Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Experience with Kilometer Scale Ensemble Data Assimilation for the COSMO Model (KENDA)

Hendrik Reich, Christoph Schraff, Andreas Rhodin, Annika Schomburg, Klaus Stephan, Africa Perianez, Roland Potthast

DWD, Offenbach

ISDA 2015, Kobe, Japan

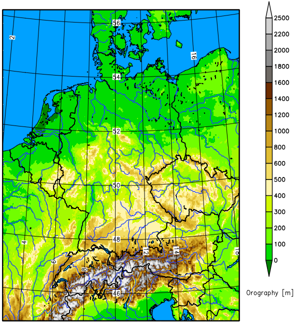
Outline

- General overview on KENDA / experimental setup
- Results
- Next steps, outlook

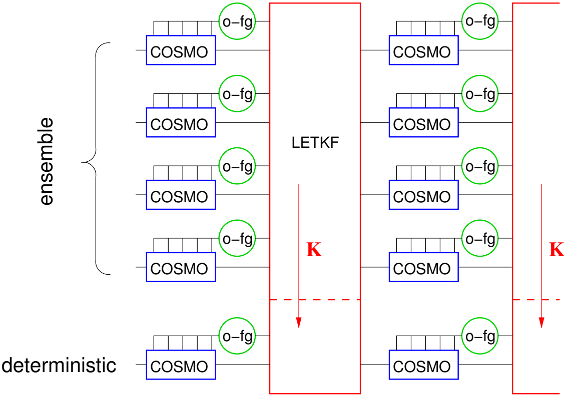
KENDA experiments: general overview

- **KENDA**: **K**ilometer Scale **En**semble **D**ata **A**ssimilation, priority project within **COSMO** (**Co**nsortium for **s**mall-scale **mo**delling)
- implementation of **LETKF** (**L**ocal **E**nsemble **T**ransform **K**alman **F**ilter) following Hunt et. al.

Orography of operational COSMO-DE domain used for KENDA-LETKF with 2.8 km horizontal resolution. The domain size is about 1170 km × 1280 km .



KENDA experiments: general overview



*KENDA system setup; 'o-fg' denotes observation minus first guess, '**K**' the Kalman Gain for the analysis mean.*

KENDA experiments: general overview

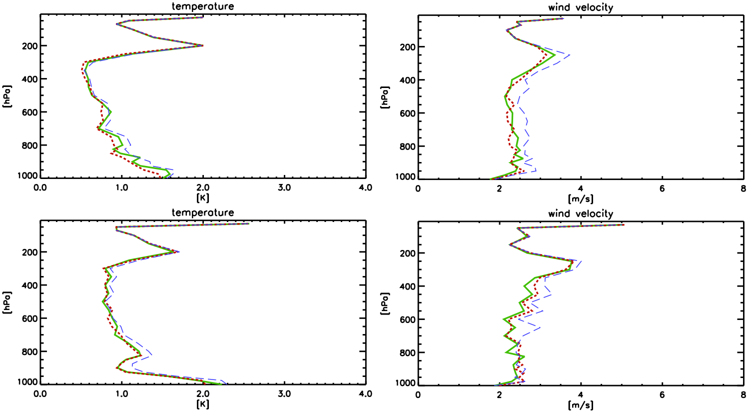
- first goal: replace (operational) nudging with deterministic LETKF analysis (second step: use as COSMO-DE EPS initial conditions)
- $\rightarrow$ focus on quality of deterministic analysis/forecast, compare with nudging (incl. LHN) as benchmark
- BACY (**ba**sic **cy**cling, bash script environment) for ICON-LETKF (global scale, $\rightarrow$talk by Ana Fernandez) and KENDA (convective scale)
- KENDA-BACY:
  - analysis cycle: LETKF incl. det analysis, nudg cycle with same obs set; verify against obs (surface/upper air)
  - forecast cycle: nudgecast (nudg analysis), free fc (w/o obs), det LETKF, verify against obs (surface/upper air)

KENDA experiments: general overview

- BC for KENDA are taken from ICON-BACY, nudging and deterministic LETKF use same BC for analysis cycle and forecast
- BC from GME-LETKF, period: 20110601-20110607
  - adaptive localization + R-matrix estimation (ens-space), relaxation methods (RTPP) tested
- BC from ICON-BACY, 2 periods:
  - 20120719-20120725,
    several experiments testing effect of soil moisture perturbations, latent heat nudging, RTPP
  - 20140517-20140528,
    compare LETKF/nudging within longer period

KENDA setup for 2012 / 2014 periods

variable / feature	**value**
ensemble size k	40
deterministic run	1
horiz. resolution ens. + det. run	2.8 km
forecast frequency / length	6 h / 24 h
analysis update frequency	1 h
vert. localis. length scale (ln p)	0.075 - 0.5
horizontal localisation $\rightarrow$ target weighted no. obs. $N_{loc}^{o_{ef}}$ $\rightarrow$ min. local. length scale r_{loc}^{min} $\rightarrow$ max. local. length scale r_{loc}^{max}	adaptive 100 50 km 100 km
multiplicative covariance inflation $\rightarrow$ lower / upper limit of ρ	adaptive 0.5 / 3.0
RTPP relaxation weight α_p	0.75



Vertical profiles of RMSE against observations (radiosonde, aircraft, and wind profiler) of 6-hour forecasts, KENDA-LETKF analyses (solid green line), nudging analyses (dotted red line), free forecast (thin dashed blue line). Left temperature; right: wind speed; upper row GME-LETKF lateral BC (for 2 - 7 June 2011), lower row ICON-LETKF lateral BC (20 - 25 July 2012)

model error: inflation/relaxation methods

- (1): compare "observed" with "expected" quantities:

$$\left\langle (y - H(x_b))(y - H(x_b))^T \right\rangle = \mathbf{R} + \rho \mathbf{H P_b H^T}$$

$$\left\langle (H(x_a) - H(x_b))(y - H(x_b))^T \right\rangle = \rho \mathbf{H P_b H^T}$$

- (2): "relaxation" methods: e.g. relaxation to prior spread (RTPS):

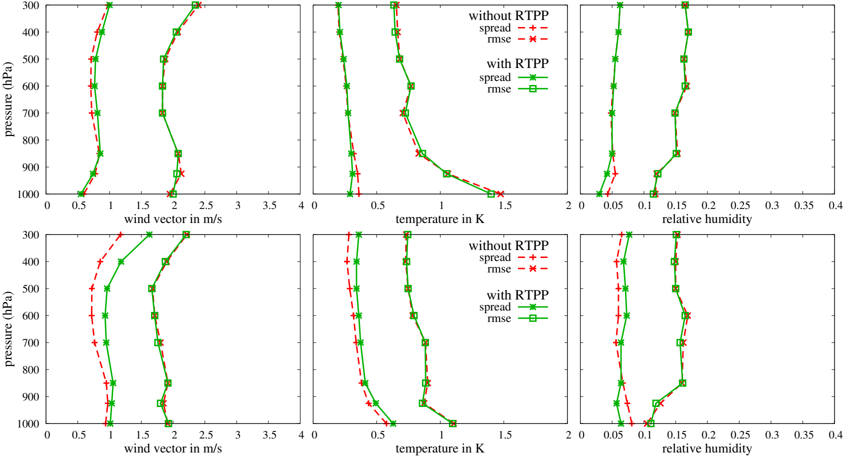
$$X_a^{i,infl} = \rho X_a^i, \; \rho = \sqrt{\alpha \frac{\sigma_b - \sigma_a}{\sigma_a} + 1}$$

  or relaxation to prior perturbation (RTPP):

$$X_a^{i,infl} = (1 - \alpha) X_a^i + \alpha X_b^i$$

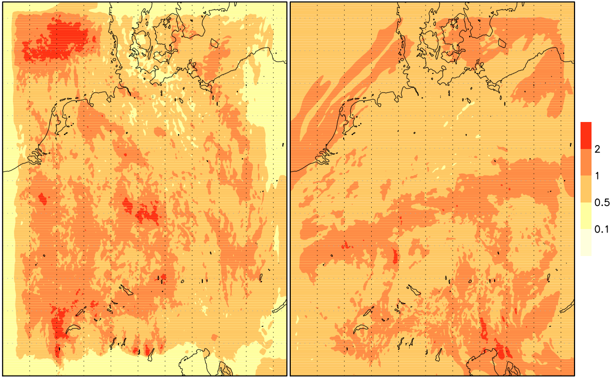
- (1) works in observation space; tries to increase/decrease spread to fulfill statistical relations
- (2) works in model space; "corrects" reduction of spread due to assimilation of observations (RTPP: similar to additive pert., "directions" of fg pert partly remain; RTPS: inflates ana pert directions)

Influence of LBC spread on benefit from RTPP



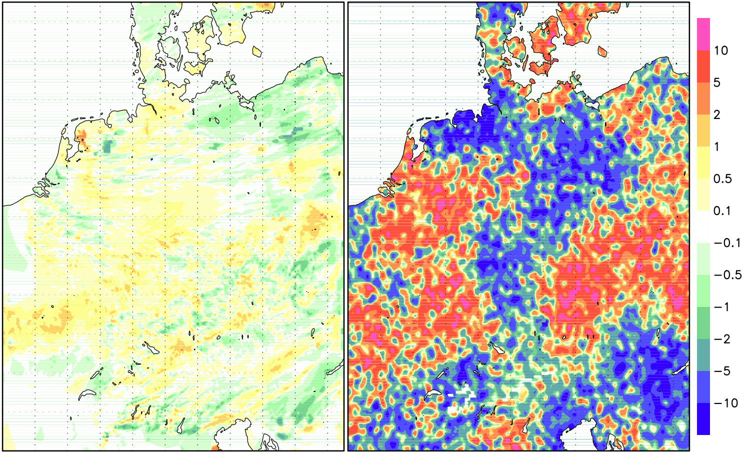
Profiles of first guess spread / RMSE against upper-air observations; with RTPP (solid green lines), without RTPP (dashed red lines). Panels are for wind vector, temperature, and relative humidity (from left to right); upper row GME lateral BC, lower row ICON BC's.

GME / ICON BC

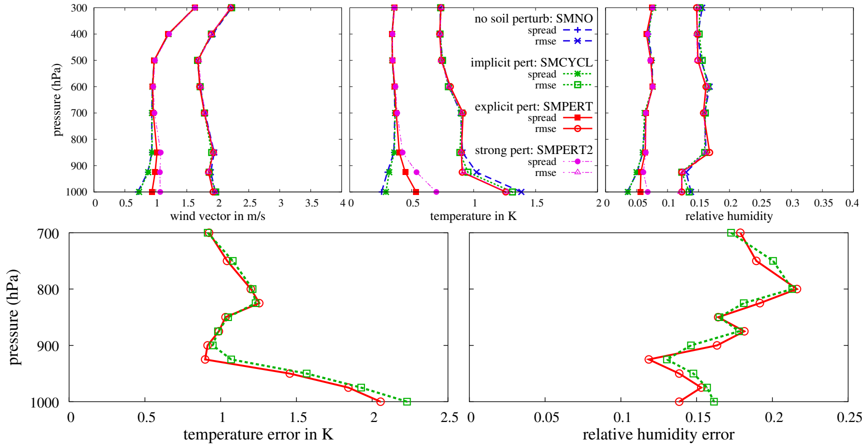


Analysis ensemble spread for the zonal wind component at model level 25 ($\approx$ 500 hPa), left panel for GME-LETKF BC (for 6 June 2011), right panel for ICON-LETKF lateral BC (24 July 2012)

Soil moisture perturbations



Spread of soil moisture (WSO in kg m^{-2}) in layer 3 (i.e. 3 - 9 cm depth) after five days of cycling, left for experiment `SMCYCL` with implicit soil moisture perturbations, right for `SMPERT` with explicit soil moisture perturbations



Upper row: first guess spread/RMSE for exp SMNO without soil moisture perturbations (dashed blue lines), SMCYCL with implicit perturbations (dotted green), SMPERT with explicit perturbations (solid red), and SMPERT2 with doubled explicit perturbations (thin dash-dotted purple). Lower row: 6-h forecast RMSE of temperature (left) and relative humidity (right) for SMCYCL (dotted green line) and SMPERT (solid red).

Assimilation of Radar-derived precipitation by LHN

Required relation:

precipitation rate $\leftrightarrow$ model variables

(observed) (info required by nudging)

precipitation $\leftrightarrow$ condensation $\leftrightarrow$ release of latent heat

$\rightarrow$ Assumption: vertically integrated latent heat release $\propto$ precipitation rate

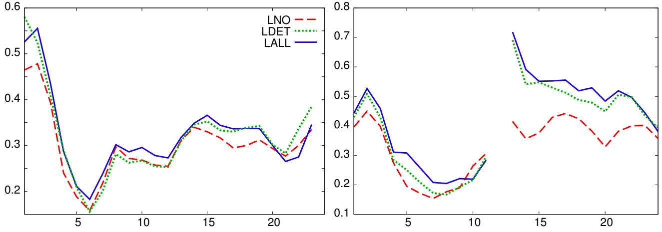
Approach: modify latent heating rates such that the model responds by producing the observed precipitation rates $\rightarrow$ Latent Heat Nudging (LHN)

$$\frac{\partial T}{\partial t} = F(t) + \left.\frac{\partial T}{\partial t}\right|_{nudging} + \left.\frac{\partial T}{\partial t}\right|_{LHN}$$

$$\Delta T_{LHN} = (\alpha - 1) \cdot \Delta T_{LH} \quad \text{with} \quad \alpha = \frac{RR_{obs}}{RR_{ref}}$$

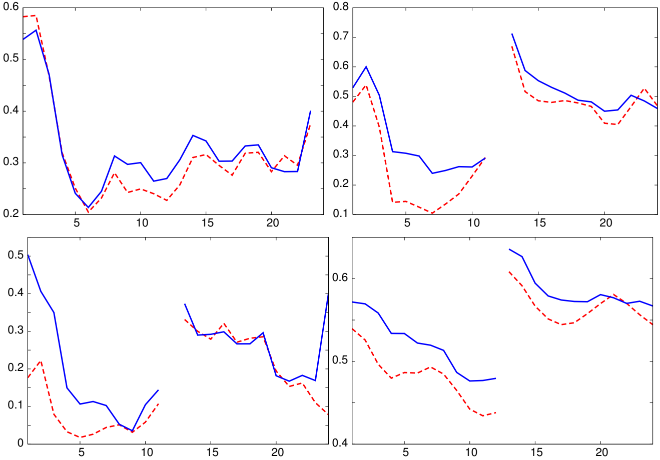
Use LHN in LETKF until assimilation of radar reflectivities is available

Effect of latent heat nudging on radar-derived precipitation rates



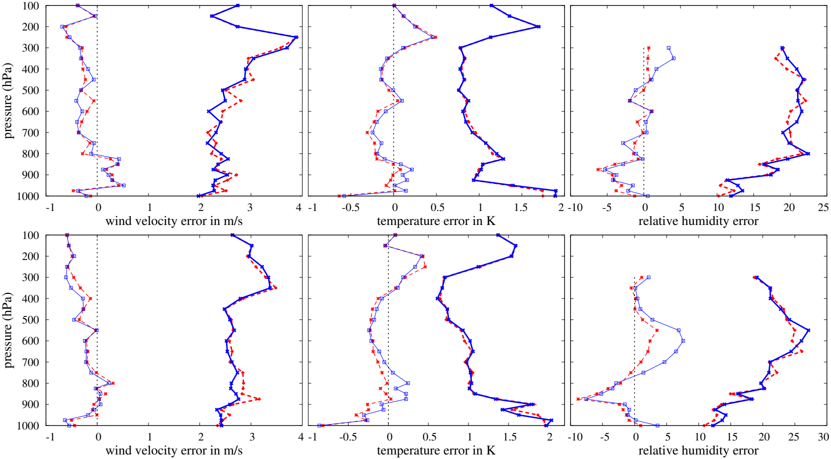
Verification of 1-hourly precipitation against radar-derived precipitation: Fraction Skill Score (FSS) with a scale of 11 grid points ($\approx$ 30 km) for a threshold of 0.1 mm/h as a function of daytime, left for 0-UTC forecast runs, right for 12-UTC forecast runs, for experiments `LNO` without LHN (dashed red line), `LDET` with LHN only in the deterministic run (dotted green), and `LALL` with LHN also in the LETKF ensemble (solid blue).(2012 period) FSS: larger values are better

KENDA-LHN vs. NUDG-LHN: precipitation



Fraction skill scores (FSS) for verification of precipitation as last figure, but for experiments KENDA-LHN (LETKF + LHN; solid blue line) and NUDGE-LHN (dashed red line). Upper row for the 6-day period T2012 in July 2012, lower row for the 12-day period T2014 in May 2014.

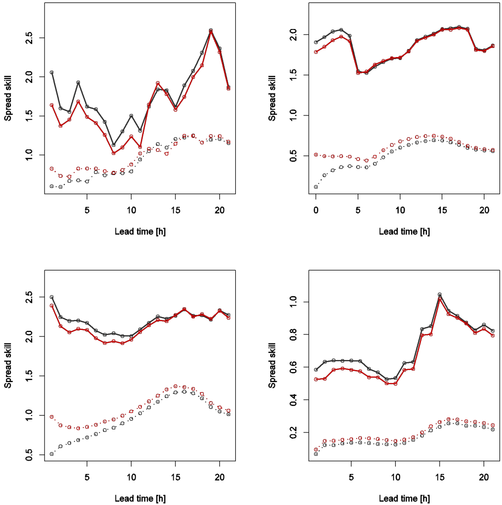
KENDA-LHN (blue) vs. NUDG-LHN (red): 6h upper air



Vertical profiles of bias and RMSE against radiosonde observations of 6-hour forecasts started from KENDA-LETKF analyses (experiment `KENDA-LHN`, solid blue line) or nudging analyses (`NUDGE-LHN`, dotted red line) for wind speed (left panels), temperature (middle), relative humidity (right), for the 6-day period `T2012` in July 2012 (upper panels) and the 12-day period `T2014` in May 2014 (lower panels).

	variable	T2012		T2014	
		rmse	bias	rmse	bias
upper air	geopotential	=	=	=	=
	temperature	=	=	(+)	(+)
	relative humidity	=	=	=	–
	wind speed	+	=	+	=
	wind direction	+	=	(+)	=
surface	2-m temperature	=	=	+	=
	2-m dew point	=	=	+	+
	10-m wind	=	=	=	=
	surface pressure	–	=	+	=
	total cloud	=	=	=	=
	low cloud	+	+	–	=
	mid-level cloud	+	+	=	=
	high cloud	–	–	–	–
radar	precip 0 UTC runs	+/(–)	+	++	++
	precip 12 UTC runs	++	++	++	++

First results from COSMO-DE EPS (thx to Richard Keane)



RMSE (solid line) and spread (dashed line) of COSMO-DE EPS using BCEPS (black) and KENDA (red) initial conditions; 10-40m Wind, T2M, 10M Wind Gusts, Precipitation. spread/rmse for KENDA is overall better

Conclusions

- ICON-BC: spread more homogenous (slightly increased), but still only 80 km resolution!
- 24 h forecast of det run, nudging, free fc: deterministic LETKF forecast overall similar /slightly better quality than nudging forecast (except high clouds)
- plots shown are for 6h forecasts, but results also hold for 12h, 18h forecasts (differences get smaller)
- LHN: no influence on upper air verification; better results for Radar verification (precipitation, 00 and 12 UTC runs)
- soil moisture perturbations: positive impact on spread/rmse close to surface

Next steps

- compute more extended periods, 30 days winter/summer
- use of additional observations (e.g. radar radial winds, SEVIRI, radar reflectivity)
- compute/investigate ensemble forecasts
- test impact of higher resolution of ICON-BC
- $\rightarrow$ more talks/posters on ICON-LETKF/KENDA/new observations: e.g. Ana Fernandez, Roland Potthast, Andreas Rhodin (DWD); Africa Perianez (Riken), Theresa Bick (Univ. Bonn)

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 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* $\mathbf{K}$:

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

- 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

Soil moisture perturbations

- perturb soil moisture (and SST) with defined spatial and temporal length scales and amplitude
- soil moisture: 2 length scales, 100 km (synoptic), 10 km (convection)
- cut-off if moisture is below zero or above capacity ($\rightarrow$ bias)
- next step: for soil moisture, limit perturbation amplitude to “available capacity” (avoid bias)

```
name = 'W_SO'              ! disturb soil moisture
scales = 0.002 100 1 24 ! 0.004 of soil capacity, 100km,
                           ! 1m vertical, 24 hour
         0.002 10 1 24     ! 0.004 of soil capacity, 10km,
                           ! 1m vertical, 24 hour
name = 'SST'            ! disturb SST
scales = 1  100 0 24 ! 1K +  100 km length scale pattern,
                        ! 24 hour
```