Ensemble Data Assimilation in Regional NWP

Massimo Bonavita
CNMCA
National Meteorological Service
Italy

Outline

- Data Assimilation at the Mesoscale
- The Ensemble Kalman Filter approach: pros & cons
- EnKF at CNMCA: Proof of concept
- EnKF at CNMCA: Results with all in-situ obs
- Lessons learned and outstanding issues

Data Assimilation at the Mesoscale

- Multiple dynamical scales (synoptic down to convective) are represented
- No static balance equation valid for all scales
- Smallest represented scales (convective) have very fast error growth (saturate ~1h) and loss of predictability, hence need for probabilistic prediction since early stages
- Common assumptions of gaussianity and linearity of errors and error growth may break down

Data Assimilation at the Mesoscale

- Observation systems with adequate spatial and temporal coverage typically provide indirect measures of model variables

- Observation systems which provide direct measurements of state variables typically lack adequate spatial and/or temporal coverage

- Model (System) error becomes a significant issue: how to treat it effectively still an open question

Ensemble Kalman Filter

- Monte Carlo implementations of Kalman Filter
- Use sample (from ensemble forecast) estimates of forecast error covariances in KF update eqs.
- Implicit (dynamical) balance relations
- Relax KF linearity assumption of forecast error covariances evolution (i.e., linear dynamics of forecast error)

Ensemble Kalman Filter

Issues in Mesoscale Data Assimilation:

Plus

- Use of dynamical, flow-dependent, balance:
    1. Avoid complex, explicit modeling of poorly known P^f;
    2. Better use of single level, sparse observations;
    3. Better use of observations with complex observation operators;
    4. Consistent dynamical update of non-observed state variables

- Relax KF linearity assumption of forecast error covariances evolution (i.e., linear dynamics of forecast error)

- Avoids the need of linearization of model & observation operators

Ensemble Kalman Filter

Issues in Mesoscale Data Assimilation:

Plus

- Provides the "best" possible initial ensemble for EPS forecasts, free!!

Ensemble Kalman Filter

Issues raised for Mesoscale Data Assimilation:

Minus

- As in KF only 1st and 2nd moments of state pdf are evolved => gaussian (or near gaussian) errors are assumed
- Linear relationship between observed and state variables over the range of forecast ensemble values
- Sensitive to model error

Ensemble Kalman Filter

Issues raised for Mesoscale Data Assimilation:

Minus

- Sensitivity to model error is particularly important, especially at very high resolution where highly-nonlinear processes are to be represented/parameterized (microphysics, turbulence, surface fluxes).
  However the same problem affects 4DVar!

Ensemble Kalman Filter

How does EnKF compare with 4DVar at the Meso-Convective Scale?

1. Difficult to perform a clean comparison in realistic settings, implementation issues overshadow fundamental results
2. Caya et al., 2005, performed one comparison in perfect model conditions (2 Km grid spacing, synthetic radar obs of radial winds and reflectivity, every 5’).

The performances were similar:

- 4DVar shows an advantage for early assimilation cycles, then EnKF gets better due to cycling covariances;
- EnKF sensitive to initial ensemble specification (should not be problematic in DA cycling)

Ensemble Kalman Filter

Different flavours of EnKF:

1. Ensemble of analyses: each member assimilates perturbed obs with direct solution of obs-space analysis eqn. (Stochastic EnKF) (Houtekamer and Mitchell, Keppene)

2. Serial, or one-observation-at-a-time, assimilation (EnSRF, EAKF, “Square Root EnKF”)

3. Local analyses for each grid column, obs selection (LETKF, “Square Root EnKF”)

EnKF at CNMCA: proof of concept

- **LETKF** (Hunt et al, 2007) approach chosen because:

    1. Algorithmically simple to code;
    2. Proven on various systems of increasing complexity and realism;
    3. Intrinsically parallel, very appropriate for current cluster computing systems;
    4. Avoids serial processing of observations (allows taking into account correlated observation errors inside local patches)

EnKF at CNMCA: proof of concept

- **LETKF** (Hunt et al, 2007) approach chosen because:

5. Same methodology for global-regional-convective scale!

For all these reasons LETKF has been chosen at DWD for the global model, at CNMCA for regional DA, and in the COSMO framework as the (tentative) basis of the next generation DA system:

KENDA project

EnKF at CNMCA: proof of concept

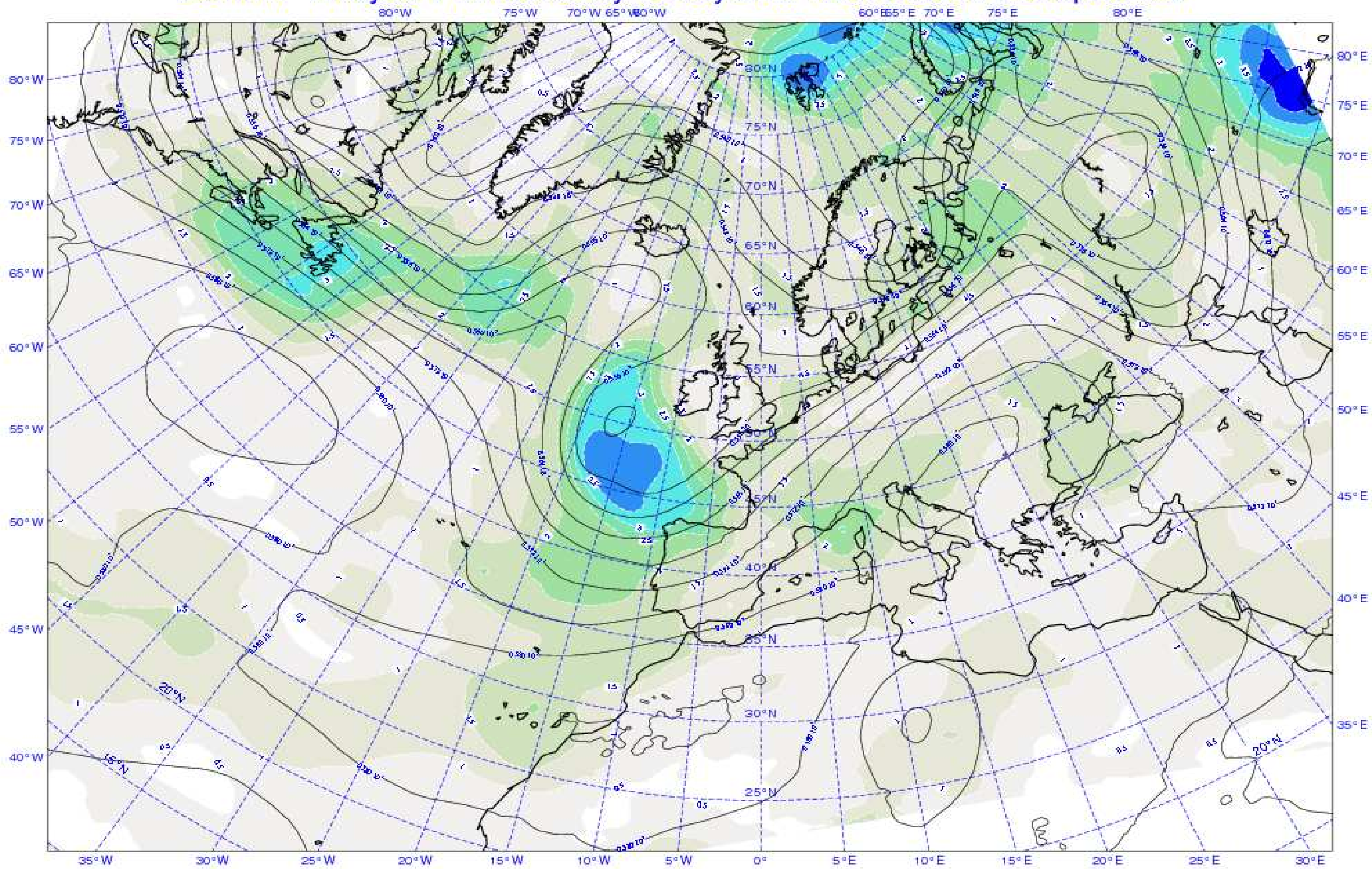
- **Preliminary results from LETKF,** CNMCA implementation (*Bonavita, Torrisi and Marcucci*, 2008,*QJRMS*)

EnKF at CNMCA: proof of concept

CNMCA Implementation

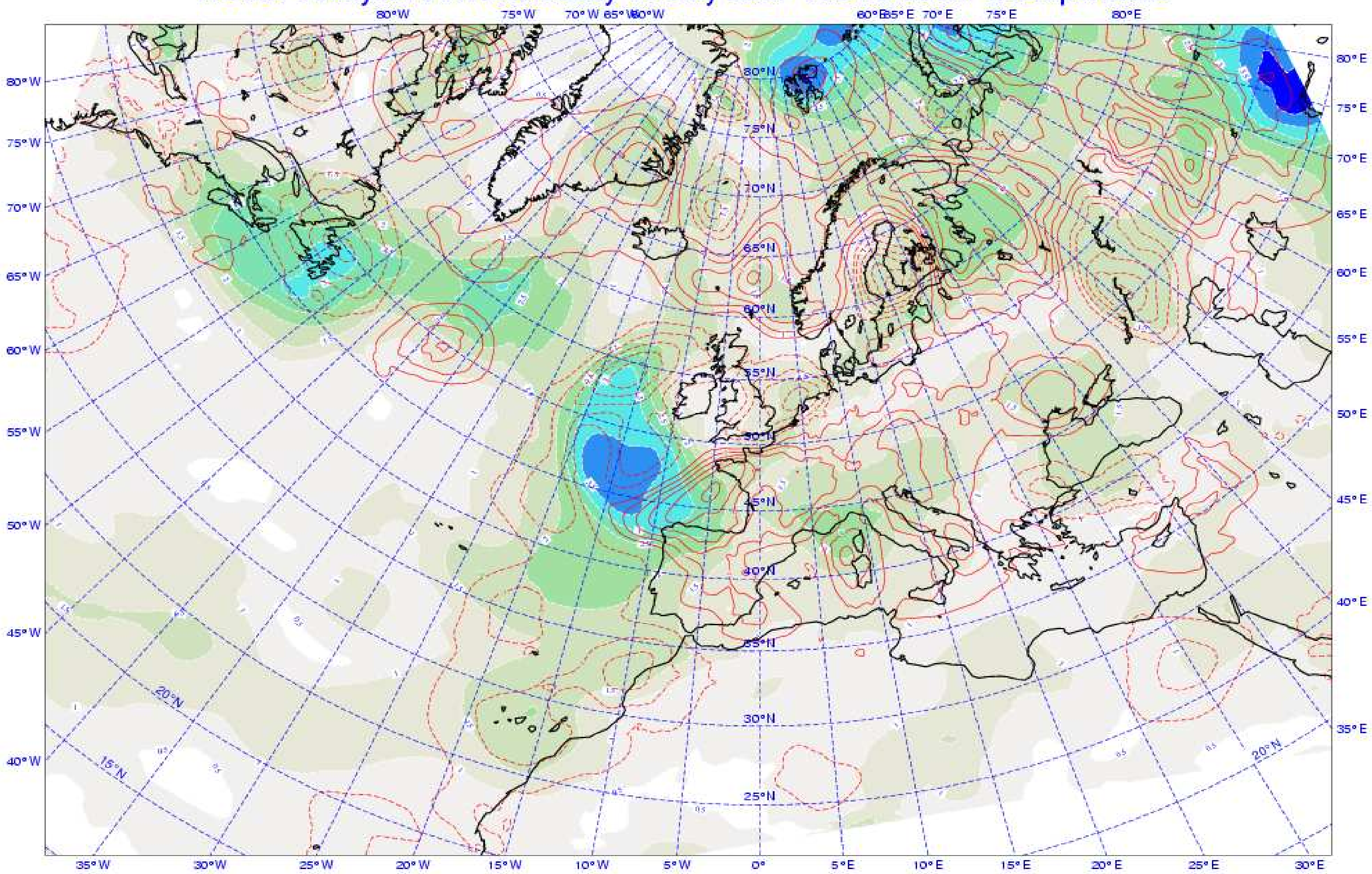
- 30 member ensemble at 0.25°(~ 28Km) grid spacing, 30 vertical levels (top at 10 hPa)
- 6-hourly assimilation cycle run for 15 days
- (T,u,v,P_s) set of control variables
- Operational 3DVar cycle run in parallel at same spatial resolution
- Observations: RAOB (Tuv), SYNOP(SP), SHIP(SP), BUOY(SP)
- 700 Km circular local patches
- Multiplicative adaptive covariance inflation, pressure dependent

ROME Analysis VT:Wednesday 18 July 2007 18UTC 500hPa temperature
ECMWF Analysis VT:Wednesday 18 July 2007 18UTC 500hPa Geopotential

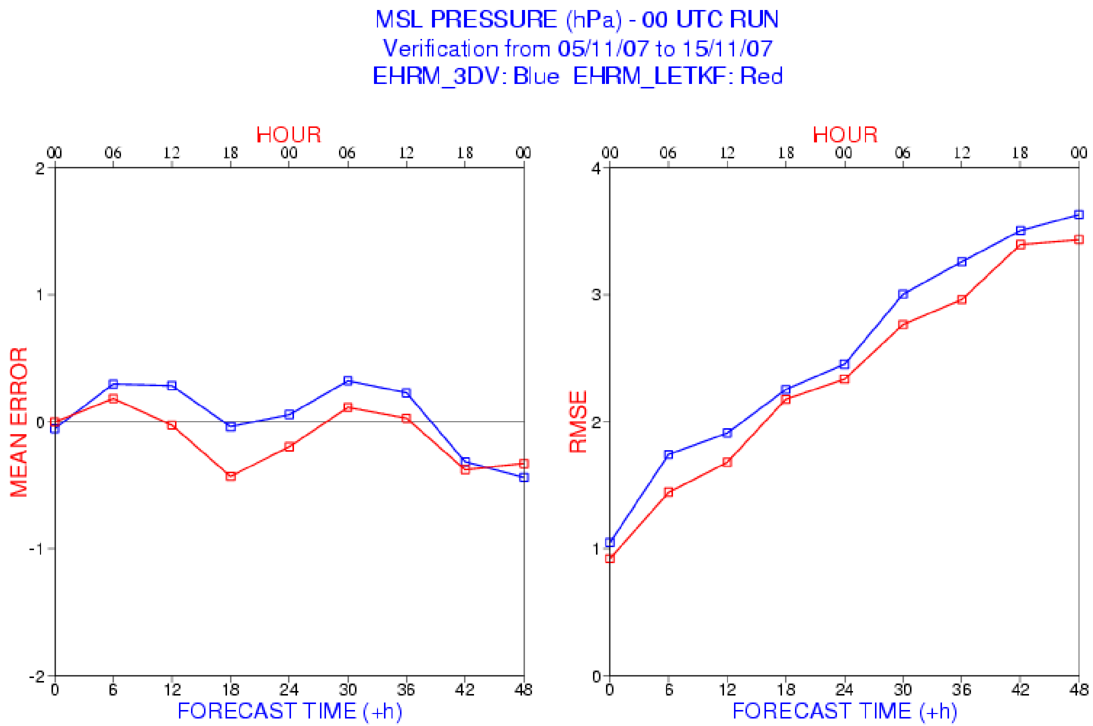


ROME Analysis VT:Wednesday 18 July 2007 18UTC 500hPa temperature

ROME Analysis VT:Wednesday 18 July 2007 18UTC 500hPa temperature



EnKF at CNMCA: proof of concept



EnKF at CNMCA: results with in-situ obs

- Results with reduced obs dataset have been found good enough to proceed to more realistic settings
- Tuning of filter parameters was also necessary

=>

- New set of experiments (in collaboration with DWD)
- Same configuration as previous experiments, but with all available in-situ observations
- Observations only at analysis time (i.e., simple LETKF, not 4D-LETKF yet)
- Best “average” observation selection radius was found to be L_{patch}=900Km

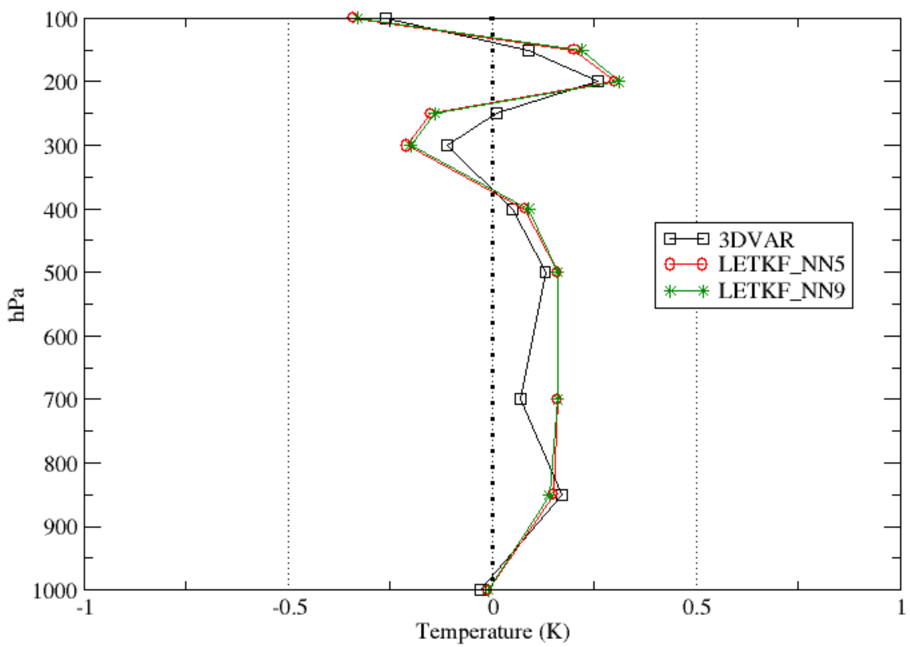
EnKF at CNMCA: results with in-situ obs

- Observation selection radius (L_{patch}) was made spatially dependent in order to reflect local observation density: this resulted in equal or marginally better scores and much better computational load balancing

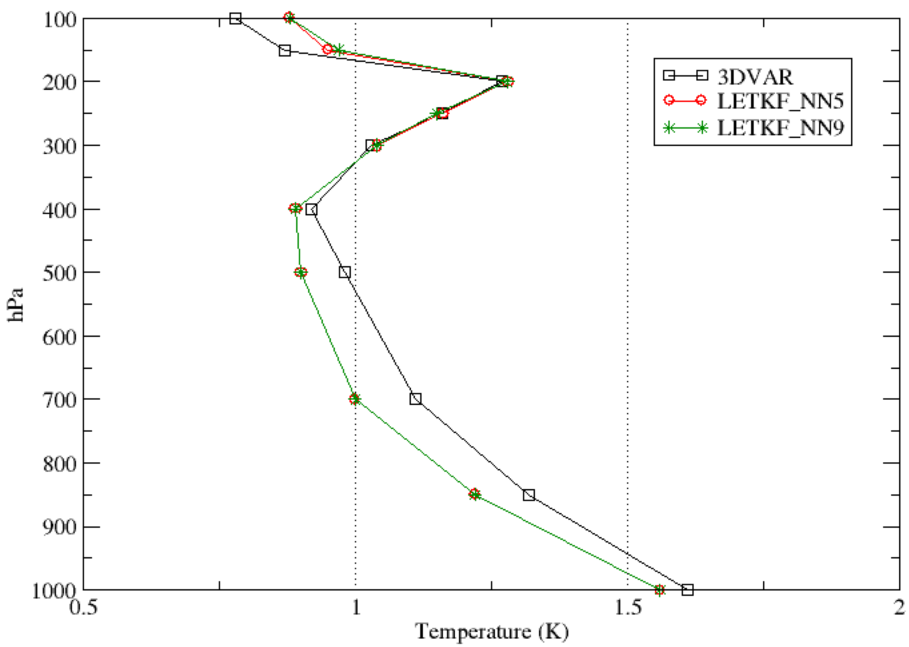
EnKF at CNMCA: results with in-situ obs

Temperature t+24h forecasts, verification vs ECMWF ana.

BIAS



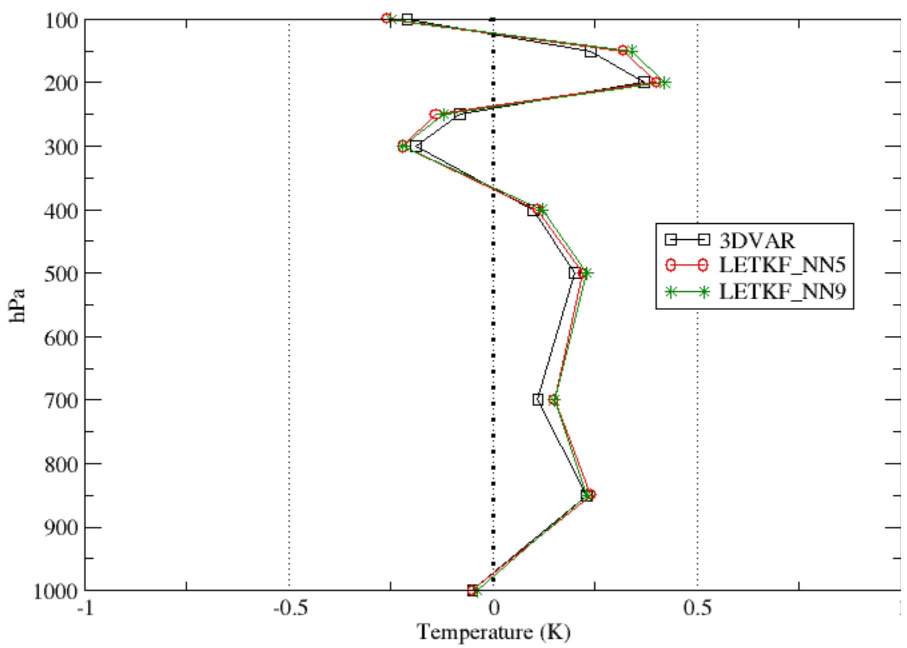
RMSD



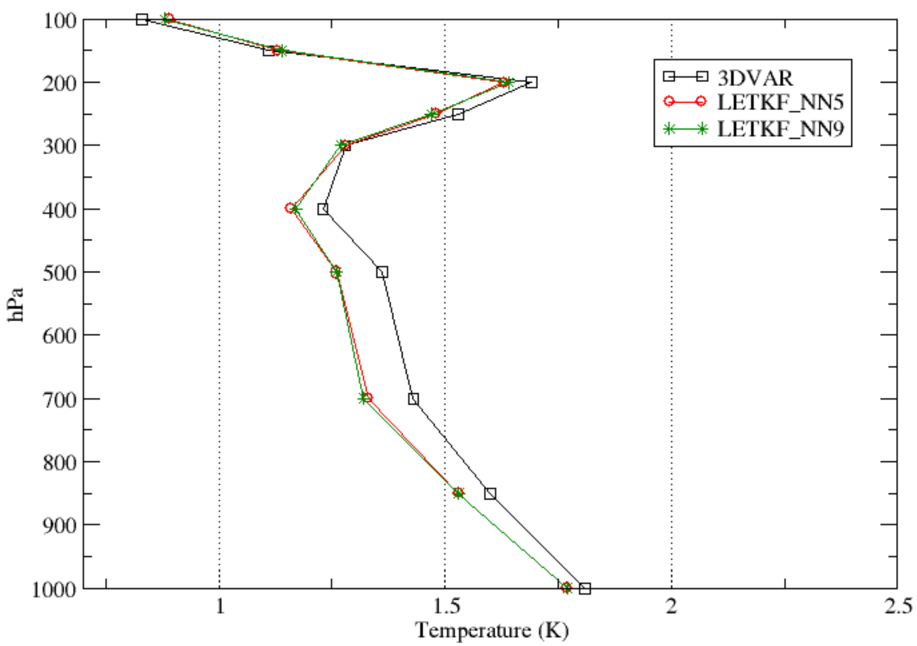
EnKF at CNMCA: results with in-situ obs

Temperature t+48h forecasts, verification vs ECMWF ana.

BIAS



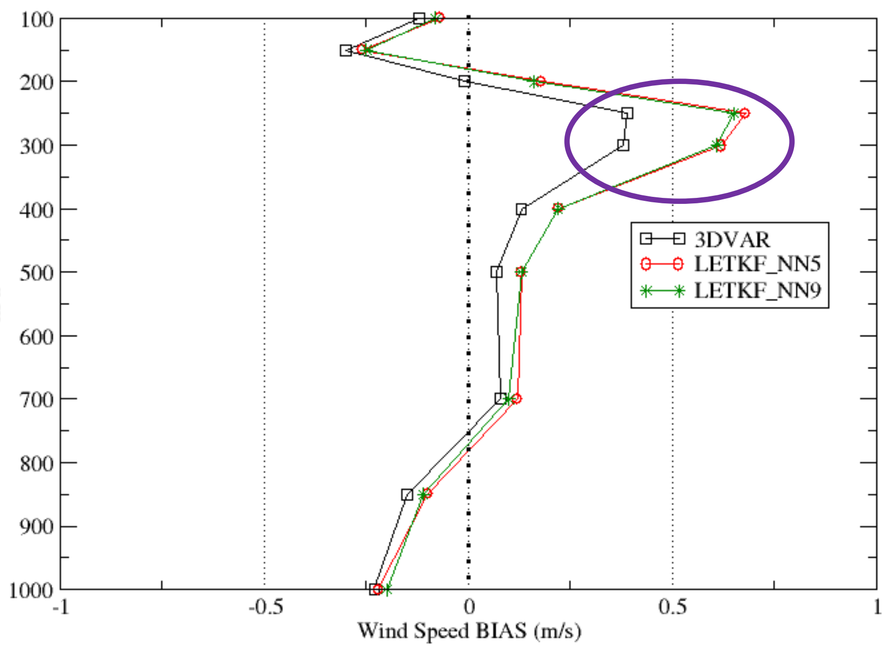
RMSD



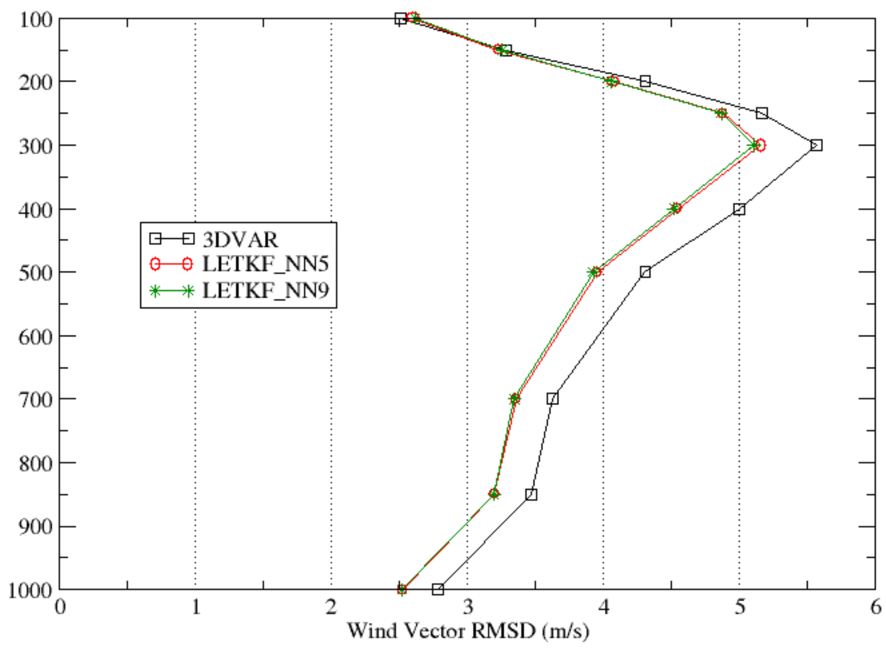
EnKF at CNMCA: results with in-situ obs

Wind t+24h forecasts, verification vs ECMWF ana.

Wind Speed BIAS

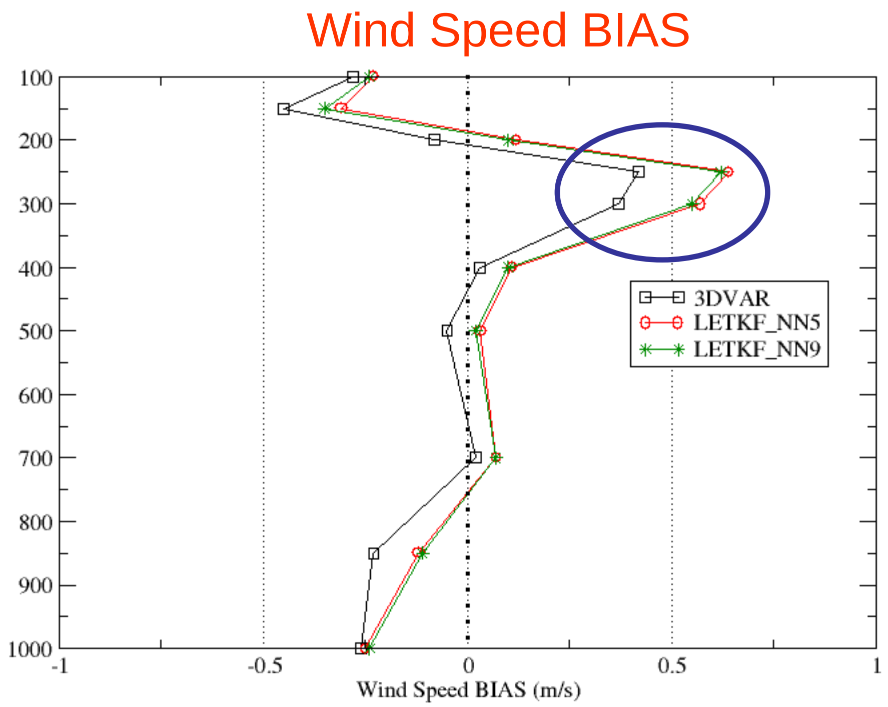


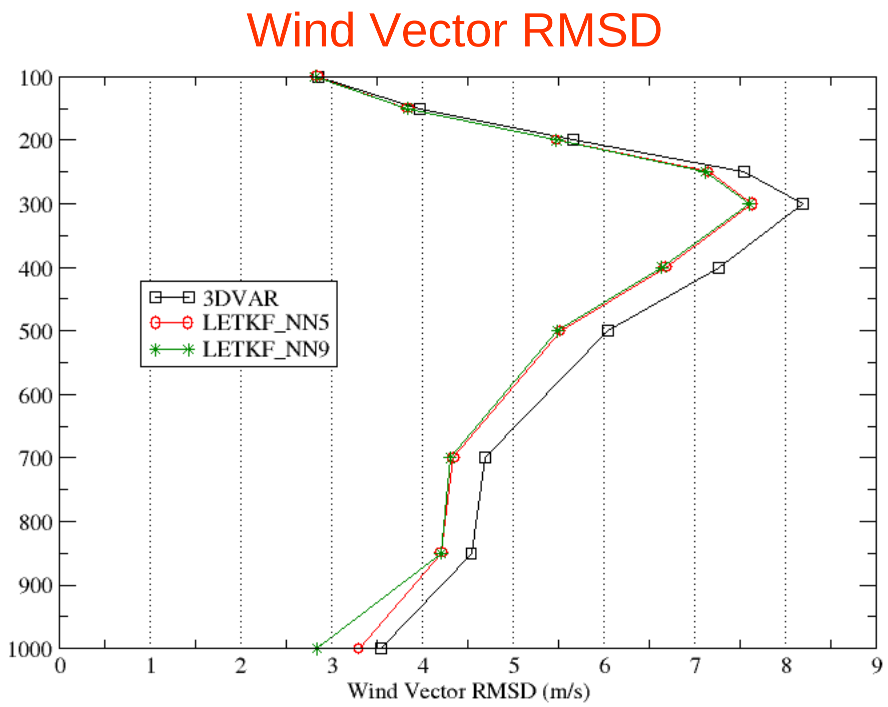
Wind Vector RMSD



EnKF at CNMCA: results with in-situ obs

Wind t+48h forecasts, verification vs ECMWF ana.





EnKF at CNMCA: results with in-situ obs

- With all in-situ obs LETKF confirms advantage over 3DVar in terms of RMSE metric
- With all in-situ obs LETKF confirms more sensitivity to model systematic errors
- Multiplicative adaptive covariance inflation seems adequate to combat filter divergence symptoms and provide a reliable first-guess ensemble

EnKF at CNMCA: Lessons learned

- EnKF full potential has not been realized yet in our implementation:
  - ❑ Improve Covariance Inflation (fully 3D model, additive model, stochastic perturb.);
  - ❑ Filtering of forecast covariances to reduce spurious correlations (reduce sampling errors);
  - ✓ Inclusion of humidity in control variables set;
  - ❑ Use of all obs over the assimilation window (4D-LETKF);
  - ❑ Use of radiances

EnKF at CNMCA: Outstanding problems

- Forecasts based on EnKF analysis consistently show equal or larger systematic errors than 3DVar initialized forecasts
- This suggests that EnKF is more sensitive to model errors than 3DVar: this is expected since EnKF analysis is linear combination of forecast ensemble => ensemble spread only represents growth of initial condition errors (i.e., it is blind to model errors)
- But in Extended KF formulation:

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

EnKF at CNMCA: Outstanding problems

- What can we do to treat model error?
  1. Wait for better models!

  … in the meantime:

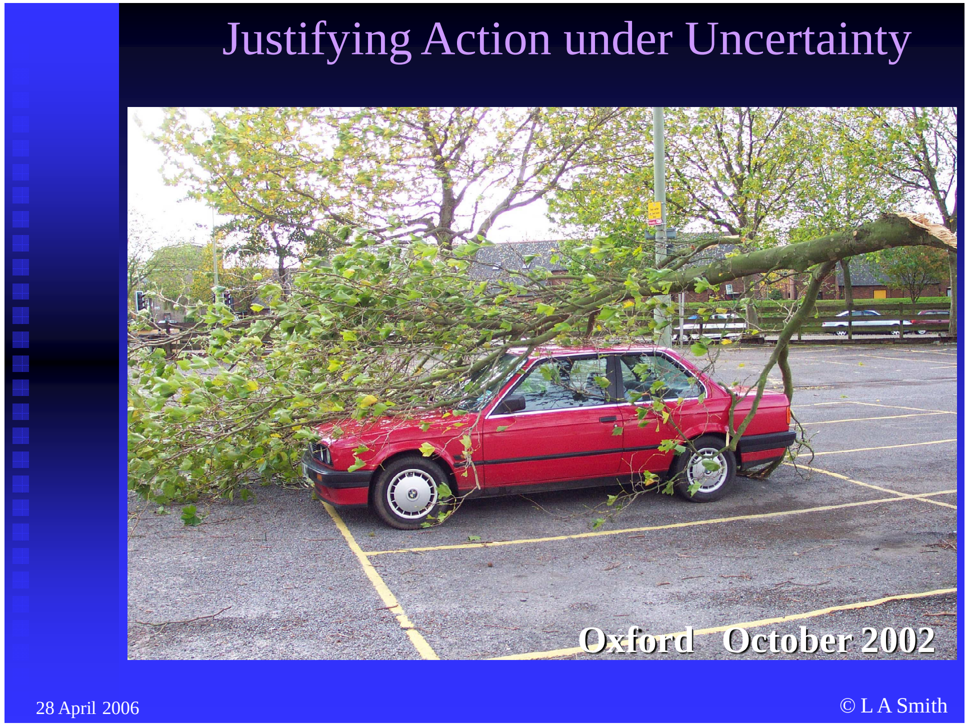
EnKF at CNMCA: Outstanding problems

In low order models (SPEEDY Model, *Molteni*, 2003) good results have been obtained with the use of additive covariance inflation + low-dim method to correct for large scale, slowly evolving model bias (*Kalnay*, 2008)

Similar method (bias correction after *Dee & DaSilva*, 1998, + additive noise + stochastic physics) has been employed in state of the art oceanic model data assimilation (GMAO Ocean EnKF, *Keppenne et al*., 2008)

…

Model Error…



From "Ensemble Prediction Systems and
Real Life Decision Support", Leonard Smith, 2006

Suggestions & Questions