Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Preliminary experiments with LETKF

H. Reich, A. Rhodin, C. Schraff

COSMO General Meeting 2010, Moscow

6-10. September 2010

LETKF basics

- Implementation following *Hunt et al., 2007*
- basic idea: do analysis in the space of the ensemble perturbations
  - computational efficient, but also restricts corrections to subspace spanned by the ensemble
  - explicit localization (doing separate analysis at every grid point, select only certain obs)
  - analysis ensemble members are locally linear combination of first guess ensemble members

LETKF experiments

- technical implementation:
  - stand-alone LETKF script environment to run COSMO-DE LETKF + diagnostics / plotting
  - toy model (Lorenz-96,40 grid points) to test LETKF components
- very preliminary experiments with successive LETKF assimilation cycles (32 ensemble members, drawn from 3dVar B-Matrix)
  - use obs from GME NetCDF feedback files (sparse density)
  - 3-hourly cycles, up to 2 days (7-8 Aug. 2009: quiet + convective day)
  - lateral boundary conditions (LBC) from COSMO-SREPS (3 * 4 members) or deterministic LBC

LETKF experiments

- results are *preliminary* because of:
  - 3h update (later $\approx$ 15 min)
  - sparse observation density
- $\rightarrow$ concentrate on general topics:
  - rms/spread of ensemble
  - noise (dps/dt and $wa500$)
  - general behaviour of LETKF (analysis increments etc.)
- test effect of parameter variation, but no fine tuning
- some (adaptive) methods to increase spread/reduce noise have been tested with toy model/LETKF

LETKF experiments

- analysed variables are $u, v, w, T, pp, qv, qcl, qci$
- analysed means that linear combination is applied to these variables (other variables taken from first guess ensemble / ensemble mean)
- verify LETKF *mean* against
  - nudging analysis (u, v, T, pp, qv)
  - observations (u, v, T)
- nudging uses much more observations
- ensemble mean is used for verification
- verification tool (deterministic/ensemble scores) is currently under development

spread (ens BC)

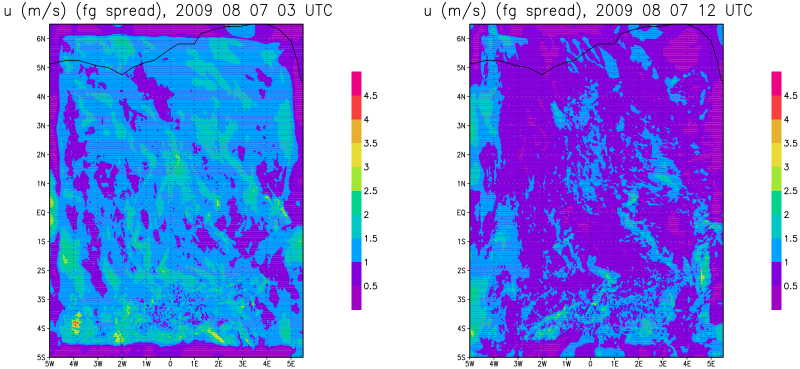


Fig.1: spread (wind component u in m/s) of first guess on 7 Aug. 2009 at 03 UTC (after 1 LETKF analysis with 3DVAR-B) (left) and at 12 UTC (after 4 analysis cycles) (right)

The large scale spread decreases and “new” spread comes in from the west due to the lateral boundary fields.

spread (det BC)

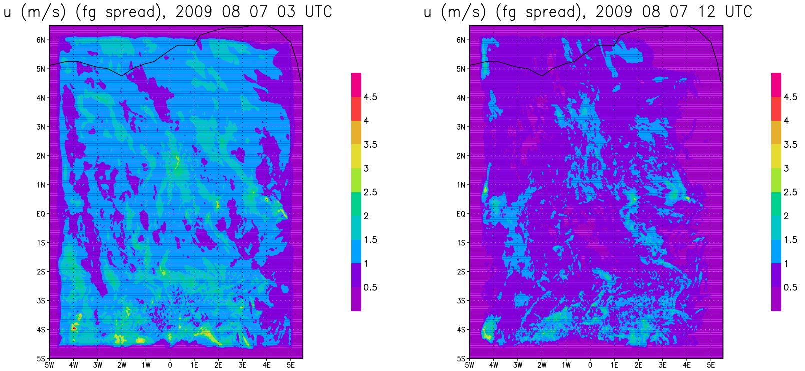


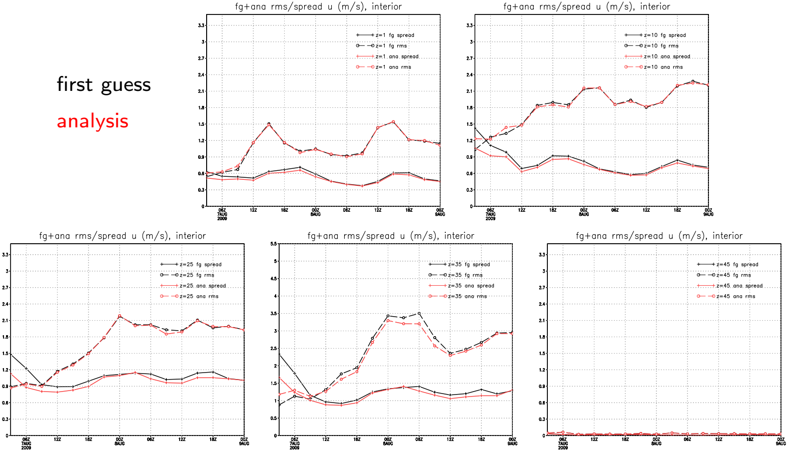
Fig.2: same as Fig.1 but with *deterministic* boundary conditions

The large scale spread decreases faster as no "new" spread comes in from the lateral boundary fields.

u rms/spread (interior), det BC

first guess

analysis



scale differs

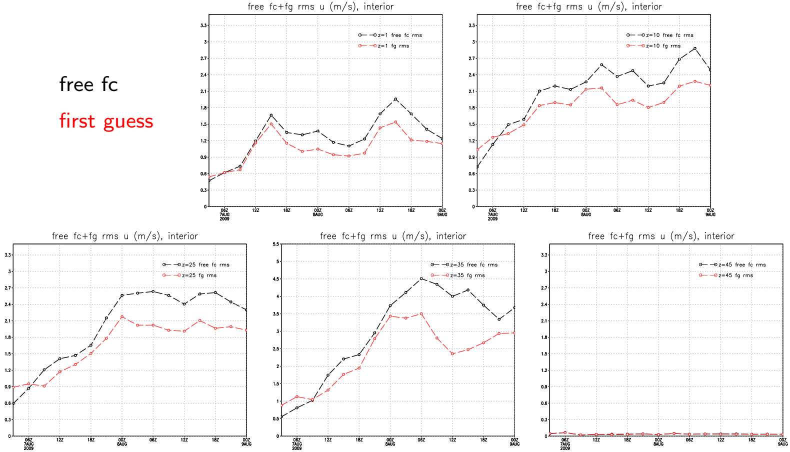
Fig.3: rms/spread of u,(m/s) (interior) of first guess and analysis; results for det BC and active vertical localization

relatively small differences between rms of first guess/analysis

u rms (interior), det BC (comparison with free forecast)

free fc

first guess

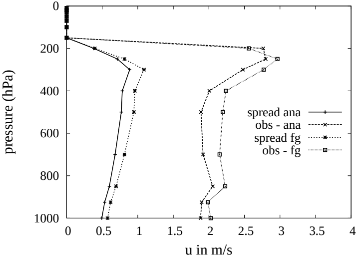


scale differs

Fig.4: rms of u,(m/s) (interior) of first guess and free forecast; results for det BC and active vertical localization

LETKF performs better than free forecast

u obs-fg/spread (time average, whole area), det BC (exp 1004)



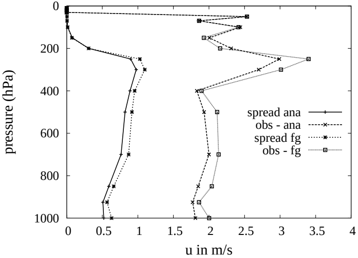


Fig.5: time average (20090807 15 UTC - 20090809 00 UTC)of obs-fg and spread of u,(m/s) (whole area), AIREP (left) and TEMP (right); results for det BC and active vertical localization (exp1004)

larger differences between analysis and first guess at observation locations, but LETKF is underdispersive

pp rms/spread (interior), det BC

first guess

analysis

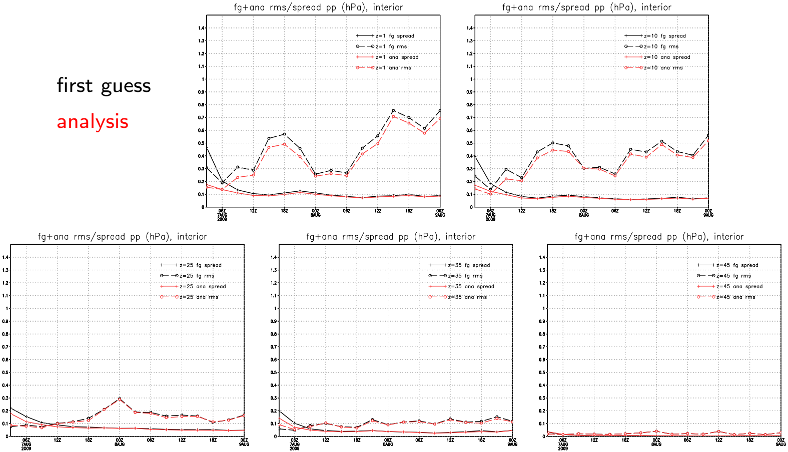


Fig.6: rms/spread of *pp*,(hPa) (interior) of first guess and analysis; results for det BC and active vertical localization

larger differences between rms of first guess/analysis at lower levels, *pp* is strongly influenced by BC

pp rms (interior), det BC (comparison with free forecast)

free fc

first guess

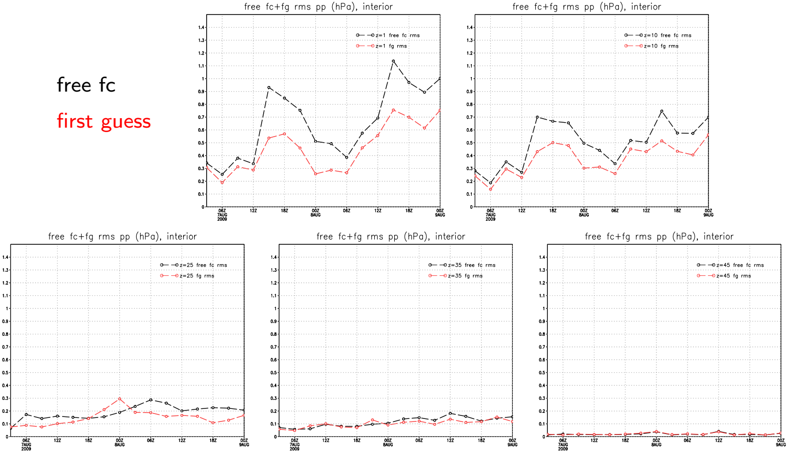


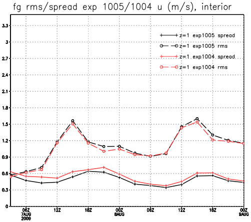
Fig.7: rms of *pp*,(hPa) (interior) of first guess and free forecast; results for det BC and active vertical localization

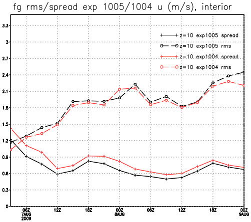
LETKF performs better than free forecast esp. at lower levels; free forecast uses "perfect" boundary conditions (same as nudging)

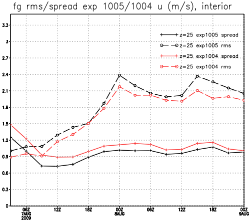
u rms/spread (interior), det BC (effect of vertical localization)

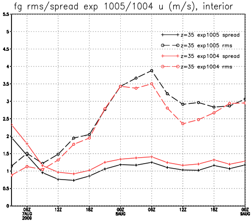
no vert loc

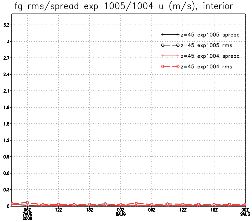
vert loc









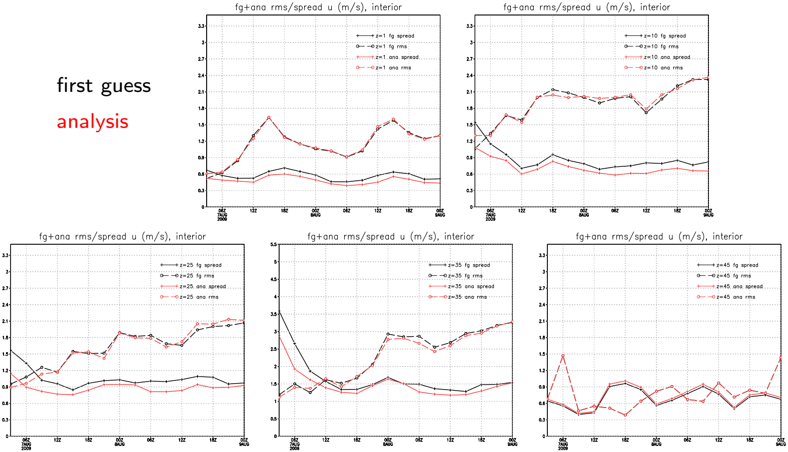


scale differs

Fig.8: intercomparison of first guess rms and spread of u,(m/s) (interior); results for det BC and active vertical localization (exp1004) and no vertical localization (exp1005)

turning off the vertical localization increases rms and decreases spread

u rms/spread (interior), ens BC



scale differs

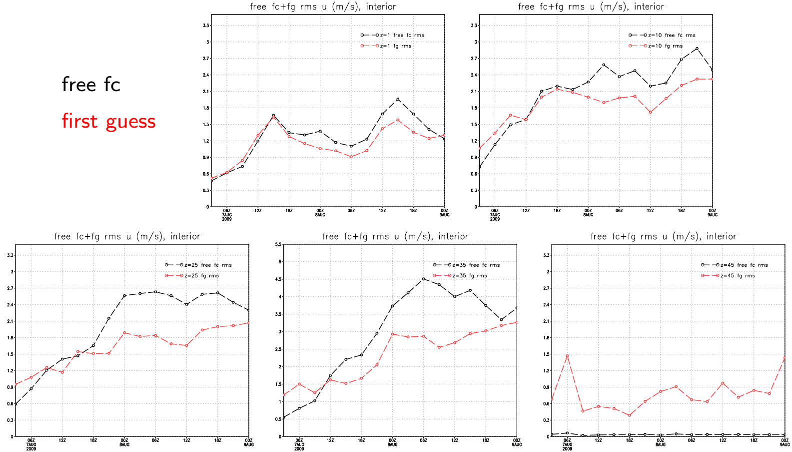
Fig.9: rms/spread of u,(m/s) (interior) of first guess and analysis; results for ens BC and active vertical localization

larger differences between rms of first guess/analysis, but sometimes rms of analysis is larger; at higher levels fields are completely determined by (upper) boundary conditions

u rms (interior), ens BC (comparison with free forecast)

free fc

first guess



scale differs

Fig.10: rms of u,(m/s) (interior) of first guess and free forecast; results for ens BC and active vertical localization

LETKF performs better than free forecast

u obs-fg/spread (time average, whole area), det BC (exp 1004)

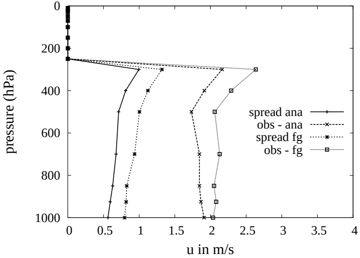


Fig.11: time average (20090807 15 UTC - 20090809 00 UTC)of obs-fg and spread of *u*,(m/s) (whole area), AIREP (left) and TEMP (right); results for det BC and active vertical localization (exp1004)

larger differences between analysis and first guess at observation locations

pp rms/spread (interior), ens BC

first guess

analysis

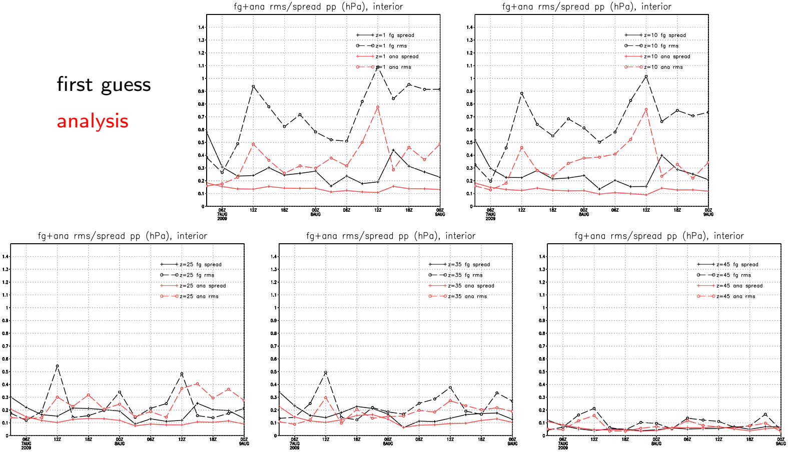


Fig.12: rms/spread of *pp*,(hPa) (interior) of first guess and analysis; results for ens BC and active vertical localization

much larger differences for pressure deviation *pp* than for *u*.

adaptive methods

- lack of spread is (partly) due to model error which is not accounted for so far
- one (simple) method to increase spread is multiplicative covariance inflation:
  - $X_b \rightarrow \rho X_b$ with $\rho > 1$
  - more advanced methods to account for model error (esp. in limited-area models) need to be developed
- problem: tuning inflation factor ρ takes much time, adaptive procedure preferable
  - (Kalnay et al.:) online estimation of inflation factor
  - compare "observed" (obs - f.g.) : $(y - H(x))$
    with "predicted" (obs - f.g.) : $(R + HP_bH^t)$

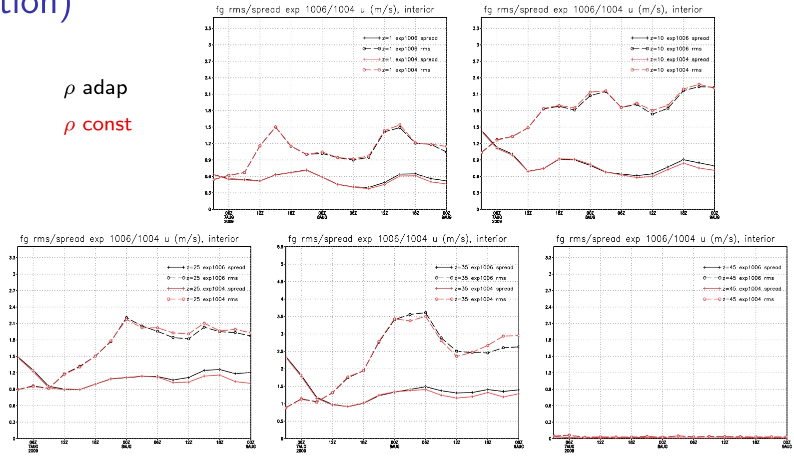
adaptive methods

- obs errors / R-matrix probably assumed incorrectly, correction needed
  - compare observed obs covariance with assumed one (in ensemble space) and correct R automatically if necessary
- both methods (est. of inflation factor / R matrix) have been tested with reasonable numerical cost and success with the toy model, and have been implemented in the COSMO LETKF
- For COSMO LETKF these methods are currently tested
- For determistic BC, positive effect visible; for ens BC more sophisticated approach necessary

u rms/spread (interior), det BC (effect of adaptive cov. inflation)

ρ adap

ρ const



scale differs

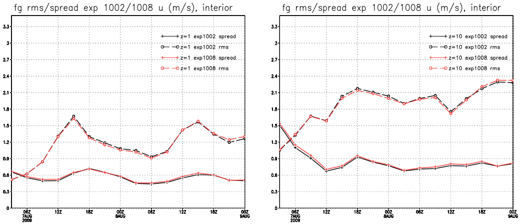
Fig.13: intercomparison of first guess rms and spread of u,(m/s) (interior); results for det BC and constant inflation factor ρ (exp1004) and adaptive covariance inflation (exp1006)

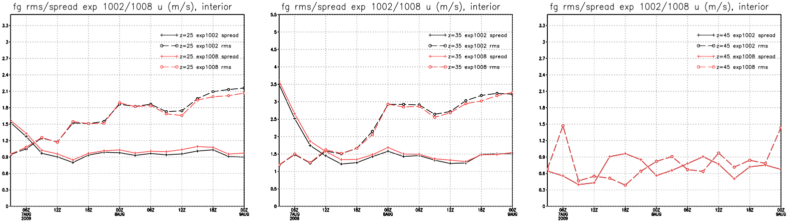
adaptive covariance inflation increases spread and decreases rms

u rms/spread (interior), ens BC (effect of adaptive cov. inflation)

ρ adap

ρ const





scale differs

Fig.14: intercomparison of first guess rms and spread of u,(m/s) (interior); results for ens BC and constant inflation factor ρ (exp1008) and adaptive covariance inflation (exp1002)

no better results with ens BC

effect of vertical localization on noise

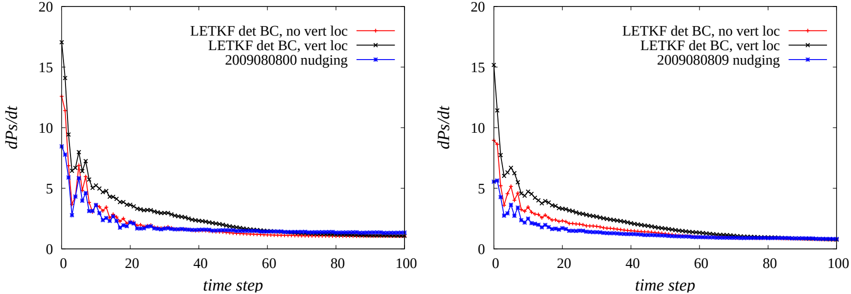
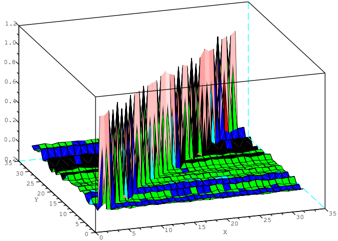


Fig.15: noise (dPs/dt in Pa/s, area mean) of one ensemble member at 20090808 00 UTC (left) and 09 UTC (right) for det BC and vertical localization switched on/off

Noise decreases for vertical localization switched off

weight matrices



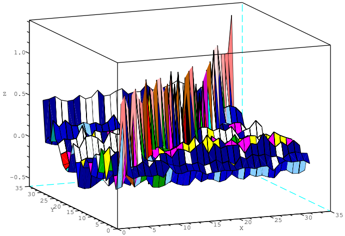


Fig.16: weight matrices (the matrix the first guess ensemble is multiplied with, for a case with "normal" number of observations (left) and with many observations (or small obs. errors; right).

off diagonal elements even for large number of obs ≤ 0.5 and diagonal elements > 0.5

hydrostatic balancing

- diagonal elements of weight matrix are larger than off diagonal elements
- $\rightarrow$ analysis ensemble k gets largest contribution from first guess ensemble member k plus (smaller) corrections from members $i \neq k$
- thus, the difference between analysis and first guess ensemble member k (the analysis increment) is small compared to the full fields
- apply hydrostatic balancing to this increment; this leaves the full fields nonhydrostatic as it should be in a nonhydrostatic model

effect of hydrostatic balancing on noise

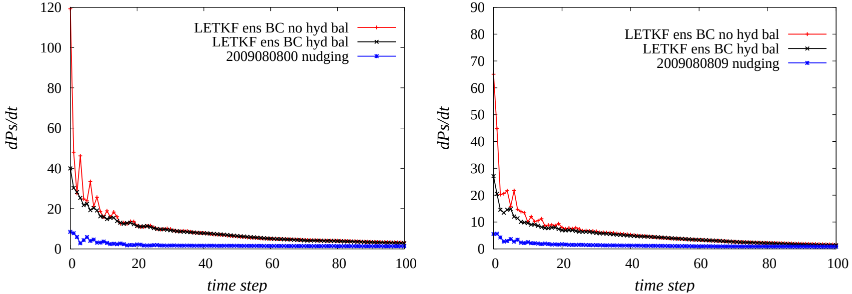


Fig.17: noise (dPs/dt in Pa/s, area mean) of one ensemble member at 20090808 00 UTC (left) and 09 UTC (right) for ens BC with hydrostatic balancing switched on/off

Noise is reduced by applying hydrostatic balancing

noise: area plots

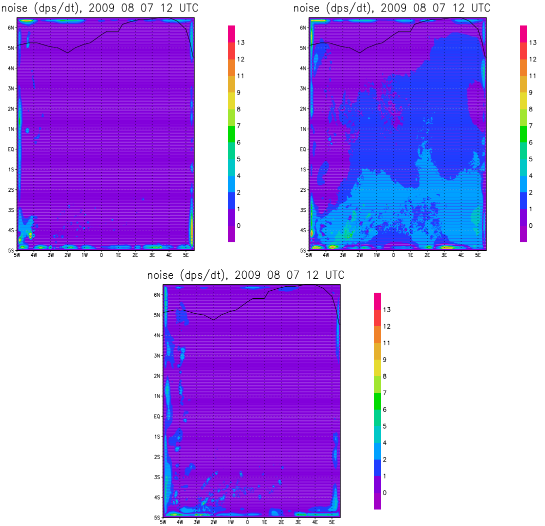


Fig.18: area plots of noise (dPs/dt) at first time step; *integration* with ens BC, *analysis* was done with determ. first guess, ens. first guess and ens. first guess with hydrostatic balancing applied.

hydrostatic balancing reduces noise in the interior, no effect at the boundaries

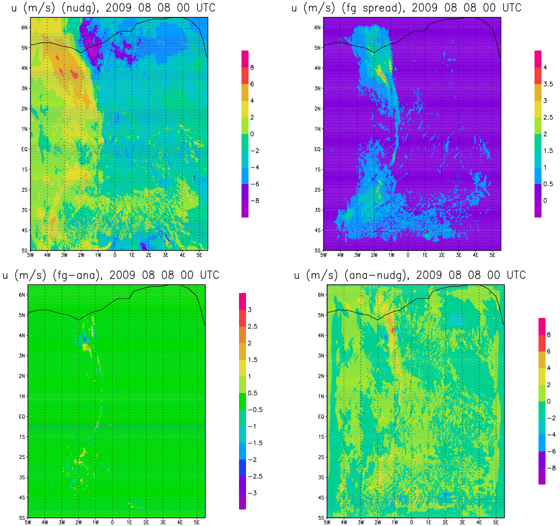


Fig.19: nudging at 2009080800 UTC, model level 1, spread (u in m/s) (det. BC) of first guess, fg minus LETKF analysis and LETKF analysis minus nudging

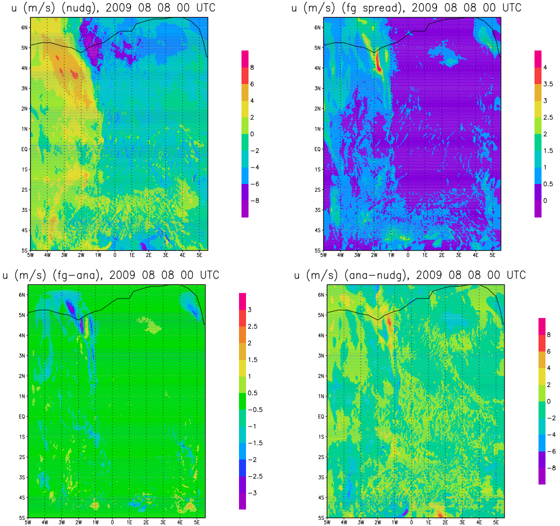


Fig.20: nudging at 2009080800 UTC, model level 1, spread (u in m/s) (ens. BC) of first guess, fg minus LETKF analysis and LETKF analysis minus nudging

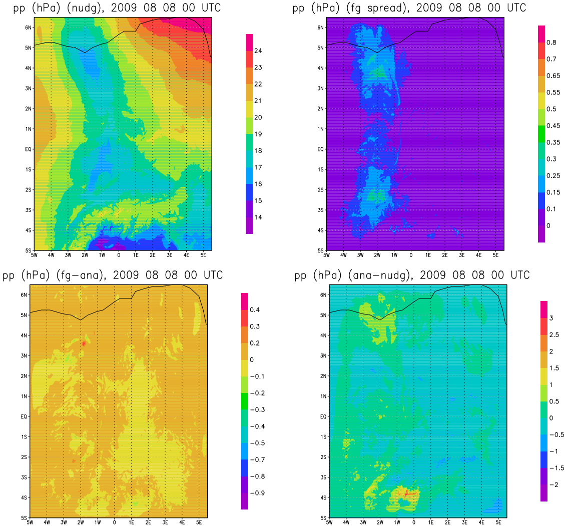


Fig.21: nudging at 2009080800 UTC, model level 1, spread (pp in hPa) (det. BC) of first guess, fg minus LETKF analysis and LETKF analysis minus nudging

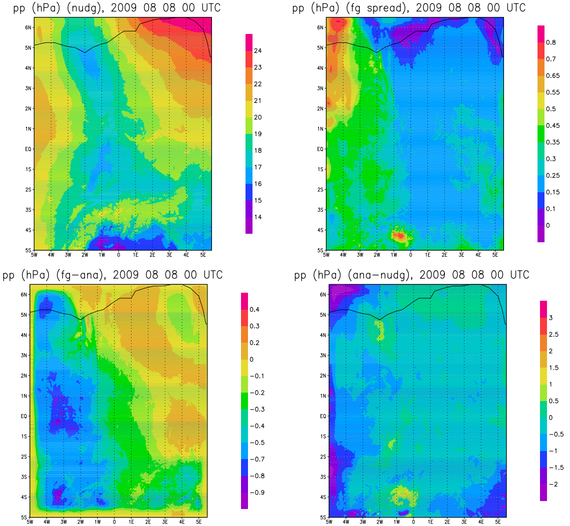


Fig.22: nudging at 2009080800 UTC, model level 1, spread (pp in hPa) (ens. BC) of first guess, fg minus LETKF analysis and LETKF analysis minus nudging

Conclusions / open questions

- noise: we already learned something from the (preliminary) experiments, but some questions remain:
  - no vertical localization: rms/spread ratio gets worse, noise slightly reduced; $\rightarrow$ use vertical localization
  - hydrostatic balancing of analysis increments reduces noise (at the beginning of integration) - effect on rms/spread has to be studied
- spread: structures (small at large scales, high values at small scales) seem to be appropiate (in cases studied so far), but amplitude too small?
- spread: influence of adaptive algorithms in case of ens BC has to be investigated

Outlook

- run with 40 members when becoming operational
- further investigations:
  - further examine and combine adaptive methods
  - try saturation adjustement
  - runs with obs from COSMO feedback files (more dense), increase update frequency, use NUMEX
  - tuning of parameters , e.g. localization length scales
  - runs with BC from global LETKF; look at spread of BC
  - model error (model perturbations): 2 projects planned (Italy/Russia) to account for model error; (stochastic) physics perturbations

LETKF Theory

- let $\mathbf{w}$ denote gaussian vector in k-dimensional ensemble space with mean 0 and covariance $\mathbf{I}/(k-1)$
- let $\mathbf{X}^b$ denote the (background) ensemble perturbations
- then $\mathbf{x} = \bar{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w}$ is the corresponding model state with mean $\bar{\mathbf{x}}^b$ and covariance $\mathbf{P}^b = (k-1)^{-1}\mathbf{X}^b(\mathbf{X}^b)^T$
- let $\mathbf{Y}^b$ denote the ensemble perturbations in observation space and $\mathbf{R}$ the observation error covariance matrix

LETKF Theory

- do analysis in the k-dimensional ensemble space

$$\bar{\mathbf{w}}^a = \tilde{\mathbf{P}}^a (\mathbf{Y}^b)^T \mathbf{R}^{-1} (\mathbf{y} - \bar{\mathbf{y}}^b)$$
$$\tilde{\mathbf{P}}^a = [(k-1)\mathbf{I} + (\mathbf{Y}^b)^T \mathbf{R}^{-1} \mathbf{Y}^b]^{-1}$$

- in model space we have

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{X}^b \bar{\mathbf{w}}^a$$
$$\mathbf{P}^a = \mathbf{X}^b \tilde{\mathbf{P}}^a (\mathbf{X}^b)^T$$

- Now the analysis ensemble perturbations - with $\mathbf{P}^a$ given above - are obtained via

$$\mathbf{X}^a = \mathbf{X}^b \mathbf{W}^a,$$

where $\mathbf{W}^a = [(k-1)\tilde{\mathbf{P}}^a]^{1/2}$