)\tilde{P}^a\right]^{1/2} \quad (\textit{Square root filter})$$

$$\overline{x}^a = \overline{x}^b + K(y^o - H(\overline{x}^b))$$

$$x^a = \overline{x}^a + X^a$$

The analysis ensemble mean is the linear combination of forecast ensemble states which best fits the observational dataset

Model Error Treatment

total forecast errors = internal errors + external error

$$\mathbf{P}_{\mathrm{i}}^{b} = \mathbf{M}_{x_{i-1}^{a}} \mathbf{P}_{\mathrm{i}}^{\mathrm{a}} \mathbf{M}_{x_{i-1}^{a}}^{\mathrm{T}} + \boxed{\mathbf{Q}}$$

Errors in initial state and their dynamical growth

Model deficiencies (model error)

BUT ensemble spread only represents growth of initial condition errors

POSSIBLE SOLUTIONS

MULTIPLICATIVE INFLATION

$$Q = 0 \qquad P^{b} \rightarrow (1+\Delta) P^{b}$$

$$(1+\Delta) = \frac{d^{T}_{o-b} d_{o-b} - Tr(R)}{Tr(HP^{b}H)}$$

+ temporal smoothing algorithm simple scalar Kalman filter approach (Li et al,2007)

ADDITIVE INFLATION

$$x_{i}^{a} = x^{a}{}_{e(i)} + rq_{i}$$

$$\overline{q}_{i} = 0$$

Additive perturbation derived from randomly selected, scaled 24-hour forecast differences

CNMCA LETKF Implementation

(OLD Implementation)

- 30 member ensemble at 0.25 (~28Km) grid spacing (EURO-HRM domain), 40 hybrid p-sigma vertical levels (top at 10 hPa)
- Initial ensemble from different EURO-HRM forecasts valid in ± 48h around start time and boundary conditions from IFS for all members (not perturbed)
- 6-hourly assimilation cycle run and (T,u,v,qv,ps) as a set of control variables
- Operational 3DVar cycle run in parallel at same spatial resolution
- Observations: RAOB (Tuvqv), SYNOP(ps), SHIP(ps), BUOY(ps), AIREP, AMDAR, ACAR, AMV, MODIS, WPROF
- Horizontal localization with 800 Km circular local patches (obs weight smoothly decay $\propto r^{-1}$)
- Vertical localization to layers whose depth increases from 0.2 scale heights at the lowest model levels to 2 scale heights at the model top.

It was found that the optimal configuration for the CNMCA LETKF implementation is
3D ADAPTIVE MULTIPLICATIVE + ADDITIVE INFLATION FACTOR

More details in:

Bonavita M, Torrisi L, Marcucci F. 2008. The ensemble Kalman filter in an operational regional NWP system: Preliminary results with real observations. *Q. J. R. Meteorol. Soc.* **134**: 1733-1744.

Bonavita M, Torrisi L, Marcucci F. 2010. Ensemble data assimilation with the CNMCA regional forecasting system. *Q. J. R. Meteorol. Soc.* **136**: 132-145.

Stochastic Physics

- LETKF outperforms 3DVAR (Bonavita et al, 2010), but forecast runs initialized by LETKF have larger biases than those starting from 3DVAR → Necessity of bias correction
- Model uncertainty could be represented also with a stochastic physics scheme (Buizza et al, 1999; Palmer et al, 2009) implemented in the prognostic model
- This scheme perturbs physics tendencies by adding perturbations, which are proportional in amplitude to the unperturbed tendencies X_c:

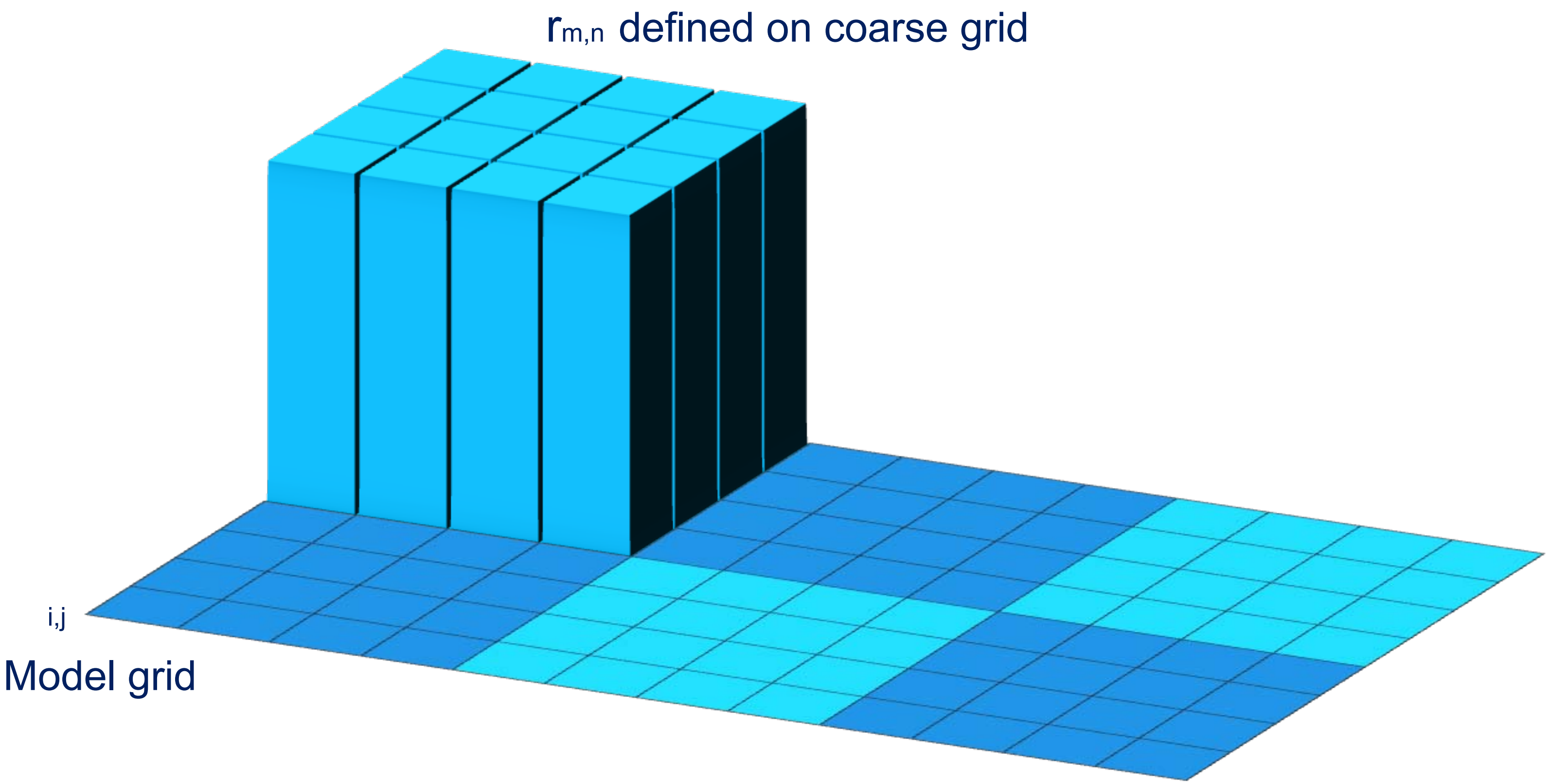
$X_p=(1+r\mu)X_c$

where r is a random number and μ is a tapering factor (μ=1 in Buizza et al, 1999)

Stochastic Physics

According to Buizza et al, 1999

Spatial correlation is imposed using the same r in a whole column and drawing r for a coarse grid with spacing DL (boxes)



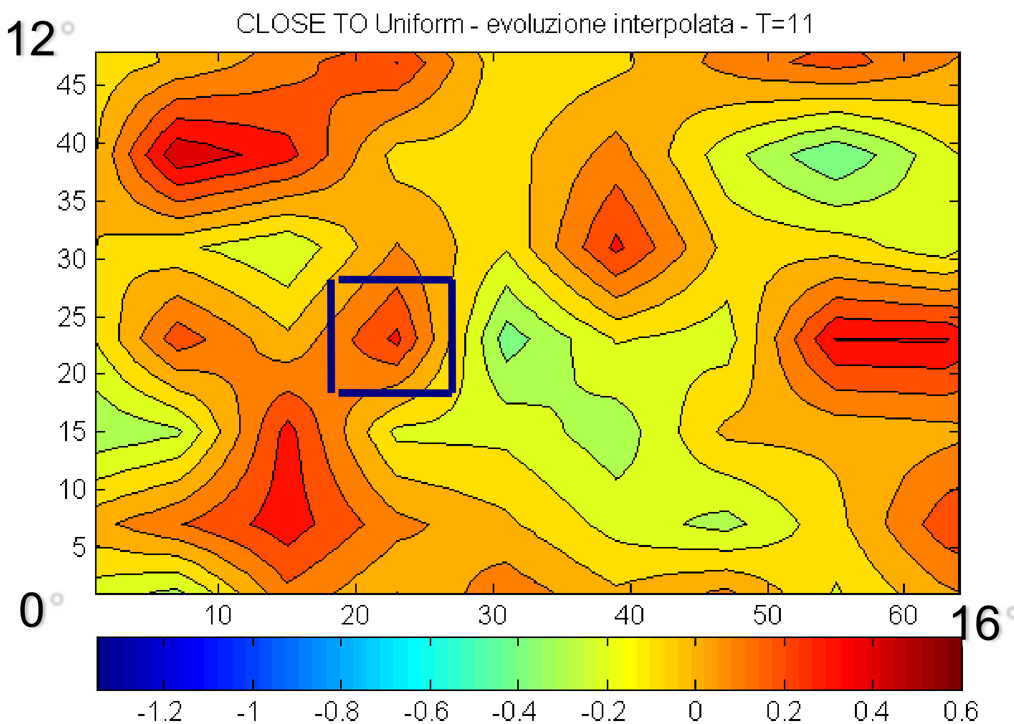
Modified to have a smoother pattern horizontally and to reduce the perturbation close to the surface and in the stratosphere

Modified Stochastic Physics

Version 1

Toy model and plots by A. Cheloni

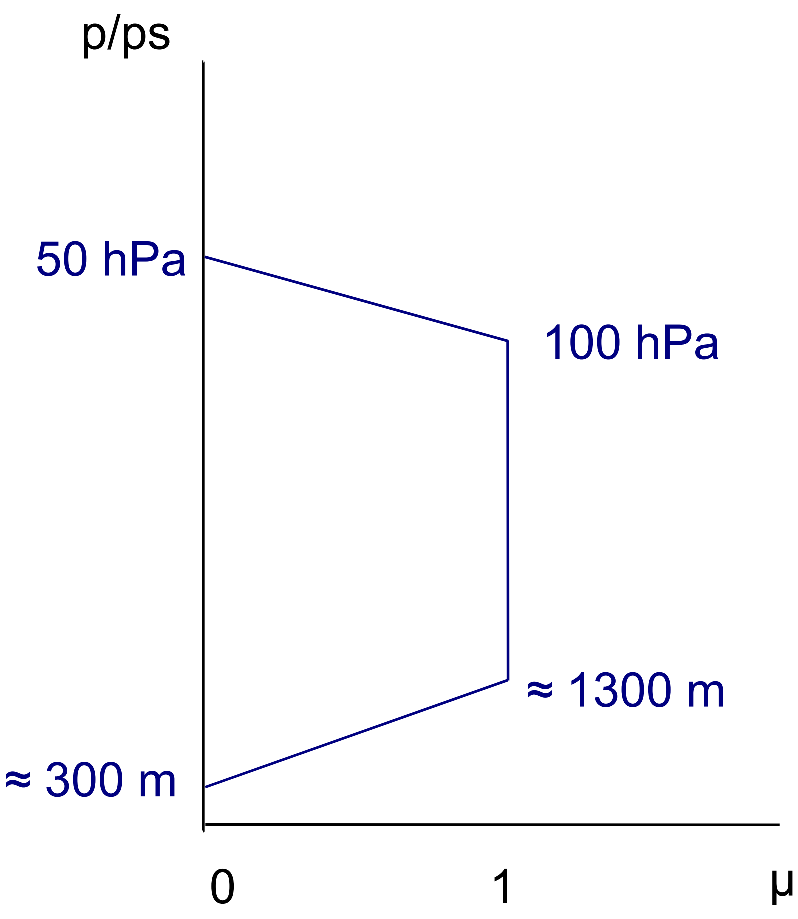
Model grid spacing: 0.25° (28 km)



2.5° coarse grid with bilin. interp.

Modified Stochastic Physics

Version 1



According to Palmer et al, 2009:
"…. For reasons of numerical stability and physical realism, the perturbations have been tapered to zero in the lowermost atmosphere and in the stratosphere.

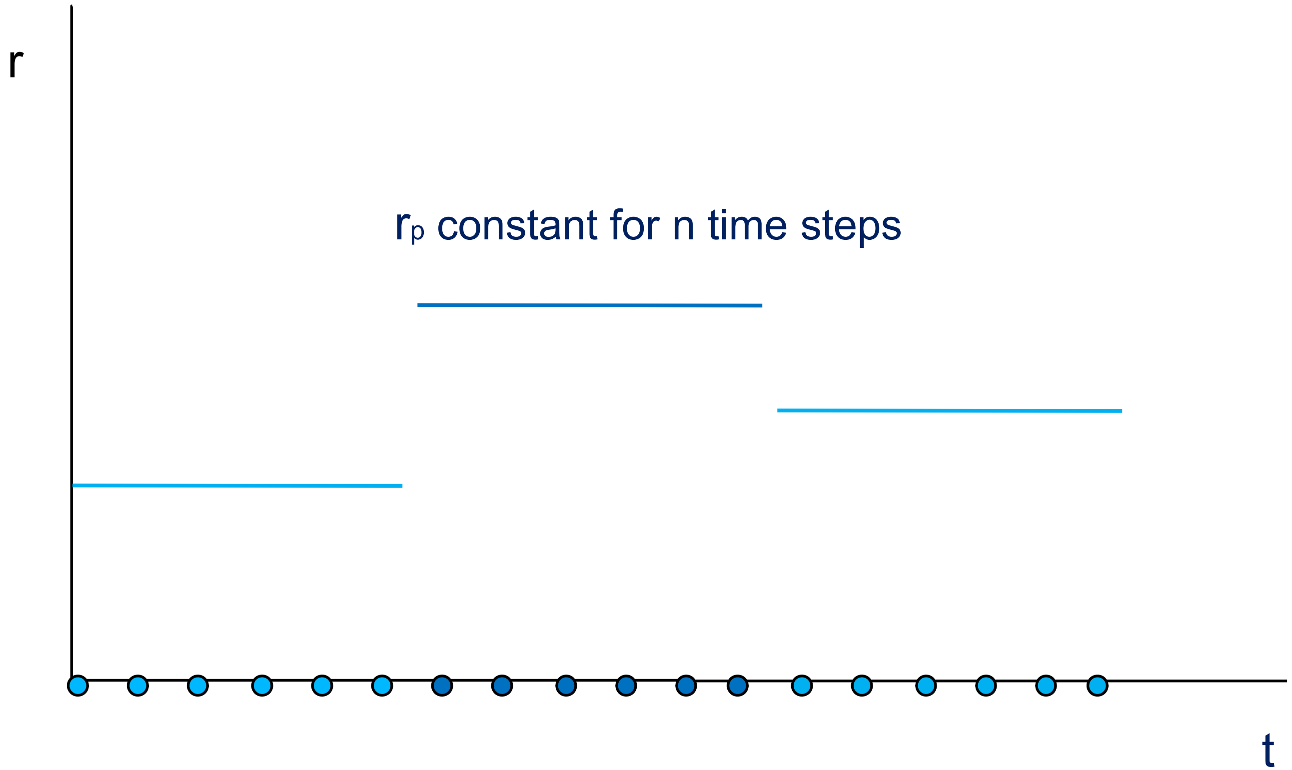
- In initial tests, tendencies were perturbed in the entire atmosphere. For standard deviations of 0.5, numerical instabilities were encountered. Further testing showed that the cause of the numerical instability are the perturbations in the lowermost part of the atmosphere.
- Radiative tendencies are expected to be relatively accurate in the stratosphere and with errors that are predominantly large scale, i.e. with correlation lengths far larger than 500 km….. "

$$X_p=(1+r\mu)X_c$$

Stochastic Physics

According to Buizza et al, 1999

Perturbations are multivariate (different r for u,v,T,qv). Temporal correlation is achieved by drawing r every n time steps (Dt)



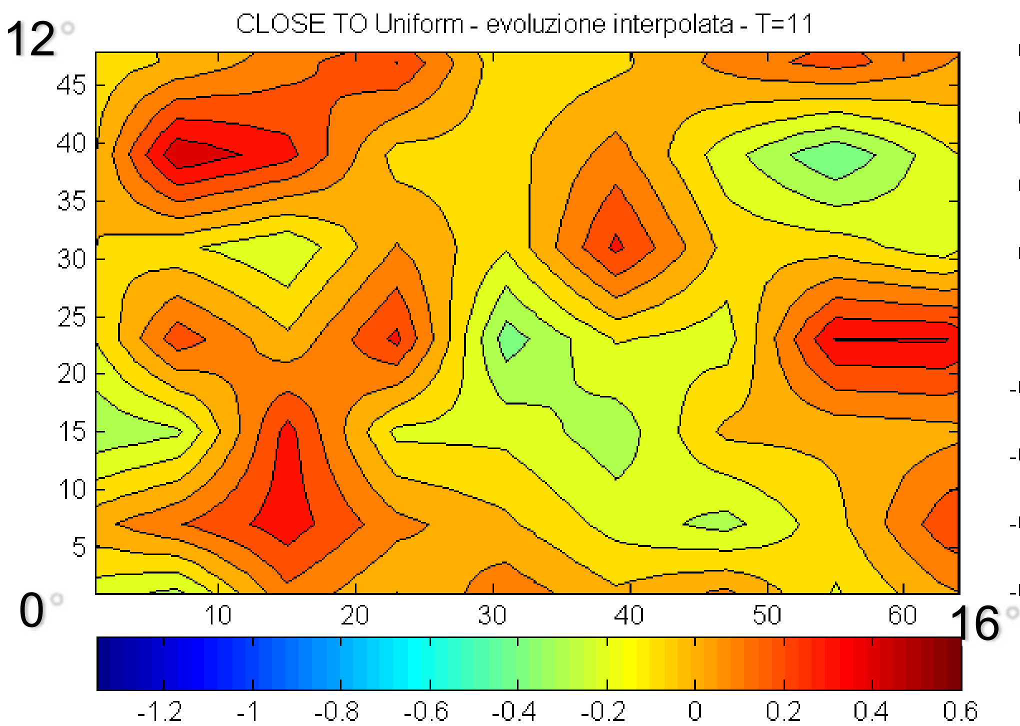
Modified to have a univariate distribution (as in Palmer et al, 1999) and a smoother pattern in time

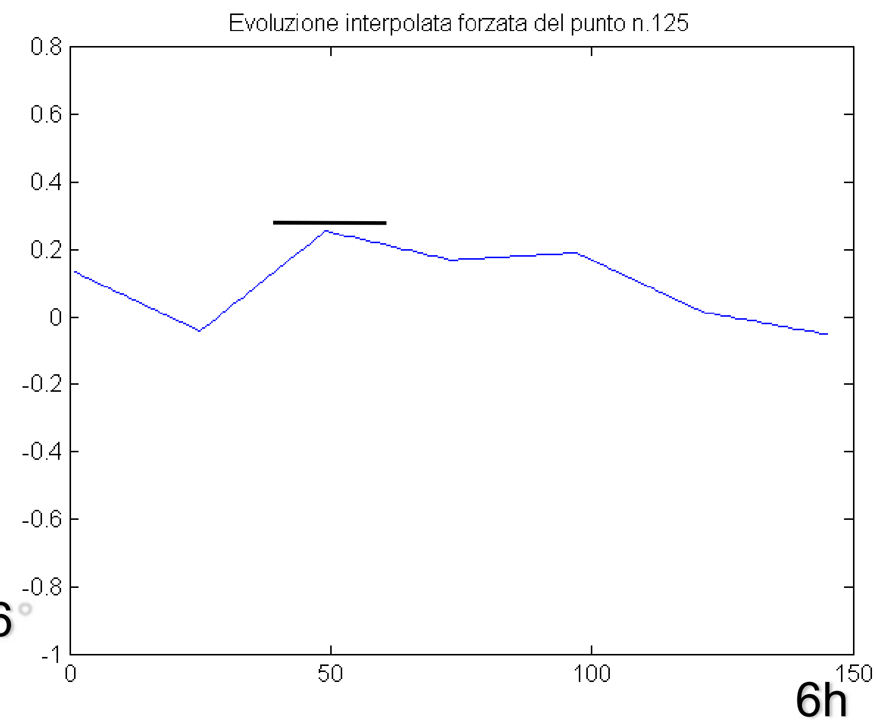
Modified Stochastic Physics

Version 1

Toy model and plots by A. Cheloni

Model grid spacing: 0.25° (28 km)   Time step: 150 s





2.5° coarse grid with bilin. interp.   1h coarse time grid with lin. interp.

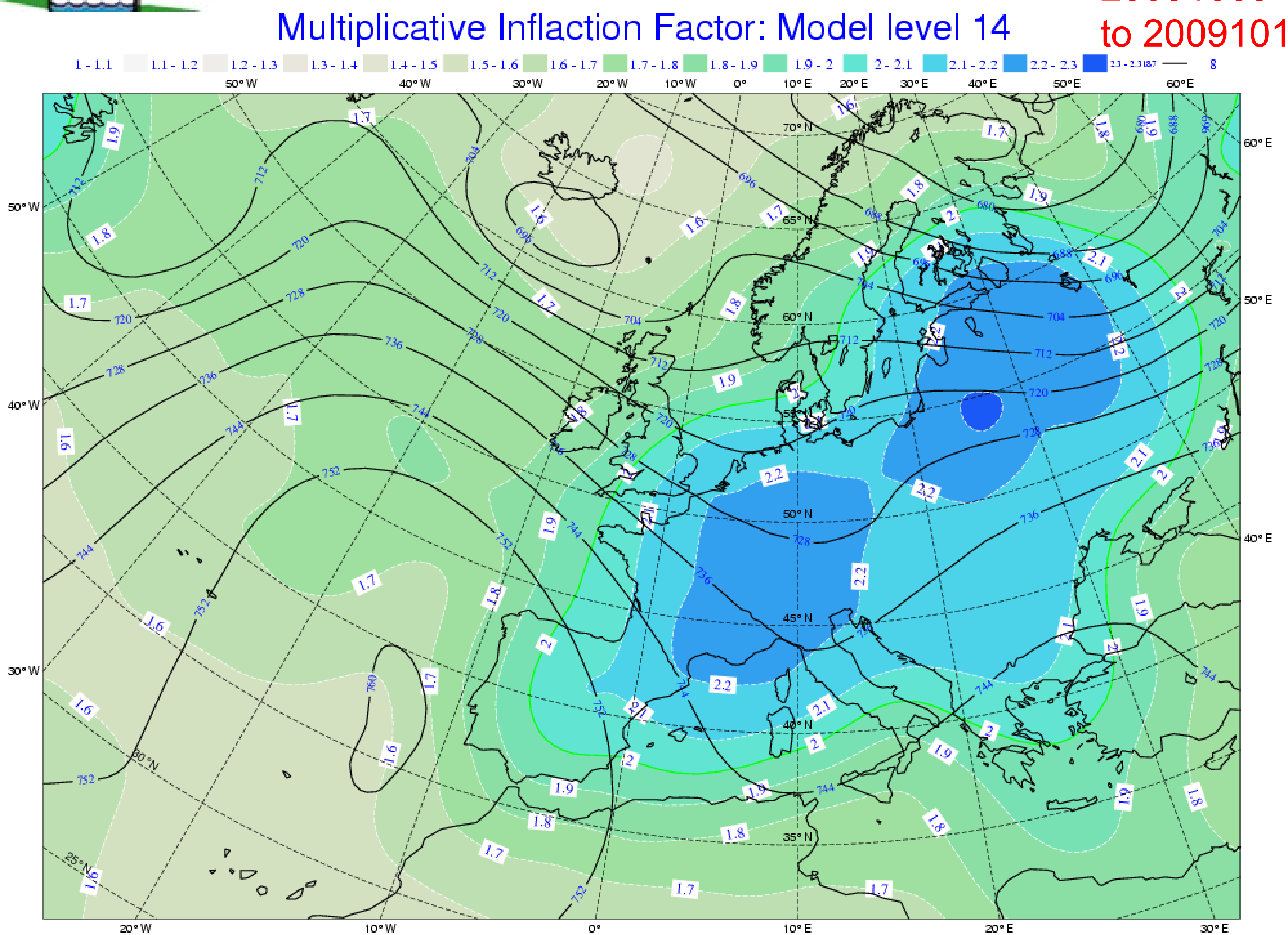
Modified Stochastic Physics

Version 1 (in blue, differences from Buizza et al, 1999)

- For all variables (u,v,T,qv), the random numbers r are drawn from a uniform distribution in the range [-0.5,0.5]
- A tapering factor μ is used to reduce r close to the surface and in the stratosphere (Palmer et al, 2009)
- The perturbations of T and qv are not applied if they lead to particular humidity values (exceeding the critical saturation value or negative values)
- Spatial correlation is imposed using the same r in a whole column and drawing r for a coarse grid with spacing DL (boxes); then they are *bilinearly interpolated* on the finer grid to have a smooth pattern in space
- Temporal correlation is achieved by drawing r every n time steps (Dt); then they are *linearly interpolated* for the intermediate steps to have a smooth pattern in time

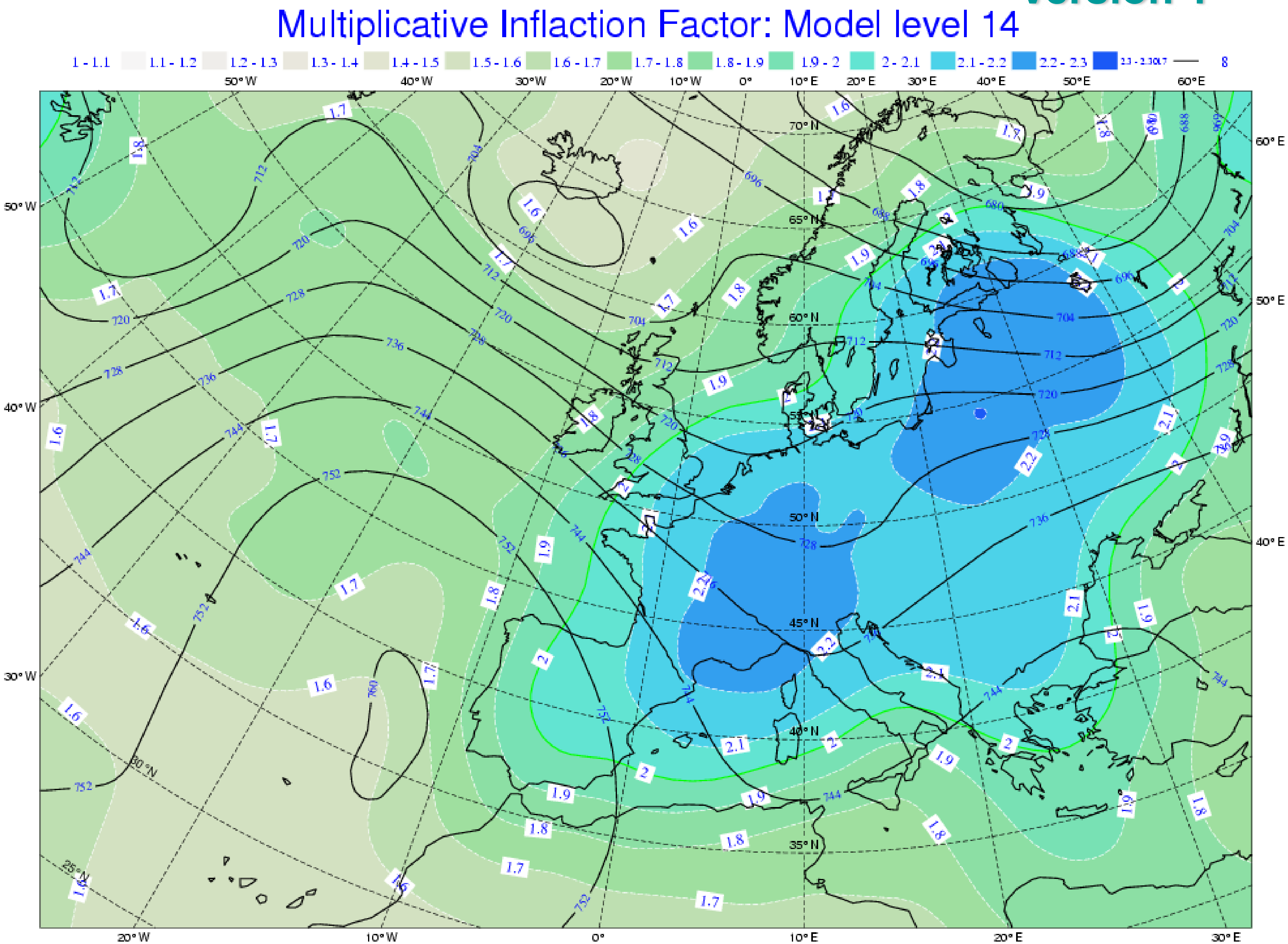
No Stochastic Physics

Averaged from 20091009-18 to 20091011-06



Modified Stochastic Physics

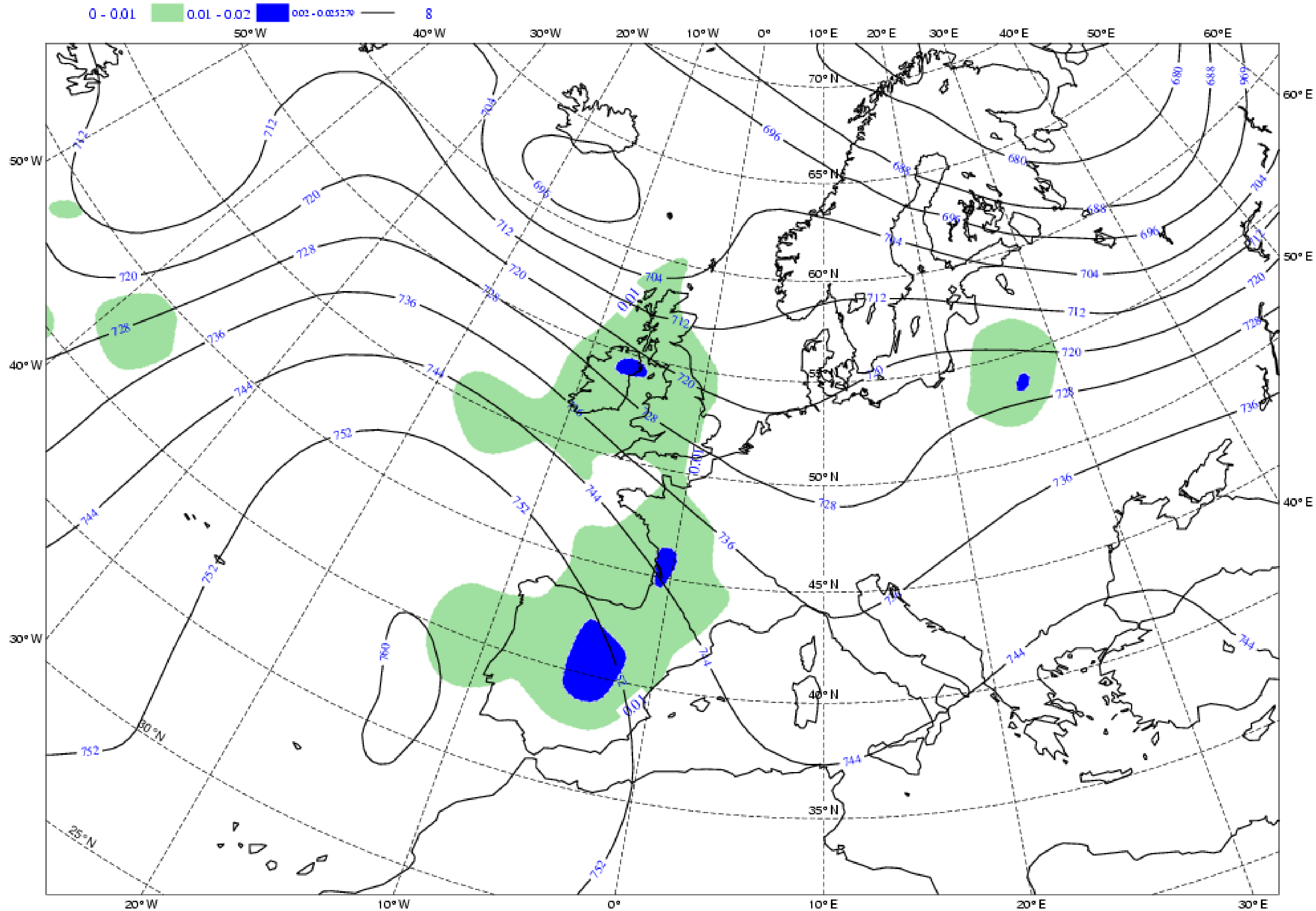
Version 1



Modified Stochastic Physics

Version 1

Multiplicative Inflaction Factor Difference (NoStochP-StochP): Model level 14



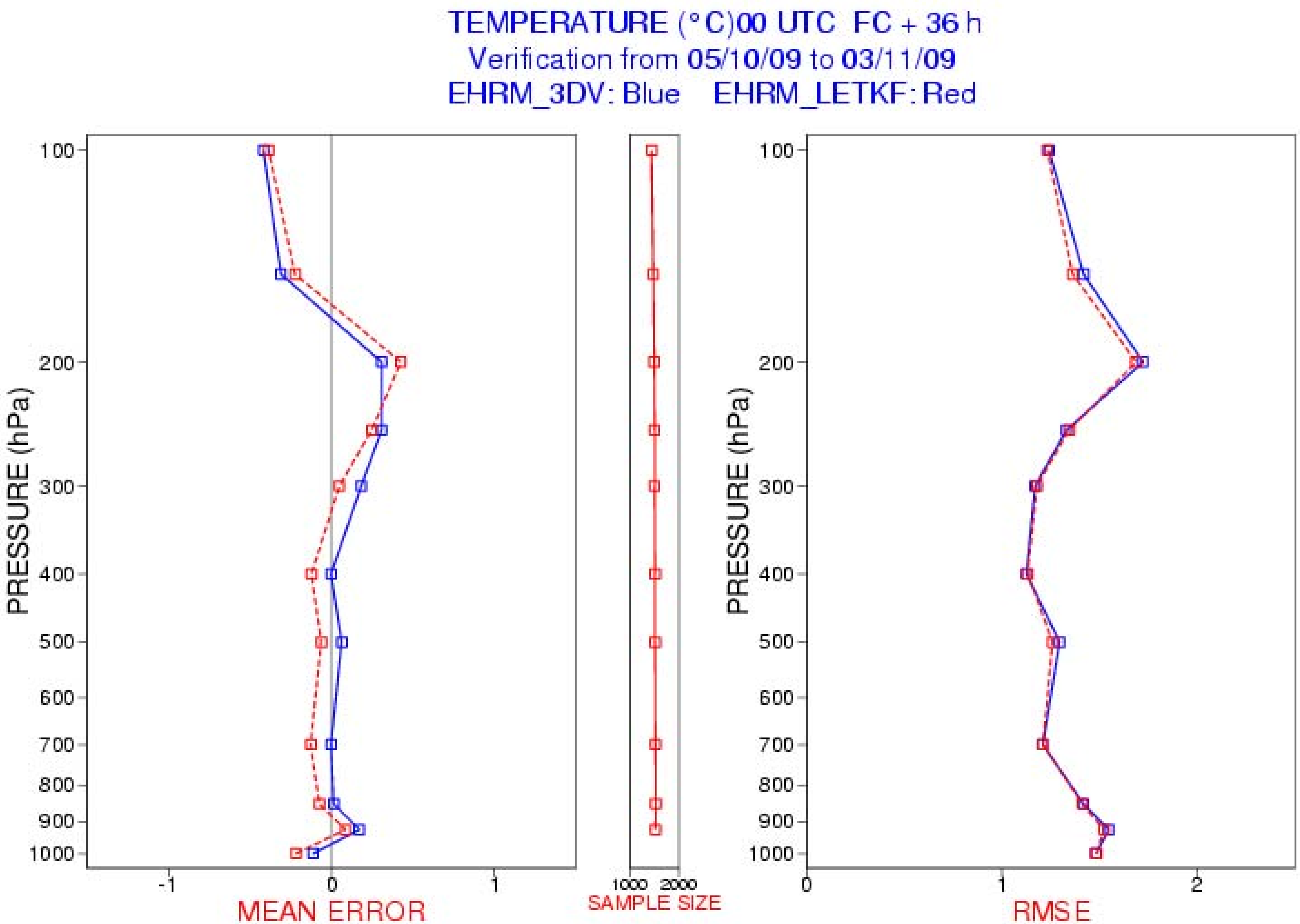
to 0.2.

Modified CNMCA LETKF Setup

- The stochastic physics scheme (version 1) was included in the HRM model and some experiments were performed to tune the length and the time scale of random number pattern. 10°/5°/2.5° and 1/3/6 hours were tested. A small improvement (especially in wind after T+24h) was found using the stochastic physics
- For a one month experiment 2.5° (5° was similar) and 1h were chosen (5 Oct-3Nov 2009) using 40 members.
- The additive inflaction factor was reduced (scal.f. 0.5 to 0.2)
- Scatterometer winds were also assimilated
- SST perturbations derived from randomly selected, scaled consecutive analysis differences were daily inserted
- A daily blending of the mean upper level analysis with the IFS one was also introduced to compensate the limited satellite data usage of the system at that levels

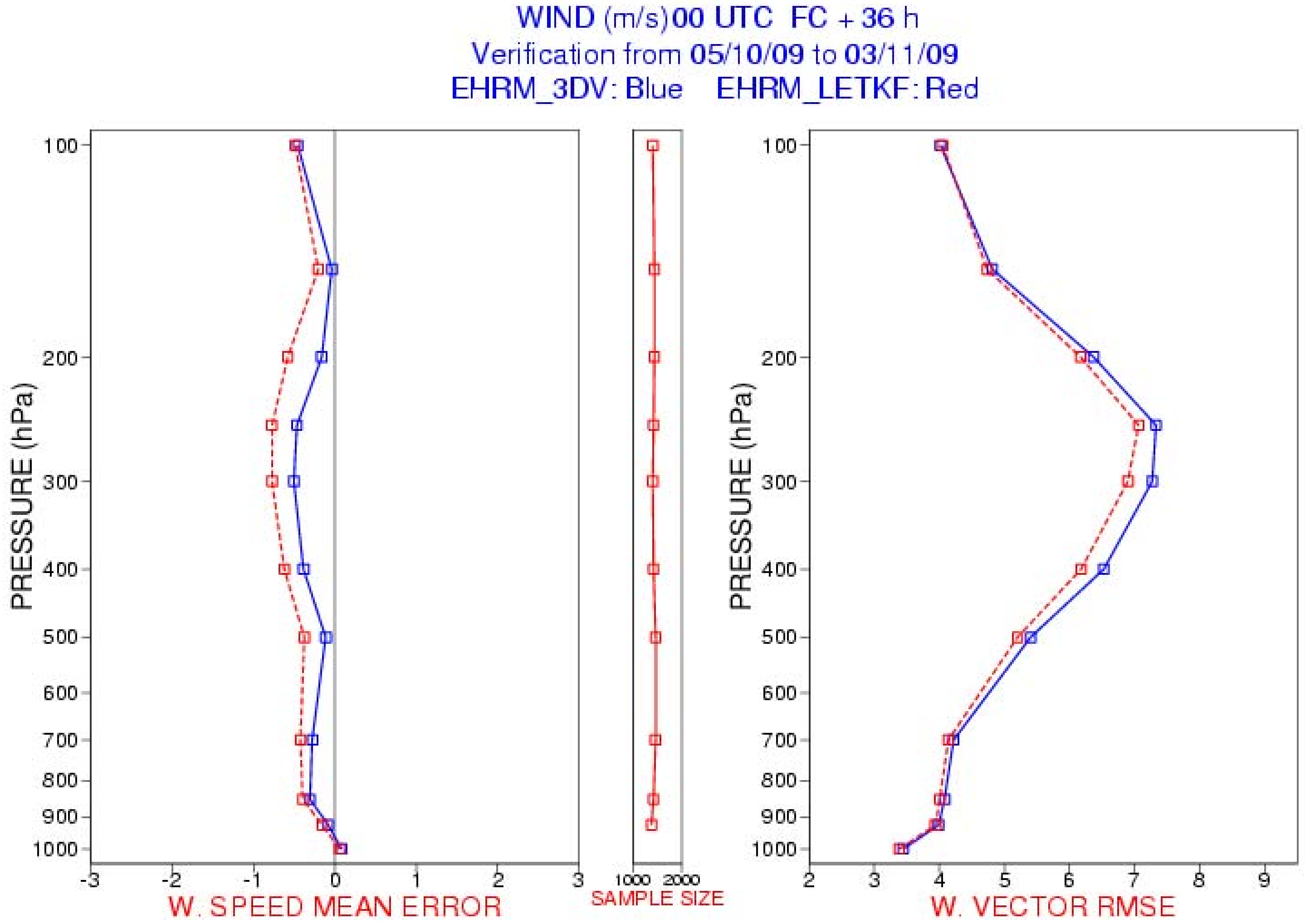
3D-VAR / LETKF Comparison

Objective verification against European radiounding stations



3D-VAR / LETKF Comparison

Objective verification against European radiounding stations



Modified Stochastic Physics

Version 2 (similar to Palmer et al, 2009)

- For all variables (u,v,T,qv), the random numbers r are drawn for each time step from a distribution close to a gaussian one with stdv=0.5 (bounded to the range 3 stdv)
- r are described with auto-regressive processes of first order (AR1) with a decorrelation time scale τ forced by gaussian random numbers η with zero mean and unit variance

$$\mathbf{r(t+Dt)=\Phi\, r(t) + \sigma\, \eta(t)}$$

where $\Phi=\exp(-Dt/\tau)$ and $\sigma=\mathrm{stdv}\,(1-\Phi^2)^{0.5}$

- The perturbations of T and qv are not applied if they lead to particular humidity values (exceeding the critical saturation value or negative)
- Spatial correlation is imposed using the same r in a whole column and drawing r for a coarse grid with spacing DL (similarly to the boxes of Buizza et al, 1999); then they are *bilinearly interpolated* on the finer grid to have a smooth pattern in space

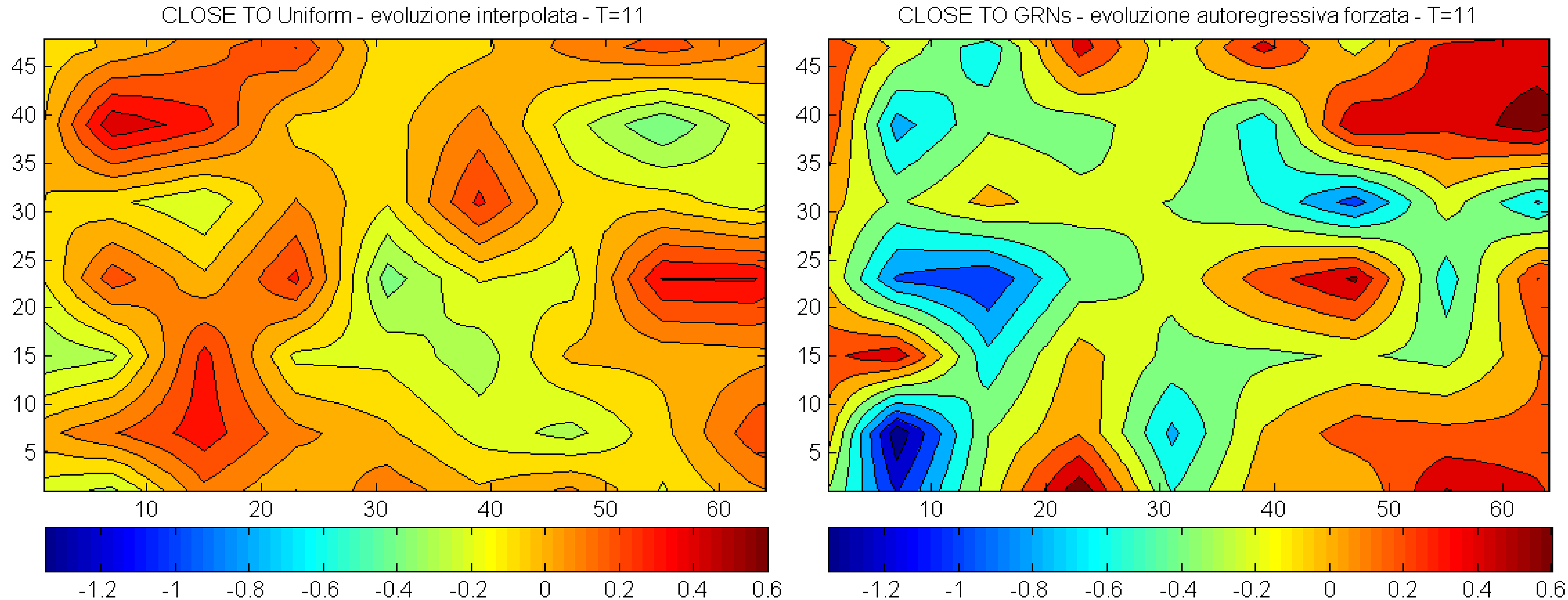
Modified Stochastic Physics

Version 2

Toy model and plots by A. Cheloni

2.5° coarse grid with bilin. interp.

Version 1: 1h coarse time grid with lin. interp. Version 2: AR1 with 1h decorr. length



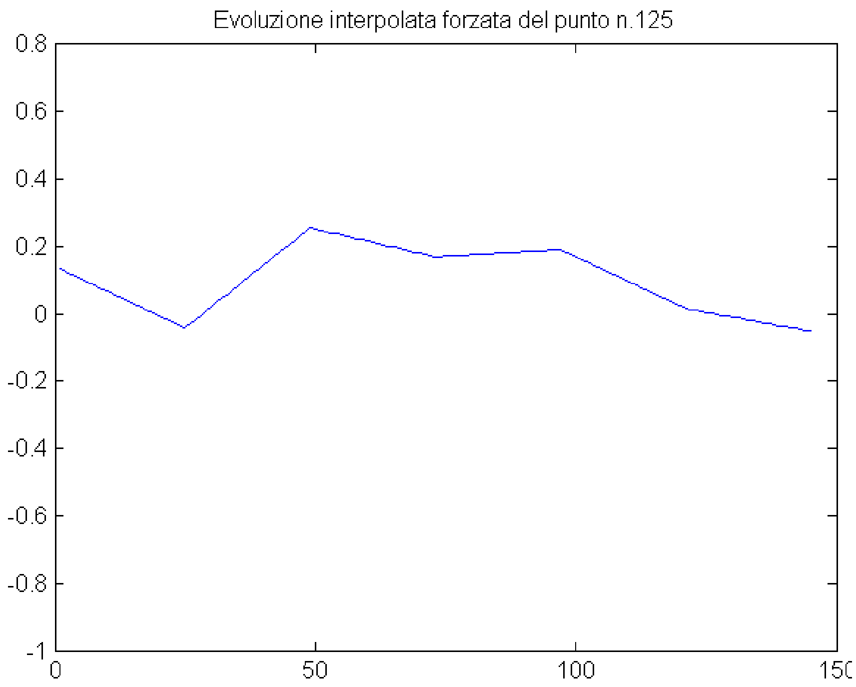
Version 2 can have larger perturbations than version 1

Modified Stochastic Physics

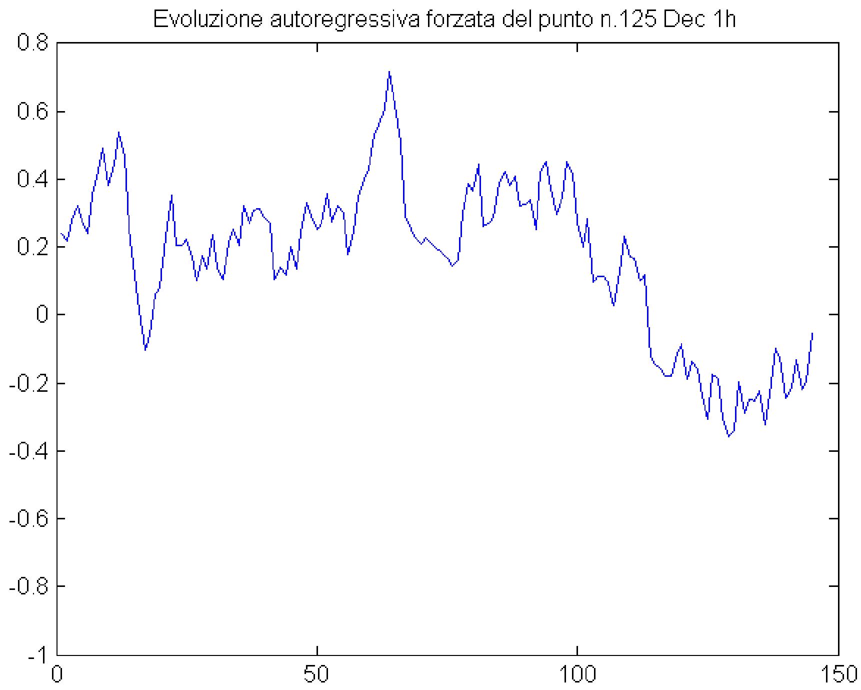
Version 2

Toy model and plots by A. Cheloni

1h coarse time grid with lin. interp.

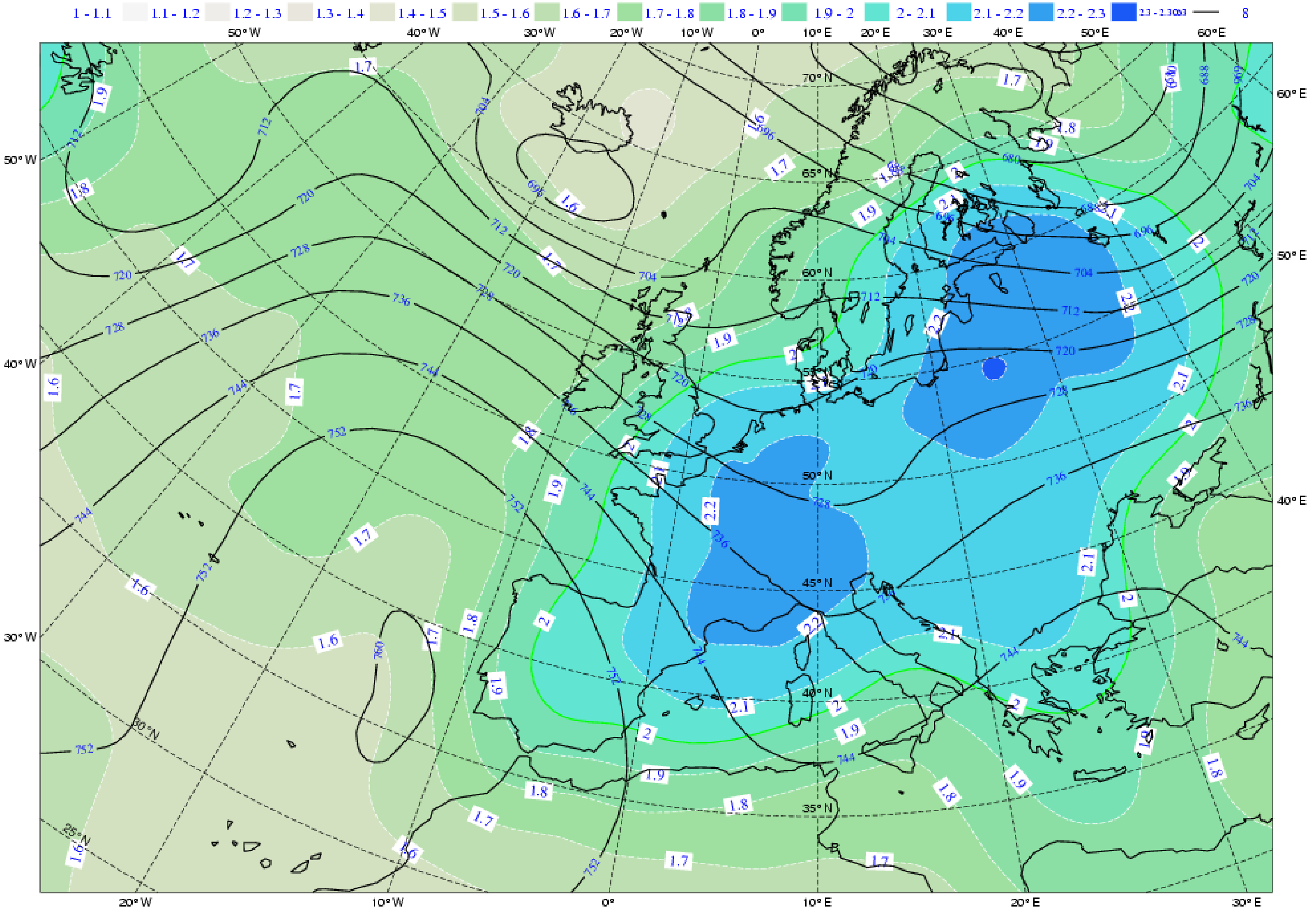


AR1 with 1h decorr. length



Modified Stochastic Physics

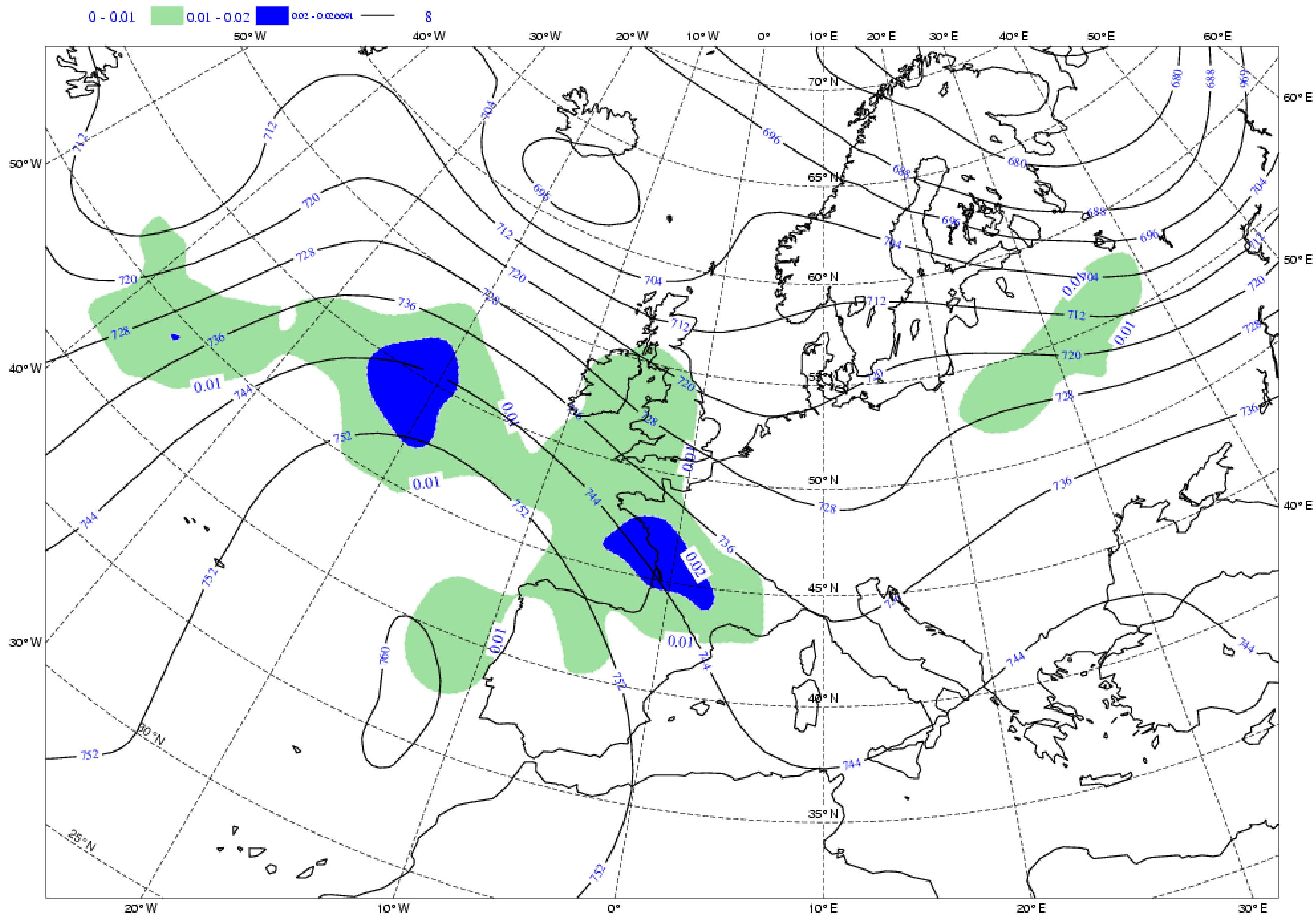
Version 2



Modified Stochastic Physics

Version 2

Multiplicative Inflaction Factor Difference (NoStochP-StochP): Model level 14



Future Developments

- Perturbation of lateral boundary conditions
- Further tuning of model error representation (stochastic physics version 2, bias correction)
- Treatment of non-linearities based on outer loop iterations
- Assimilation of AMSUA radiance observations
- Use of adaptive covariance localization

Thank you for your attention!