Comparison of KENDA with nudging and impact of latent heat nudging

Hendrik Reich, Christoph Schraff, Andreas Rhodin, Klaus Stephan, Roland Potthast

DWD, Offenbach

COSMO General Meeting 2015, Wroclaw, Poland
Outline

- General overview on experiments
- Results
- Remaining problems, next steps
KENDA experiments: general overview

- **KENDA**: Kilometer Scale Ensemble Data Assimilation
- implementation of **LETKF** (Local Ensemble Transform Kalman Filter) 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. Domain and resolution will be increased soon (COSMO-D2, 2.2 km resolution).
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)
- → focus on quality of deterministic analysis/forecast, compare with nudging (incl. LHN) as benchmark
- **BACY (basic cycling, bash script environment) for ICON-LETKF and KENDA**
- **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), det LETKF, verify against obs (surface/upper air/radar precipitation)
- speed ≈ 2.0 for KENDA, but needs large hard disk storage
KENDA experiments: general overview

- BC for KENDA are taken from ICON-BACY, nudging and deterministic LETKF use same BC for analysis cycle and forecast.

- ICON-BC (80 km resolution for ensemble members, 40 km for deterministic 3dVar-run):
  - 20120719-20120725, several experiments testing effect of soil moisture perturbations, latent heat nudging, RTPP
  - 20140517-20140615, compare LETKF/nudging within longer period
  - Preliminary tests with 20 km resolution BC from ICON-NEST; spread increased
### KENDA setup for 2014 periods

<table>
<thead>
<tr>
<th>variable / feature</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ensemble size $k$</td>
<td>40</td>
</tr>
<tr>
<td>deterministic run</td>
<td>1</td>
</tr>
<tr>
<td>horiz. resolution ens. + det. run</td>
<td>2.8 km</td>
</tr>
<tr>
<td>forecast frequency / length</td>
<td>6 h / 24 h</td>
</tr>
<tr>
<td>analysis update frequency</td>
<td>1 h</td>
</tr>
<tr>
<td>vert. localis. length scale (ln p)</td>
<td>0.075 - 0.5</td>
</tr>
<tr>
<td>horizontal localisation</td>
<td>adaptive</td>
</tr>
<tr>
<td>→ target weighted no. obs. $N^\text{eff}_{\text{loc}}$</td>
<td>100</td>
</tr>
<tr>
<td>→ min. local. length scale $r^\text{min}_{\text{loc}}$</td>
<td>50 km</td>
</tr>
<tr>
<td>→ max. local. length scale $r^\text{max}_{\text{loc}}$</td>
<td>100 km</td>
</tr>
<tr>
<td>multiplicative covariance inflation</td>
<td>adaptive</td>
</tr>
<tr>
<td>→ lower / upper limit of $\rho$</td>
<td>0.5 / 3.0</td>
</tr>
<tr>
<td>RTPP relaxation weight $\alpha_p$</td>
<td>0.75</td>
</tr>
</tbody>
</table>
model error: inflation/relaxation methods

(1): compare “observed” with “expected” quantities:

\[
\begin{align*}
\langle (y - H(x_b))(y - H(x_b))^T \rangle &= \mathbf{R} + \rho \mathbf{HP}_b \mathbf{H}^T \\
\langle (H(x_a) - H(x_b))(y - H(x_b))^T \rangle &= \rho \mathbf{HP}_b \mathbf{H}^T
\end{align*}
\]

(2): “relaxation” methods: e.g. relaxation to prior spread (RTPS):

\[
X^{i,\text{infl}}_a = \rho X^i_a, \quad \rho = \sqrt{\alpha \frac{\sigma_b - \sigma_a}{\sigma_a} + 1}
\]

or relaxation to prior perturbation (RTPP):

\[
X^{i,\text{infl}}_a = (1 - \alpha) X^i_a + \alpha X^i_b
\]

(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)
Spread-skill ratio $r_s = \frac{\sigma_{\tilde{f} - \tilde{f}}}{\sqrt{\sigma_{f_d - o}^2 - \sigma_{od}^2}}$, where $\sigma_{od}^2$ is the (diagnostic) observation error variance, $\sigma_{\tilde{f} - \tilde{f}}$ the spread and $\sigma_{f_d - o}^2$ the RMSE.

→ With sufficient spread in the boundary conditions, RTPP (plus multiplicative inflation) gives reasonable spread-skill ratio
Effect of LHN on radar-derived precipitation rates

- Fraction Skill Score (FSS) of 1-hourly precipitation (11 grid points, $\approx 30\text{ km}$, 0.1 mm/h threshold), 00/12-UTC forecast (left/right)
- **KENDA** without LHN
- **KENDA-LDET** with LHN only in the det run
- **KENDA-LHN** with LHN also in the LETKF ensemble

Similar results for all scales and forecast start times.
KENDA-LHN vs. NUDGE-LHN: precipitation

FSS as before, 0.1 mm/h and 1.0 mm/h threshold (upper/lower row), 00/12-UTC forecast (left/right)

- **KENDA-LHN** (LETKF + LHN)
- **NUDGE-LHN** (nudging + LHN)
- **KENDA** (LETKF without LHN)
- **NUDGE** (nudging without LHN).

Similar results for all scales and forecast start times.

Hendrik Reich, Christoph Schraff, Andreas Rhodin, Klaus Stephan, Roland Potthast

Comparison of KENDA with nudging and impact of latent heat nudging.
KENDA-LHN (blue) vs. NUDGE-LHN (red): 6h upper air

Vertical profiles of bias and RMSE against radiosonde observations; 6-hour forecasts, wind speed and direction (upper row), temperature and relative humidity (lower row),) started from:

- KENDA-LHN analyses,
- NUDGE-LHN analyses
## Surface verification results

<table>
<thead>
<tr>
<th>experiment</th>
<th>PS [hPa]</th>
<th>T2M [K]</th>
<th>TD2M [K]</th>
</tr>
</thead>
<tbody>
<tr>
<td>KENDA-LHN</td>
<td>.53</td>
<td>2.03</td>
<td>3.33</td>
</tr>
<tr>
<td>NUDGE-LHN</td>
<td>.55</td>
<td>2.06</td>
<td>3.54</td>
</tr>
<tr>
<td>KENDA</td>
<td>.56</td>
<td>2.10</td>
<td>3.55</td>
</tr>
<tr>
<td>NUDGE</td>
<td>.56</td>
<td>2.15</td>
<td>3.89</td>
</tr>
</tbody>
</table>

- Root mean square errors (RMSE) of surface pressure ('PS'), 2-m temperature ('T2M'), and 2-m dewpoint depression ('TD2M') against observations from surface stations.
- Each of the RMSE values shown is an average over the 21 RMSE values valid for the forecast lead times from 1 to 21 hours.

KENDA gives clearly better results for TD2M and slightly better results for T2M and PS (with LHN).
Soil moisture perturbations: Soil Moisture Index (SMI)

MeteoSwiss:

\[ SMI = \frac{WSO - PWP}{FC - PWP} \]  

(PWP = Plant Wilting Point, FC = Field Capacity)

DWD:

\[ SMI = \frac{WSO - ADP}{PV - ADP} \]  

(ADP = Air Dryness Point, PV = Pore Volume)
SMI plots

SMI (area mean of det run, ensemble mean and spread using Eq. (1) for layer 5 (54 cm), layer 4 (18 cm), layer 1 (0.5 cm)) → spread is too large, in layer 5 mean and det run diverge
Conclusions

- ICON-BC: sufficient amount of spread at boundaries, but still only 80 km resolution! → preliminary tests with 20 km resolution BC from ICON-NEST; spread increased

- 24 h forecast of det run, nudging: deterministic LETKF forecast overall similar /slightly better quality than nudging forecast (except relative humidity), especially better results for precipitation

- plots shown are for 6h forecasts, but results also hold for 12h, 18h forecasts (differences get smaller)

- LHN: nearly no influence on upper air verification (wind slightly better); better results for Radar verification (precipitation, 00 and 12 UTC runs)

- soil moisture perturbations: positive impact on spread/rmse close to surface; but seems to introduce bias!! → tune parameters
Next steps

- include SST, SNOW analysis
- compute winter period
- COSMO-D2 experiment, using ICON-NEST with 20 km resolution as BC
- tests with pattern generator
- use of additional observations, e.g. radar radial winds, SEVIRI, radar reflectivity (Theresa Bick, paper will be submitted soon)
- compute/investigate ensemble forecasts
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

Hendrik Reich, Christoph Schraff, Andreas Rhodin, Klaus Stepha

Comparison of KENDA with nudging and impact of latent heat nudging
LETKF Theory

- do analysis in the $k$-dimensional ensemble space

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

- in model space we have

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

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

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

where $W^a = [(k - 1)\tilde{P}^a]^{1/2}$
it’s possible to obtain a *deterministic run* via

\[ x_{a}^{\text{det}} = x_{b}^{\text{det}} + K \left[ y - H(x_{b}^{\text{det}}) \right] \]

with the *Kalman gain* \( K \):

\[
K = X_{b} \left[ (k - 1)I + Y_{b}^{T}R^{-1}Y_{b} \right]^{-1} Y_{b}^{T}R^{-1}
\]

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
Assimilation of Radar-derived precipitation by LHN

Required relation:

precipitation rate ↔ model variables
   (observed)    (info required by nudging)

precipitation ↔ condensation ↔ release of latent heat

→ **Assumption**: vertically integrated latent heat release ∝ precipitation rate

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

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

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

Use LHN in LETKF until assimilation of radar reflectivities is available
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 (→ 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
```
Comparison of KENDA with nudging and impact of latent heat nudging