LETKF for COSMO-DE: recent developments

H. Reich, A. Rhodin, R. Potthast, C. Schraff

DWD, Offenbach

HERZ meeting 24. June 2013, Munich

Outline

- Multistep analysis
  - motivation and status
  - theory
- Status of KENDA
  - general setup
  - first results
  - finished and current experiments
  - radar operator: status & plans
- Outlook, open questions, discussion

multistep analysis: motivation

also known as *successive, serial or batch assimilation*, but so far used for computational/algorithmic reasons.
For COSMO-LETKF various motivations (not completely independent) to use *multistep analysis*:

- local / nonlocal observations (e.g. Radiances, Christoph's idea)
- in relation with *adaptive localization*: different observation densities (conventional / radar)
- different observed scales (synoptic / convectional scale), observation errors

status: technically implemented / tested in COSMO-LETKF

- next step: test with radar data

paper: (together with África Periáñez); prove equivalence of 1-step/multi-step for (ensemble) KF; investigate effect of localization

multistep analysis: theory

Theorem

For the standard Kalman Filter with analysis $\varphi^{(a)}$ at time t, and the multistep Kalman Filter with analysis $\varphi^{(a,\xi)}$ for $\xi = 1, ..., q$ we have

$$\varphi^{(a)} = \varphi^{(a,q)} \quad \textit{and} \quad B^{(a)} = B^{(a,q)}.$$

Theorem

For the covariance matrices $B^{(a)} := Q^{(a)}(Q^{(a)})'$ generated by the classical EnKF and the covariance matrix $B^{(a,q)} := Q^{(a,q)}(Q^{(a,q)})'$ by the multi-step EnKF we have

$$B^{(a)} = B^{(a,q)}.$$

multistep analysis: theory

We define

$$A_1 := (Q^{(b)})'(H^{(1)})'(R^{(1)})^{-1}H^{(1)}Q^{(b)}$$

and

$$A_2 := (Q^{(b)})'(H^{2)})'(R^{(2)})^{-1}H^{(2)}Q^{(b)}.$$

Theorem (Multistep-EnKF Equivalence)

Assume that the observation operators $H^{(1)}$ and $H^{(2)}$ for two different sets of measurements satisfy $A_1 A_2 = A_2 A_1$. Then the analysis ensemble generated by the multi-step EnKF with square root filter is identical to the analysis ensemble generated by the classical EnKF.

LETKF general setup

	GME	**COSMO**
ensemble member	40 + 1 (3dVar)	40 + 1 (det run)
horizontal resolution (ens)	ni128 ($\approx$ 60 km)	2.8 km
horizontal resolution (det)	ni256 ($\approx$ 30 km)	2.8 km
horiz. local. length scale	300 km	100 km
vert. local. length scale (ln p)	0.3 (0.075-0.5)	0.3 (0.075-0.5)
adapt. horiz. local.	not tested	tested (new exp)
additive model error	T (3dVar **B**)	F
(adaptive) inflation	T	T
conventional obs	T	T
Radiances	T (AMSU-A)	F
GPS-RO	new exps	F
Radar data	F	operator implemented
cloud height	F	Annika's talk
update frequency	3h	1h ($\rightarrow$ 30/15 min)

KENDA status: summary

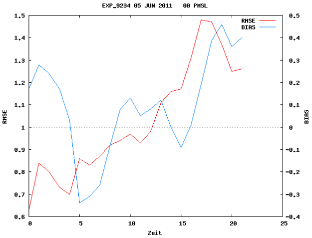
3 experiments so far:

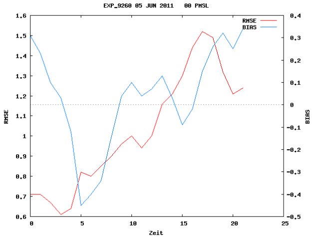
- 9125 (base experiment)
- 9203 (modified observation errors)
- 9259 (saturation adjustement switched on, slightly modified observation errors, no assimilation of T2M, RH2M, (weak) adaptive localization used

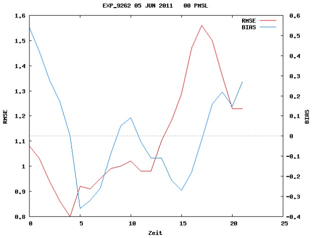
Verification:

- deterministic forecast: Klaus Stephan runs forecast up to 21 h, results comparable to nudging
- EPS: Richard Keane has run KENDA Ensemble from exp9259 for 2011060100 UTC; rmse larger than for COSMO-DE EPS, spread smaller

RMSE/BIAS of deterministic KENDA forecasts



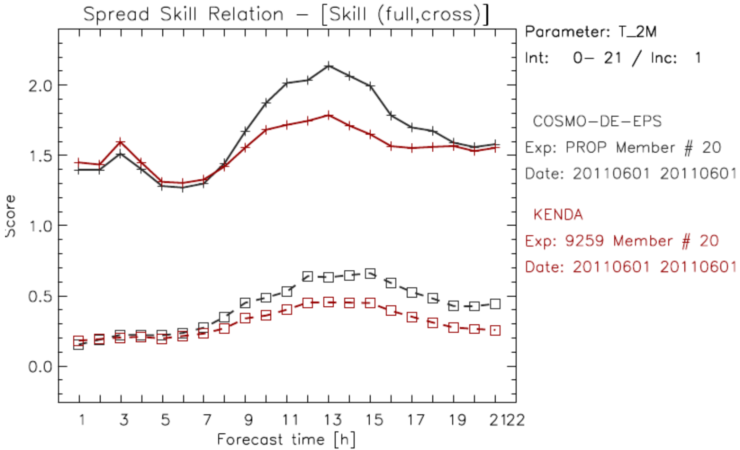




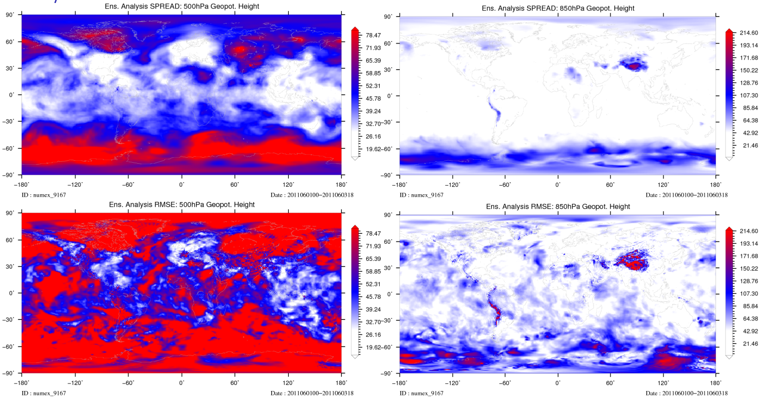
RMSE and BIAS of surface pressure, verified against SYNOP stations for LETKF, nudging and free forecast

LETKF comparable to nudging

KENDA EPS: comparison with COSMO-DE EPS



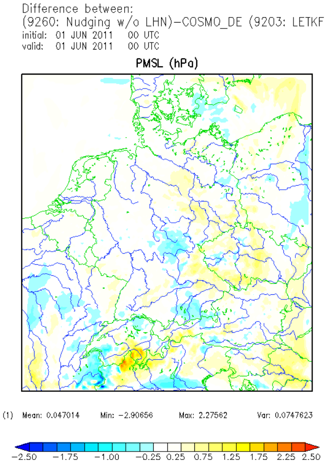
RMSE/SPREAD of GME-LETKF

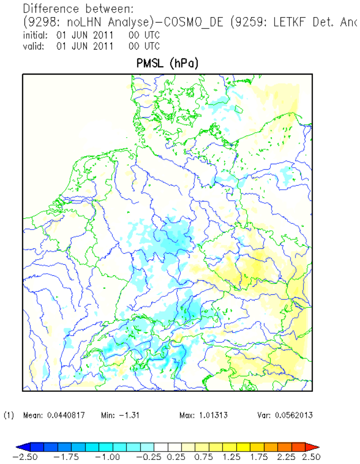


SPREAD and RMSE of GME-LETKF analysis (geop. height, 500 and 850 hPa)

very low SPREAD over Europe and other data-rich areas → BC for COSMO-LETKF will also suffer from lack of spread; test/tune adaptive methods

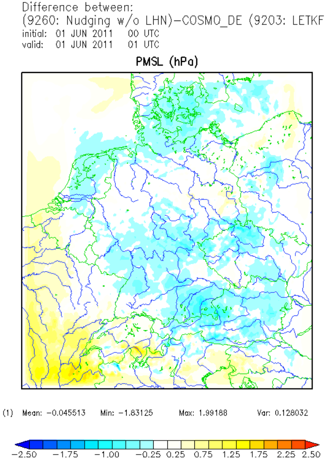
KENDA: influence of boundary conditions I

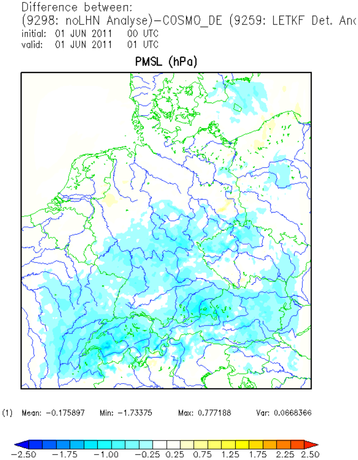




left: difference between nudging w/o LHN and LETKF with BC from COSMO-EU (nudging) and deterministic GME-LETKF (LETKF;currently 3dVar); right: same BC (GME-LETKF). (PMSL at 00 forecast time)

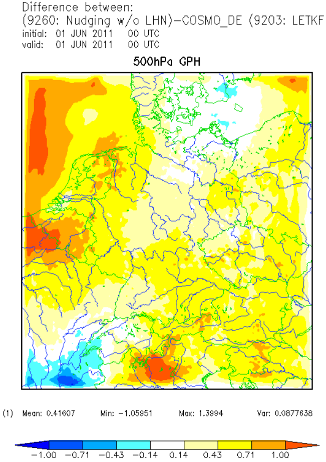
KENDA: influence of boundary conditions II

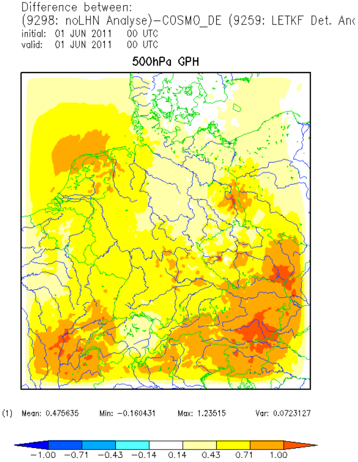




same as before, but for 01 forecast time

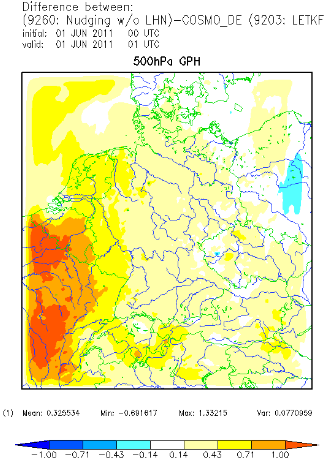
KENDA: influence of boundary conditions III

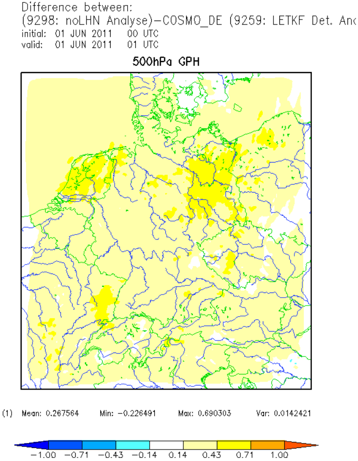




same as before, but for geopotential at 500 hPa, 00 forecast time

KENDA: influence of boundary conditions IV





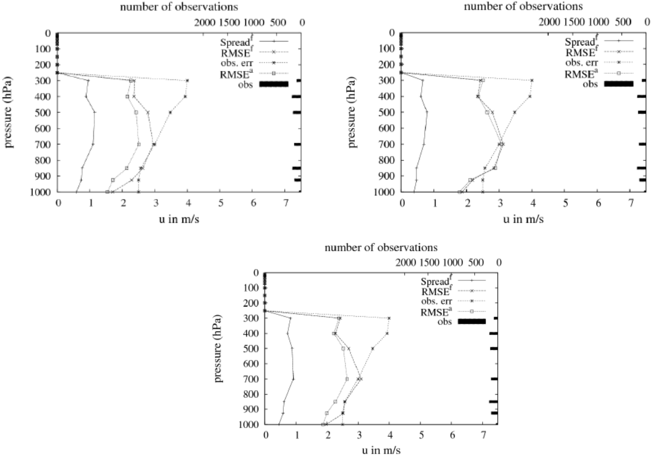
same as before, but for geopotential at 500 hPa, 01 forecast time

radar operator

First results from Yuefei Zeng:

- radar data (radial wind) assimilated for 2011053118 UTC (1 analysis only, no cycling)
- 3 Experiments $E0, E1, E2$.
- all experiments use conventional data, settings are:
  - $E0$ radar passive
  - $E1$ radar active , localization length scale $100km$ for conventional/radar
  - $E3$ radar active , localization length scale $100km$ for conventional data, $20km$ for radar
- no multistep analysis, but different localization radii used within 1 analysis

radar operator



verification against AIREP for *u* wind component, Experiments E0, E1, E2

Outlook

- *multistep analysis*: test with radar data (together with Yuefei Zeng), continue with theoretical work (paper), tests with toy models (more advanced models than Lorenz 95 needed)
- *technical (data base) problems* need to be solved to run experiments...; stand alone ($\approx 1 week$) as alternative (store data in *ECFS*, use Thomas' stand-alone script)
- first results from KENDA (summary):
  - *deterministic*: in general comparable with nudging, but differences for surface pressure/geopotential (hydrostatic balancing, BC?)
  - *ensemble* not as good as COSMO-DE EPS, esp. lack of spread (due to BC / interior?); but only 1 forecast evaluated...
- *additional observations*: radar obs (radial winds, reflectivity), cloud height (Annika's talk)

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

LETKF Theory

- it's possible to obtain a *deterministic run* via

$$\mathbf{x}_a^{det} = \mathbf{x}_b^{det} + \mathbf{K}\left[\mathbf{y} - H(\mathbf{x}_b^{det})\right]$$

with the *Kalman gain* **K**:

$$\mathbf{K} = \mathbf{X}_b\left[(k-1)\mathbf{I} + \mathbf{Y}_b^T\mathbf{R}^{-1}\mathbf{Y}_b\right]^{-1}\mathbf{Y}_b^T\mathbf{R}^{-1}$$

- the deterministic analysis is obtained on the same grid as the ensemble is running on; the *analysis increments* can be interpolated to a higher resolution