Representing model error and observation error uncertainty for data assimilation of polarimetric radar measurement

Yuefei Zeng [a]

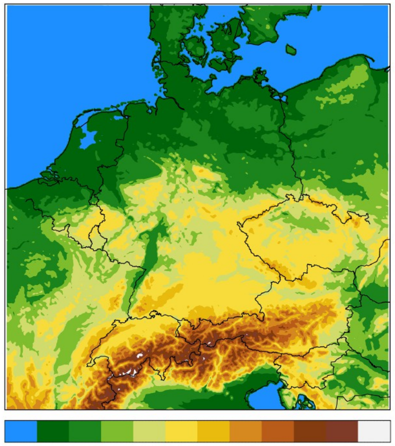
a) Meteorologisches Institut, Ludwig-Maximilians Universität (LMU), Munich, Germany

b) Deutscher Wetterdienst (DWD), Offenbach, Germany

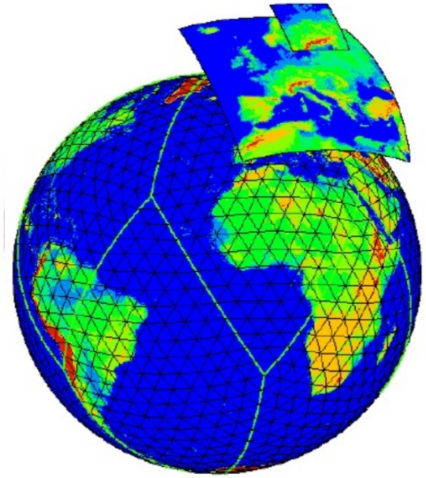
PIs: Tijana Janjic[a], Axel Seifert[b], Daniel Klocke[b]

Background information

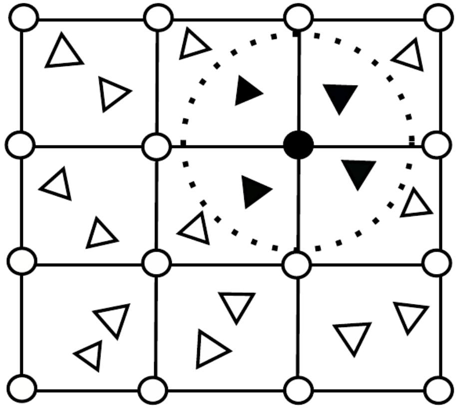
- **K**ilometre-scale **EN**semble **D**ata **A**ssimilation (**KENDA**) system at DWD since March 2017 (Schraff et al. 2016, QJRMS) in operational mode:
  - Convection-permitting model: **CO**nsortium for **S**mall-scale **MO**deling (**COSMO**, Baldauf et al. 2011, MWR) with horizontal resolution 2.8 km
  - Lateral boundary conditions: **ICO**sahedral **N**onhydrostatic (**ICON**, Zängl et al. 2015, QJRMS)
  - Data assim. scheme: **L**ocal **E**nsemble **T**ransform **K**alman **F**ilter (**LETKF**, Hunt et al. 2007, Phy. D)
  - Operational radar network: 16 C-band Doppler radars



Orography COSMO-DE domain (m)



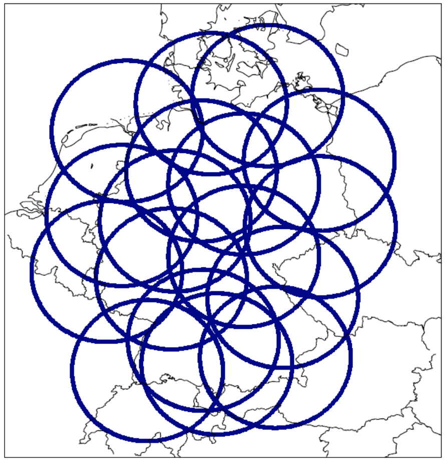
ICON grid structure with nested domains



LETKF : ● Analysis grid point

▲ Observation

…….. Localization radius



Radar network

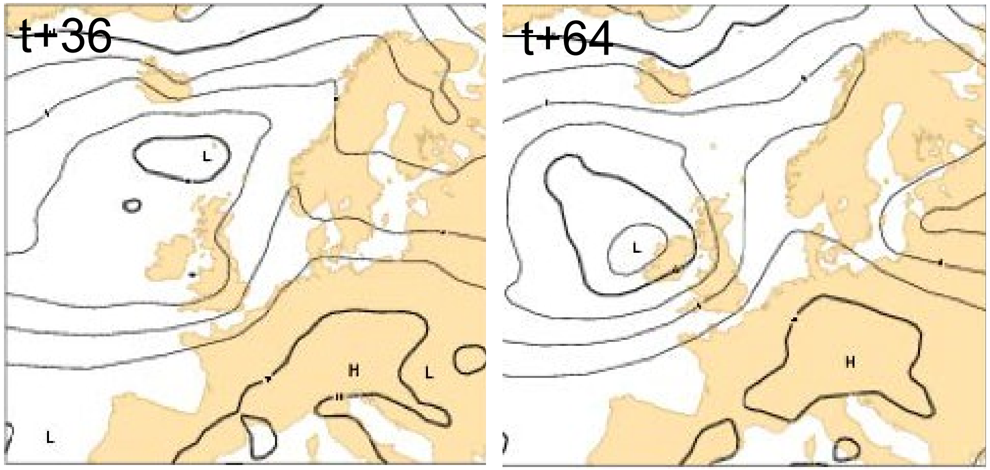
Outline

- Approaches to represent model error
- Alternatives for representing subgrid-scale model error
- Idealized setup for radar data assimilation
- Summary and outlook

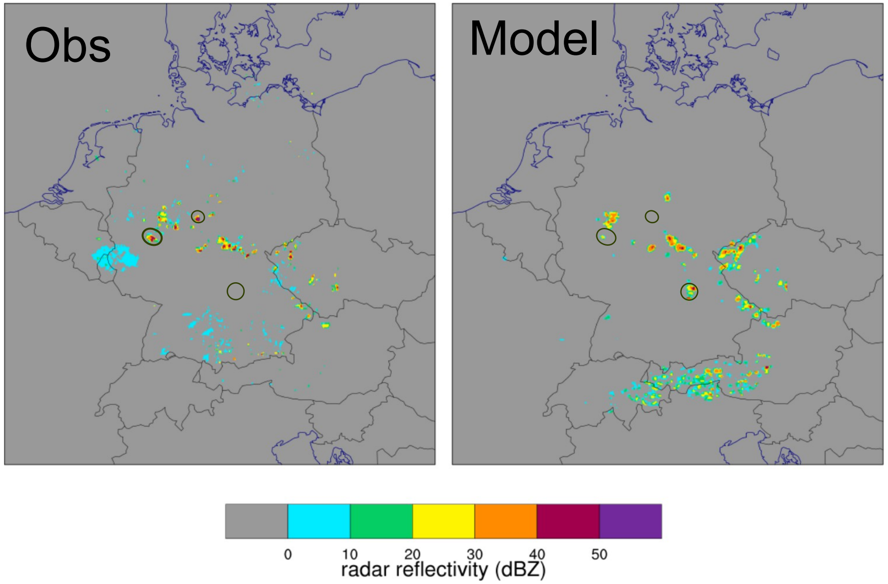
yuefei.zeng@lmu.de

Approaches to represent model error

- The performance of **EnKF** algorithms strongly depends on quality of background error covariance **B**-matrix which should
  - be sufficiently large to account for sampling and model error
  - appropriately describe correlations between variables
  - capture large- and small-scale features of model error



surface pressure charts



Research questions

- How to account for multiscale model error in **B**-matrix in convective-scale assimilation? (Zeng et al. 2018 & 2019a, JAMES)

yuefei.zeng@lmu.de

Approaches to represent model error

Additive Noise (applied to analysis ensemble members): $\boldsymbol{x}^{a(i)} \leftarrow \boldsymbol{x}^{a(i)} + \alpha_a \boldsymbol{\eta}^{(i)}$

- Random draws from climatological **B**-matrix for global EnVar data assim. system for **ICON**

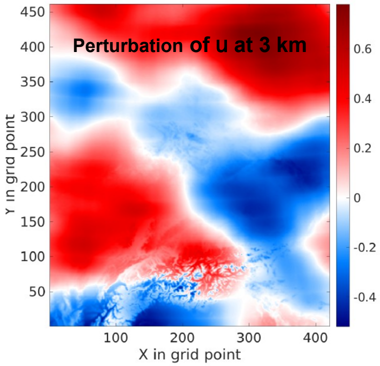
  Construction of **B**-matrix by **NMC**-method (Zeng et al. 2018, JAMES):

  $$\boldsymbol{B} \approx \frac{1}{2} < (\boldsymbol{x}^{t_1} - \boldsymbol{x}^{t_2})(\boldsymbol{x}^{t_1} - \boldsymbol{x}^{t_2})^T >,$$ t1= 48 h, t2 = 24 h

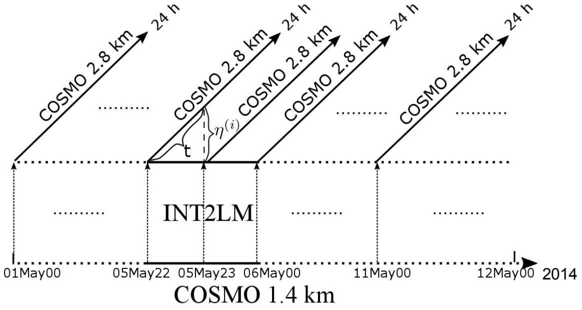
  $\boldsymbol{B}^{\frac{1}{2}}$ adapted to finer resolution of regional **COSMO** and save as $\tilde{\boldsymbol{B}}^{\frac{1}{2}}$

  $$\boldsymbol{\eta}^{(i)} = \tilde{\boldsymbol{B}}^{\frac{1}{2}} \gamma \ , \ \gamma \in \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$$

  - provide synoptic uncertainty, called “**LAN**” (**L**arge-scale **A**dditive **N**oise)
  - u, v, T, qv and p perturbed
  - α = 0.1



- Random draws from set of model truncation error (Zeng et al. 2019a, JAMES)

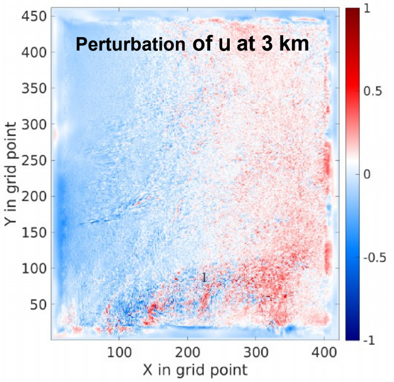


$$\boldsymbol{\eta} = \mathcal{T}\{\mathcal{M}^H[\boldsymbol{x}^H(t_k - t)]\} - \mathcal{M}^L\{\mathcal{T}[\boldsymbol{x}^H(t_k - t)]\},$$

$\mathcal{M}^H$ &$\mathcal{M}^L$: high & low resolution model

$\mathcal{T}$: interpolation operator

t_k random, t = 1 h

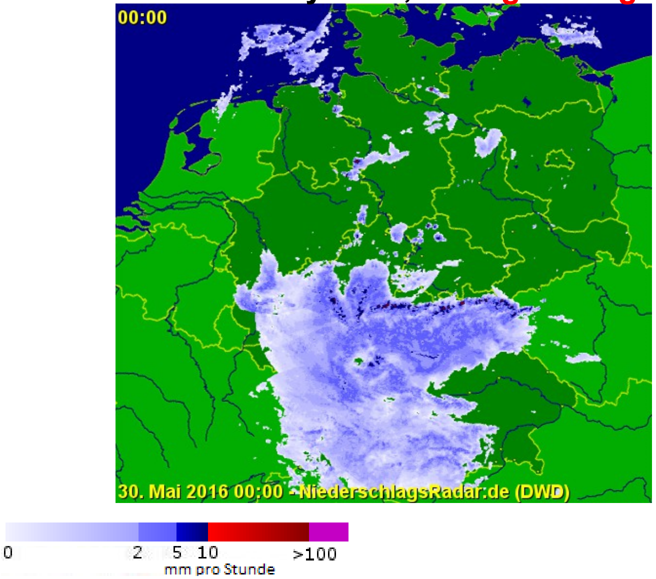


- provide unresolved uncertainty, called “**SAN**” (**S**mall-scale **A**dditive **N**oise)
- u, v, w, T and qv perturbed
- α = 1.25

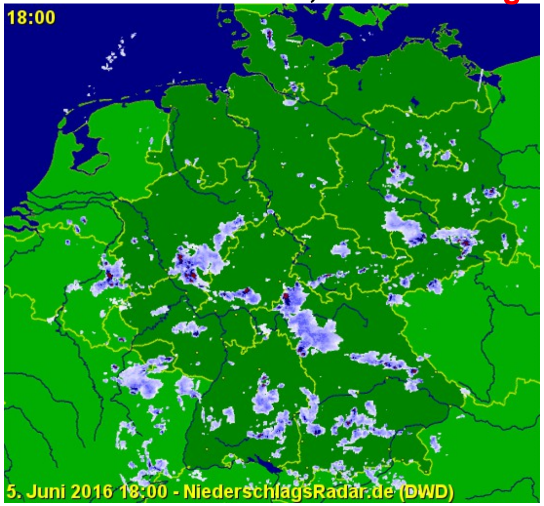
Experimental design

Weather situation: a two-week period (27 May – 9 June 2016) over Germany

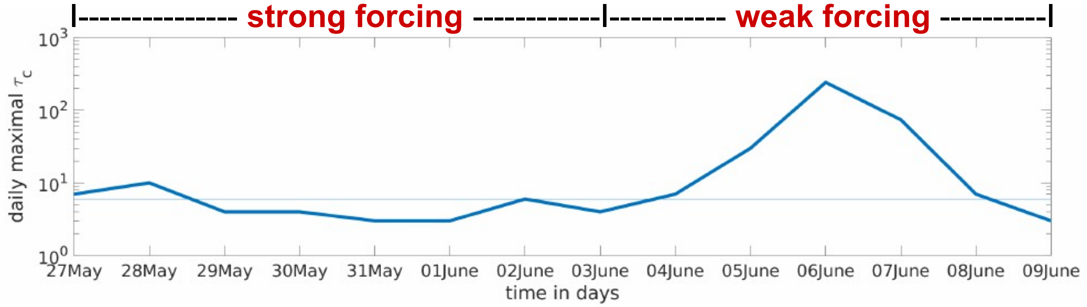
00:00 30 May 2018, strong forcing



18:00 05 June 2018, weak forcing



τc :convective adjustment time-scale [h] (Keil et al. 2014)



If max. $\tau c \leq 6$ h, strong forcing;
If max. $\tau c > 6$ h, weak forcing

Objective verification skill scores:

- **F**ractions **S**kill **S**core (**FSS**, Roberts & Lean 2008, MWR) :

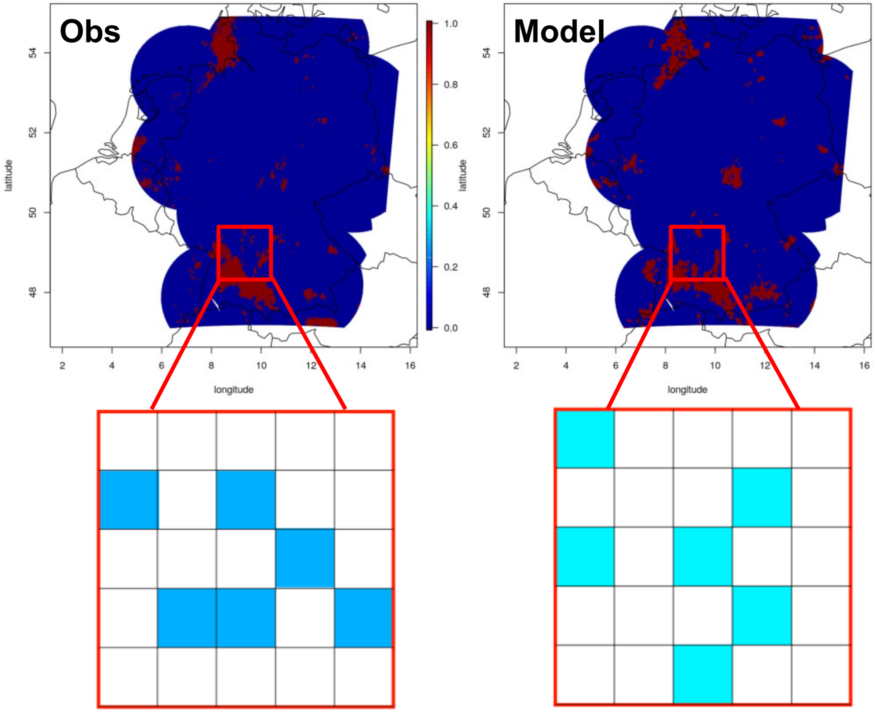
  for a threshold value, compare forecast fractions with observed fractions over different scales

$$FSS = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}(P_{fcst} - P_{obs})^2}{\frac{1}{N}\sum_{i=1}^{N}P_{fcst}^2 + \frac{1}{N}\sum_{i=1}^{N}P_{obs}^2}$$

  [0,1], perfect score 1

Example: precip. rate = 5 mm/h and scale of 5 grid points

If >= 5 mm/h, = 1; If < 5 mm/h, = 0



Pobs = 6/25 = 0.24

Pfcst = 6/25 = 0.24

Experimental design

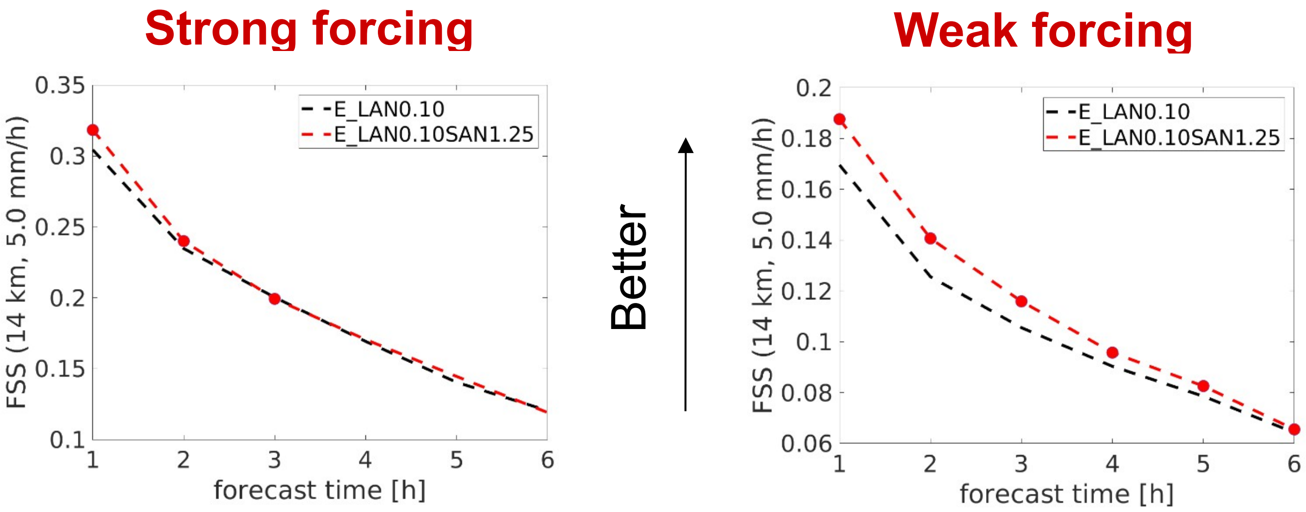
Set-up:

Exp	LAN (α = 0.1)	SAN (α = 1.25)
E_LAN0.10	✓	✗
E_LAN0.10SAN1.25	✓	✓

Observations: SYNOP, **AIREP**, **TEMP**, **PROF** + radar reflectivity
Data assim. Scheme: LETKF (also for radar reflectivity using forward operator **EMVORADO** (Zeng et al. 2014, JTECH; Zeng et al. 2016, QJRMS)
Assimilation window: one hour
Size of ensemble: 40 members, and 20 members are used for 6-h ensemble forecasts, initiated at 10, 11, …, 18:00 UTC
Localization: adaptive horizontal localization for conventional data, constant horizontal localization (16 km) for reflectivity
Observation error: adaptive for conventional data, constant (10 dBZ) for reflectivity
Period: 00:00 27 May – 00:00 03 June 2016 (strong forcing)
00:00 03 June – 00:00 10 June 2016 (weak forcing)

yuefei.zeng@lmu.de

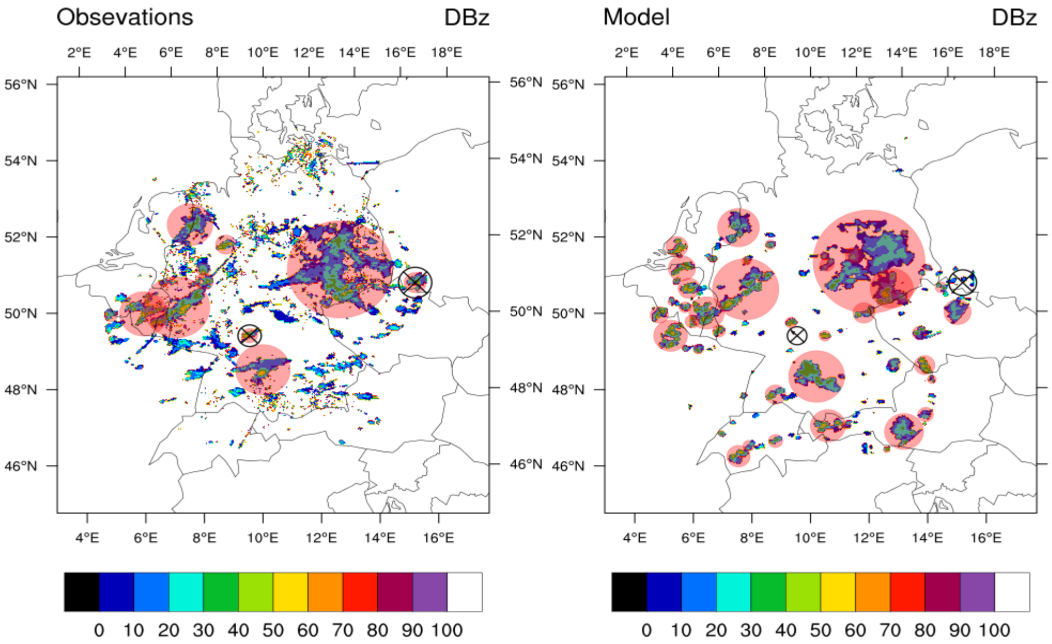
Results



E_LAN0.10 is **reference run**. Differences with statistical significance indicated with dots.

Alternatives for representing subgrid-scale model error

- New warm-bubble technique developed at DWD (Zeng et al. 2019b, MWR):



- **P**hysically based **S**tochastic **P**erturbations scheme (**PSP**, Kober and Craig 2016)

$$\left(\frac{\partial \Phi}{\partial t}\right)_{\text{total}} = \left(\frac{\partial \Phi}{\partial t}\right)_{\text{param}} + \alpha_{\text{tuning}}\, \eta \,\frac{1}{\tau_{\text{eddy}}} \frac{l_{\text{eddy}}}{\Delta x_{\text{eff}}} \sqrt{\overline{\Phi'^2}}$$

$\Phi \in T, q_v, w$; $\tau_{\text{eddy}} = 10$ minutes; $\Delta x_{\text{eff}} = 5\Delta x$; $\alpha_{\text{tuning}} = 7.2$; $l_{\text{eddy}} = 1$ km

$\sqrt{\overline{\Phi'^2}}$ is the subgrid standard deviation; η is a two-dimensional random field

Alternatives for representing subgrid-scale model error

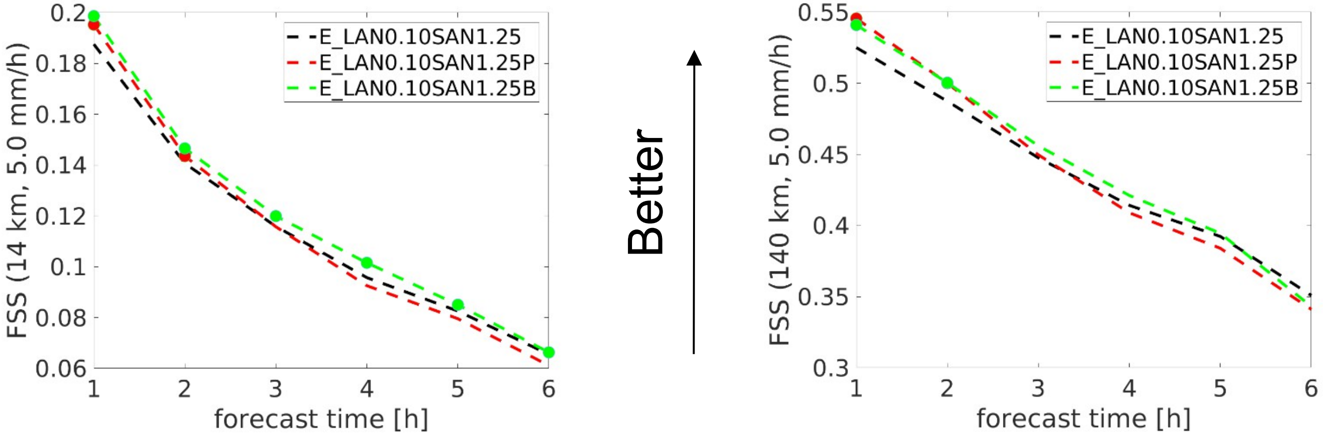
(Zeng et al. 2019b, MWR):

Set-up:

Exp	LAN (α = 0.1)	SAN (α =1.25)	PSP	Warm bubble
E_LAN0.10SAN1.25	✓	✓	✗	✗
E_LAN0.10SAN1.25P	✓	✓	✓	✗
E_LAN0.10SAN1.25B	✓	✓	✗	✓

Results

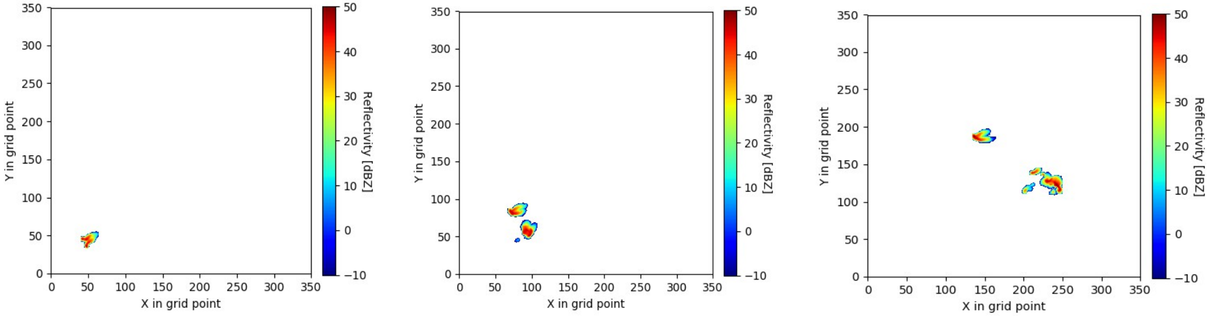
Weak forcing



E_LAN0.10SAN1.25 is **reference run**. Differences with statistical significance indicated with dots.

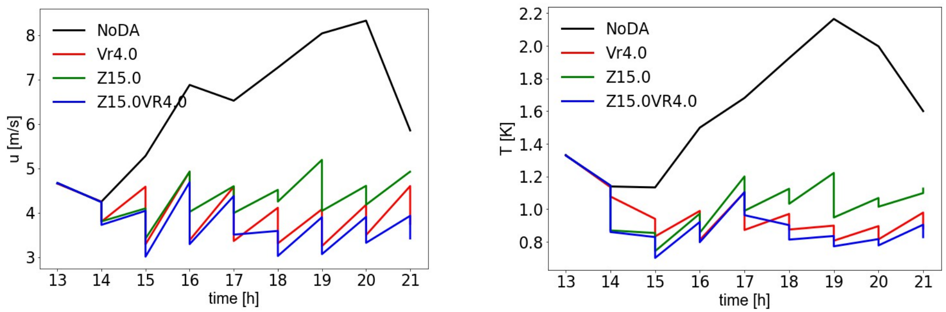
Idealized setup for radar data assimilation

Supercell simulation

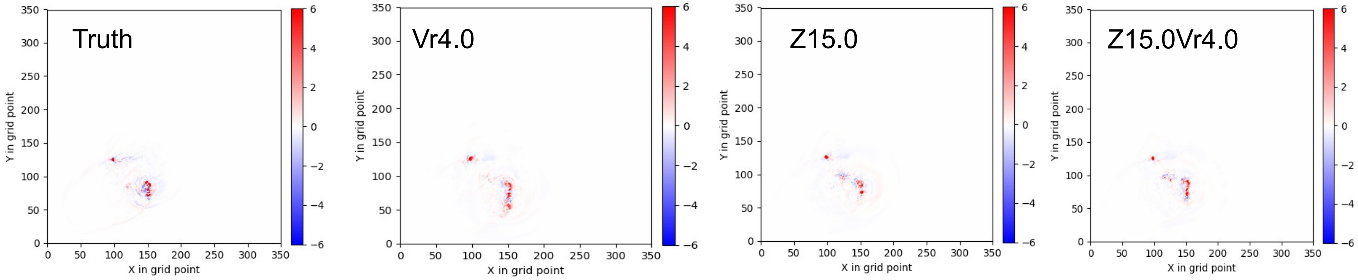


Assimilation results

RMSE



w at 10 km



Summary and future work

- The combination of large- and small-scale additive noise outperforms large-scale additive noise alone in both strong and weak forcing situations. The improvements are especially significant in weak forcing situation

- Further improvements can be achieved by using the PSP scheme or warm bubble technique

- Assimilation of radar radial winds and reflectivity exhibits possitive impacts in idealized setup

Future work:

- Representation of model error in microphysics and obs. error in polarimetric radar measurement
- Sensitivity to errors in observation operator
- Transition to ICON

yuefei.zeng@lmu.de

Selected literature

Zeng, Y., T. Janjic, A. de Lozar, S. Rasp, U. Blahak, A. Seifert and G. C. Craig (2019b), Comparison of methods accounting for subgrid-scale model error in convective-scale data assimilation, Mon. Wea. Review, under review

Zeng, Y., T. Janjic, M. Sommer, A. de Lozar, U. Blahak and A. Seifert (2019a), Representation of model error in convective-scale data assimilation: additive noise based on model truncation error, J. Adv. Model. Earth Syst., https://doi.org/10.1029/2018MS001546

Zeng, Y., T. Janjic, A. de Lozar, U. Blahak, H. Reich, C. Keil, and A. Seifert (2018),Representation of model error in convective-scale data assimilation: additive noise, relaxation methods and combinations, J. Adv. Model. Earth Syst., 10, 2889-2911

Zeng, Y., T. Janjic, Y. Ruckstuhle, and M. Verlaan (2017), Ensemble-type Kalman filter algorithm conserving mass, total energy and enstrophy, Quart. J. Roy. Meteor. Soc., 143, 2902-2914.

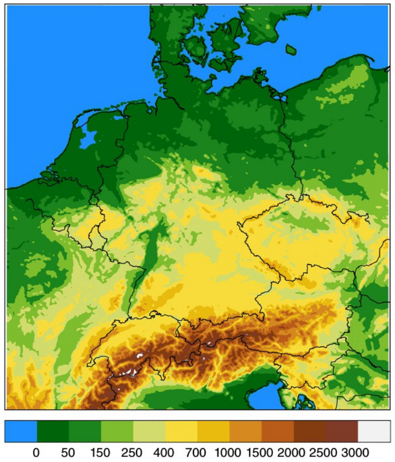
Zeng, Y., U. Blahak, and D. Jerger (2016), An efficient modular volume-scanning radar forward operator for NWP models: description and coupling to the COSMO model, Quart. J. Roy. Meteor. Soc., 142, 3234-3256.

Zeng, Y., U. Blahak, M. Neuper, and D. Jerger (2014), Radar Beam Tracing Methods Based on Atmospheric Refractive Index, J. Atmos. Ocean. Tech, 31, 2650-2670.

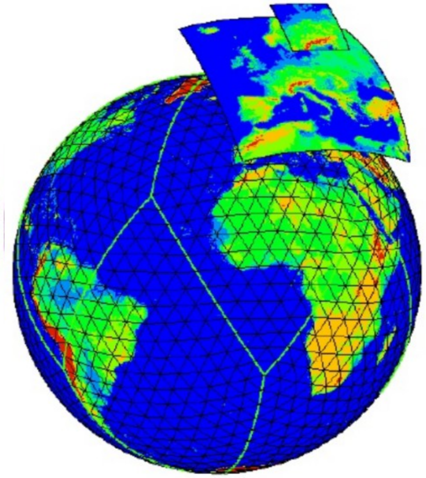
yuefei.zeng@lmu.de

Background information

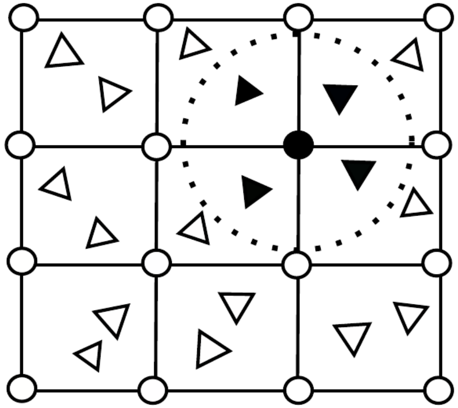
- **K**ilometre-scale **EN**semble **D**ata **A**ssimilation (**KENDA**) system at DWD since March 2017 (Schraff et al. 2016, QJRMS) in operational mode:
  - Convection-permitting model: **CO**nsortium for **S**mall-scale **MO**deling (**COSMO**, Baldauf et al. 2011, MWR) with horizontal resolution 2.8 km
  - Lateral boundary conditions: **ICO**sahedral **N**onhydrostatic (**ICON**, Zängl et al. 2015, QJRMS)
  - Data assim. scheme: **L**ocal **E**nsemble **T**ransform **K**alman **F**ilter (**LETKF**, Hunt et al. 2007, Phy.D)
  - Obs. assim.: **SYNOP**, **AIREP**, **TEMP** , **PROF**, **MODES**; radar reflectivity



Orography COSMO-DE domain (m)



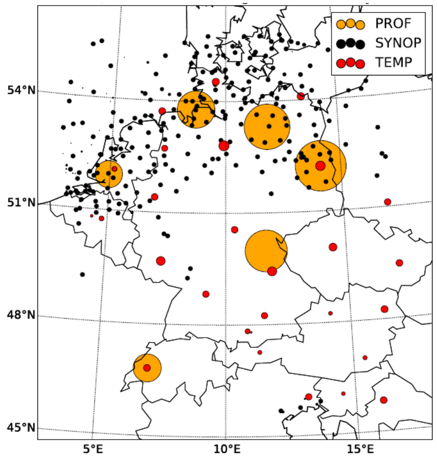
ICON grid structure with nested domains



LETKF : ● Analysis grid point

▲ Observation

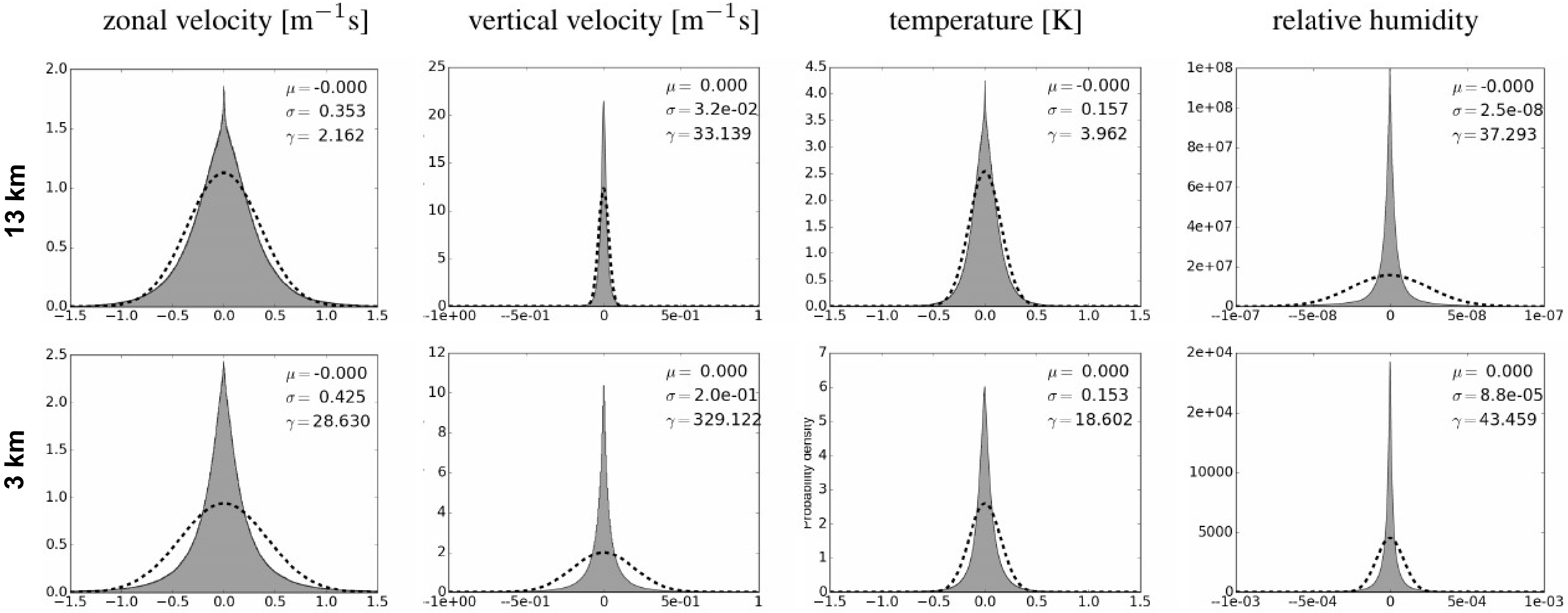
.... Localization radius



Obs. assimilated

Approaches to represent model error

Histogram of model truncation error samples for SAN



…………: Gaussian distribution

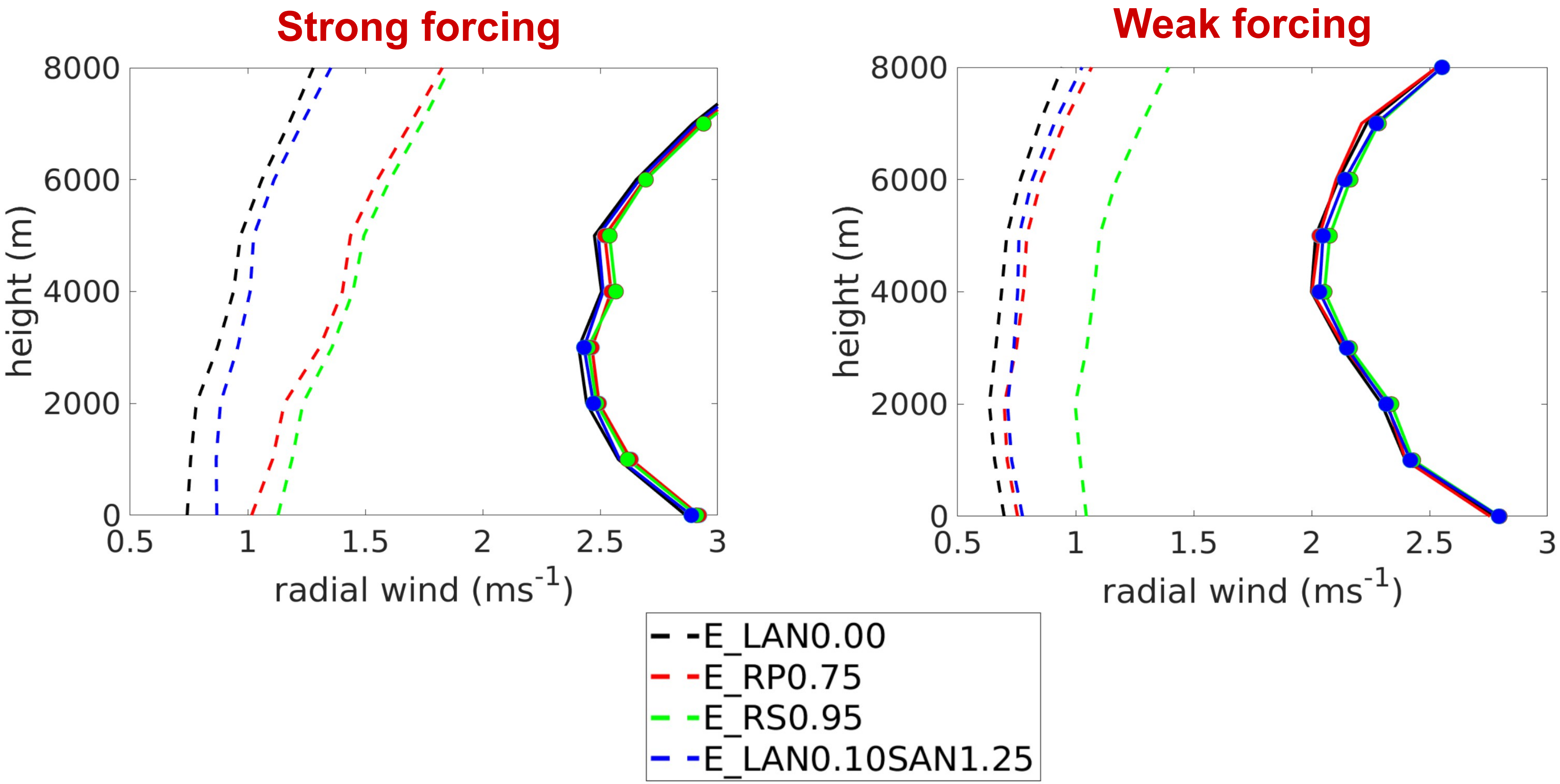
μ : mean

σ : standard deviation

γ : kurtosis, γ = 0 if Gaussian distribution

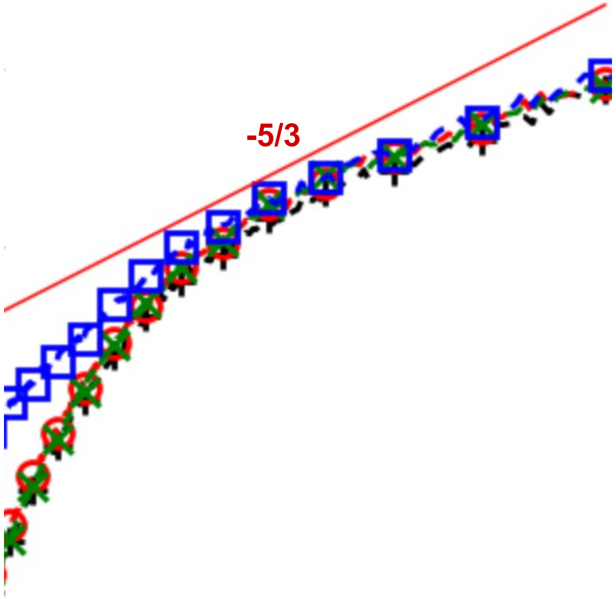
Results

Spread (dashed lines) and RMSE (solid lines) of background ensemble

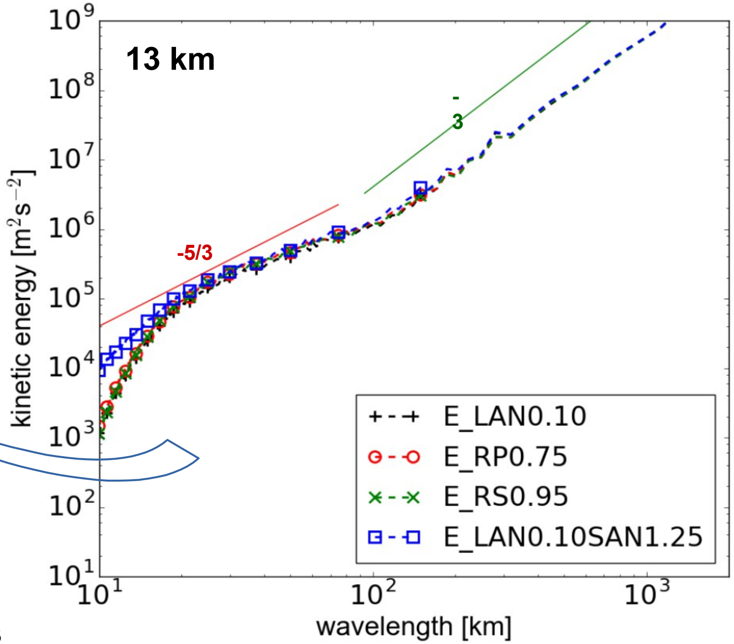


E_LAN0.10 is **reference run**. Differences with statistical significance indicated with dots.

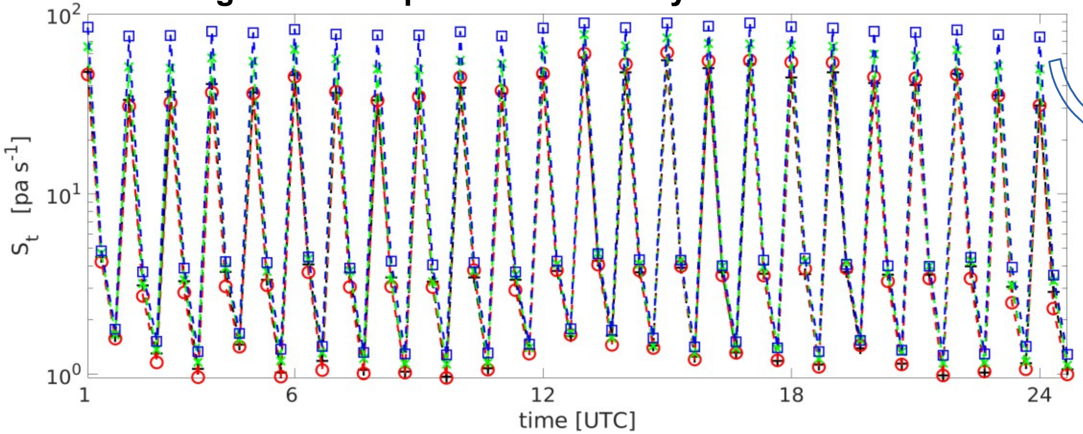
Results

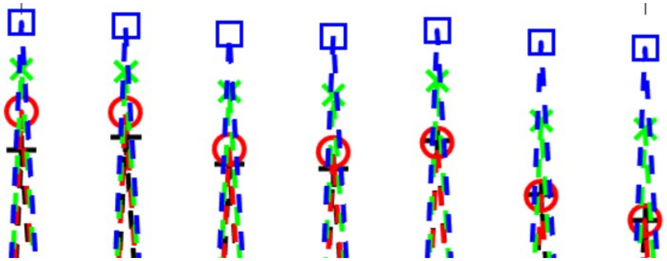


Averaged kinetic energy spectra of analysis ensemble members



Averaged surface pressure tendency of ensemble members

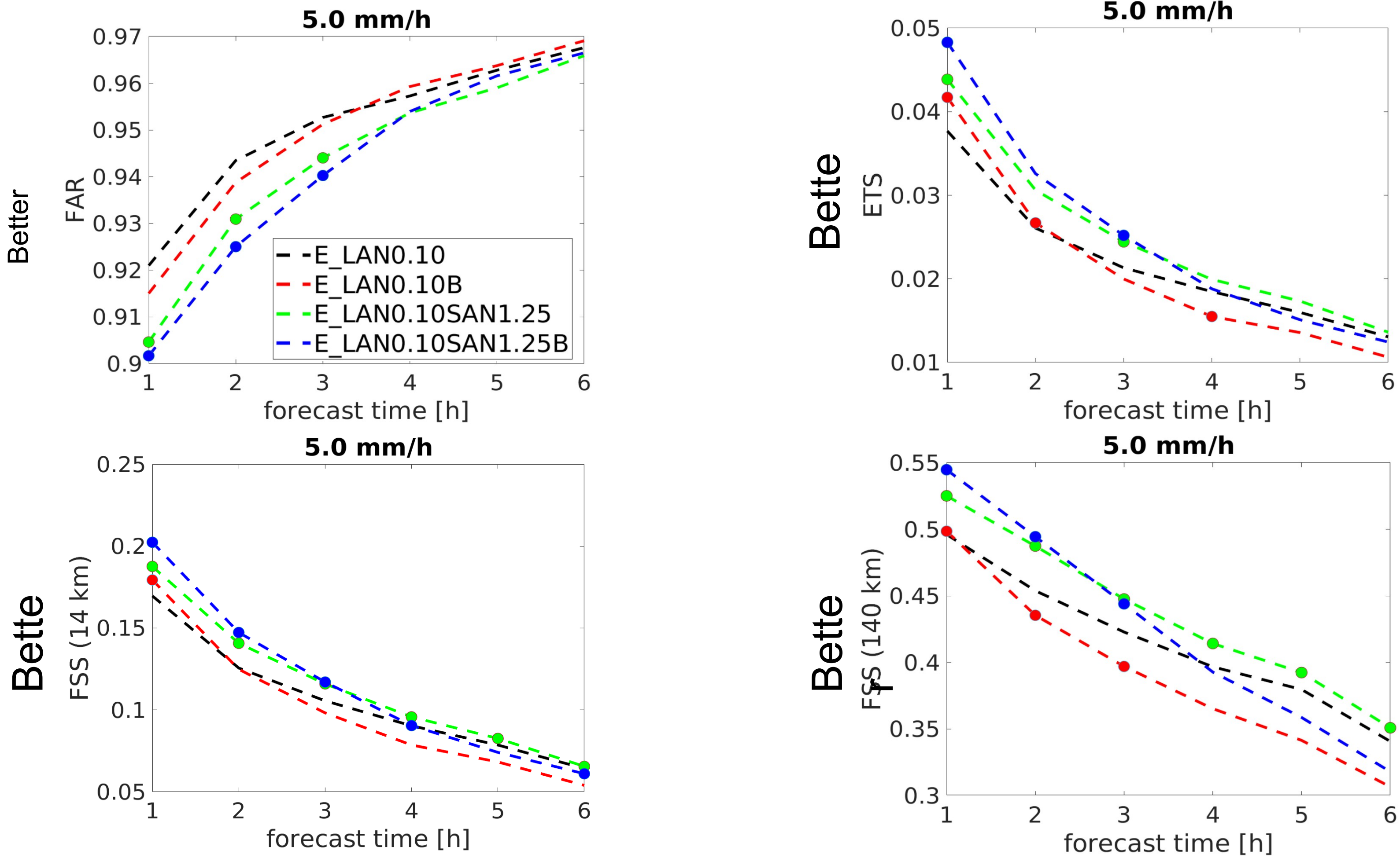




yuefei.zeng@lmu.de

Results

Weak forcing



E_LAN0.10 is **reference run**. Differences with statistical significance indicated with dots.

yuefei.zeng@lmu.de

Approaches to represent model error

1. Relaxation methods:

- **R**elaxation **T**o **P**rior **P**erturbations (**RTPP**, Zhang et al. 2004)

$$\mathbf{X}^a \leftarrow (1-\alpha_p)\mathbf{X}^a + \alpha_p \mathbf{X}^b$$

Ensemble perturbations:

$$\mathbf{X}^a = \left[\mathbf{x}^{a(1)} - \overline{\mathbf{x}}^a, \mathbf{x}^{a(2)} - \overline{\mathbf{x}}^a, \ldots, \mathbf{x}^{a(N)} - \overline{\mathbf{x}}^a\right] \qquad \mathbf{X}^b = \left[\mathbf{x}^{b(1)} - \overline{\mathbf{x}}^b, \mathbf{x}^{b(2)} - \overline{\mathbf{x}}^b, \ldots, \mathbf{x}^{b(N)} - \overline{\mathbf{x}}^b\right]$$

$\alpha_p = 0.75$

- **R**elaxation **T**o **P**rior **S**pread (**RTPS**, Whitaker and Hamill 2012)

$$\sigma^a \leftarrow (1-\alpha_s)\sigma^a + \alpha_s \sigma^b \iff \mathbf{X}^a \leftarrow \left(\alpha_s \frac{\sigma^b - \sigma^a}{\sigma^a} + 1\right)\mathbf{X}^a, \quad \alpha_s = 0.95$$

Ad hoc, usually used to account for unknown model error !

Experimental design

Set-up:

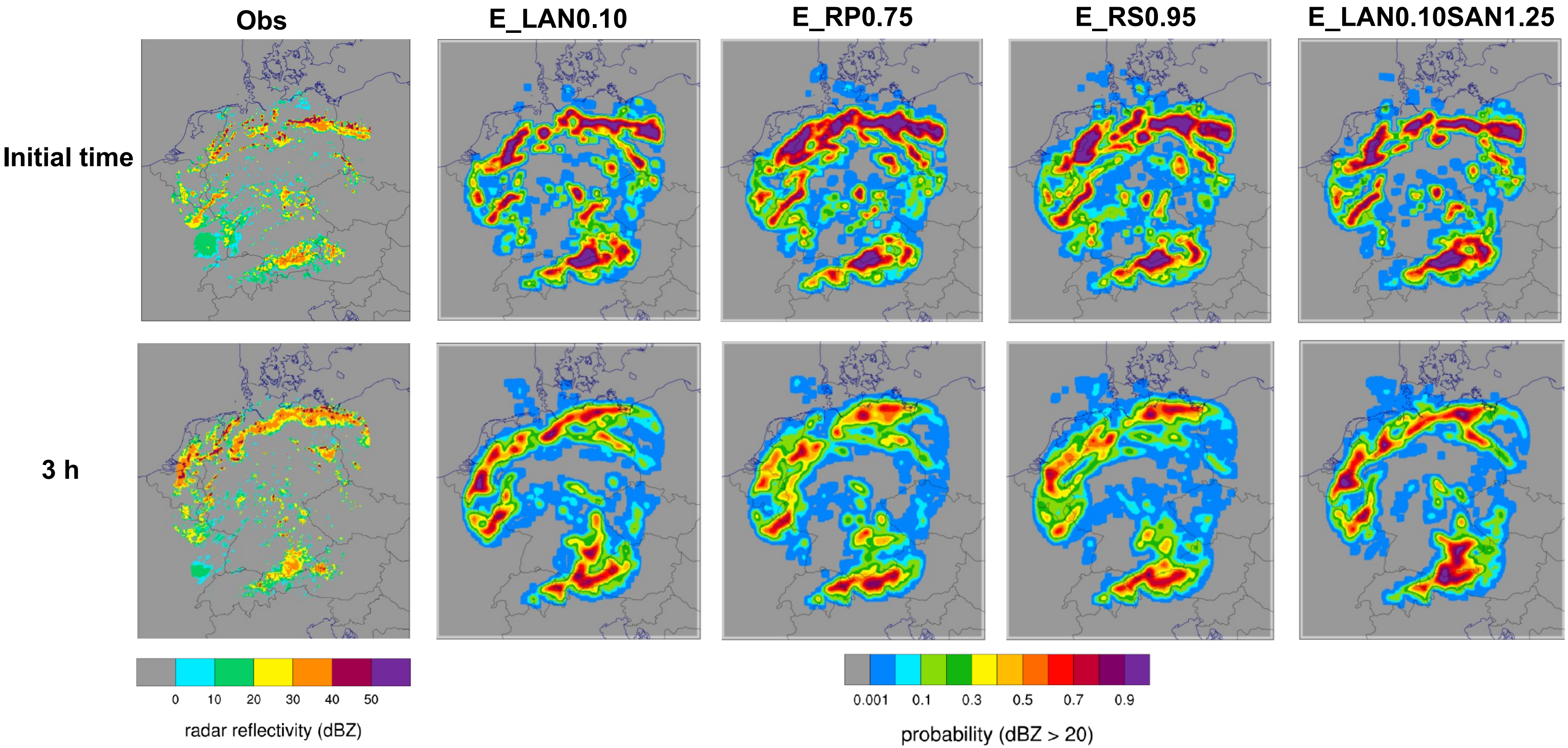
Exp	LAN (α = 0.1)	RTPP (α = 0.75)	RTPS (α = 0.95)	SAN (α = 1.25)
E_LAN0.10	✓	×	×	×
E_RP0.75	×	✓	×	×
E_RS0.95	×	×	✓	×
E_LAN0.10SAN1.25	✓	×	×	✓

Observations: SYNOP, **AIREP**, **TEMP**, **PROF** + radar reflectivity
Data assim. Scheme: LETKF (also for radar reflectivity using forward operator **EMVORADO** (Zeng et al. 2014, JTECH; Zeng et al. 2016, QJRMS)
Assimilation window: one hour
Size of ensemble: 40 members, and 20 members are used for 6-h ensemble forecasts, initiated at 10, 11, …, 18:00 UTC
Localization: adaptive horizontal localization for conventional data, constant horizontal localization (16 km) for reflectivity
Observation error: adaptive for conventional data, constant (10 dBZ) for reflectivity
Period: 00:00 27 May – 00:00 03 June 2016 (strong forcing)
00:00 03 June – 00:00 10 June 2016 (weak forcing)

yuefei.zeng@lmu.de

Results

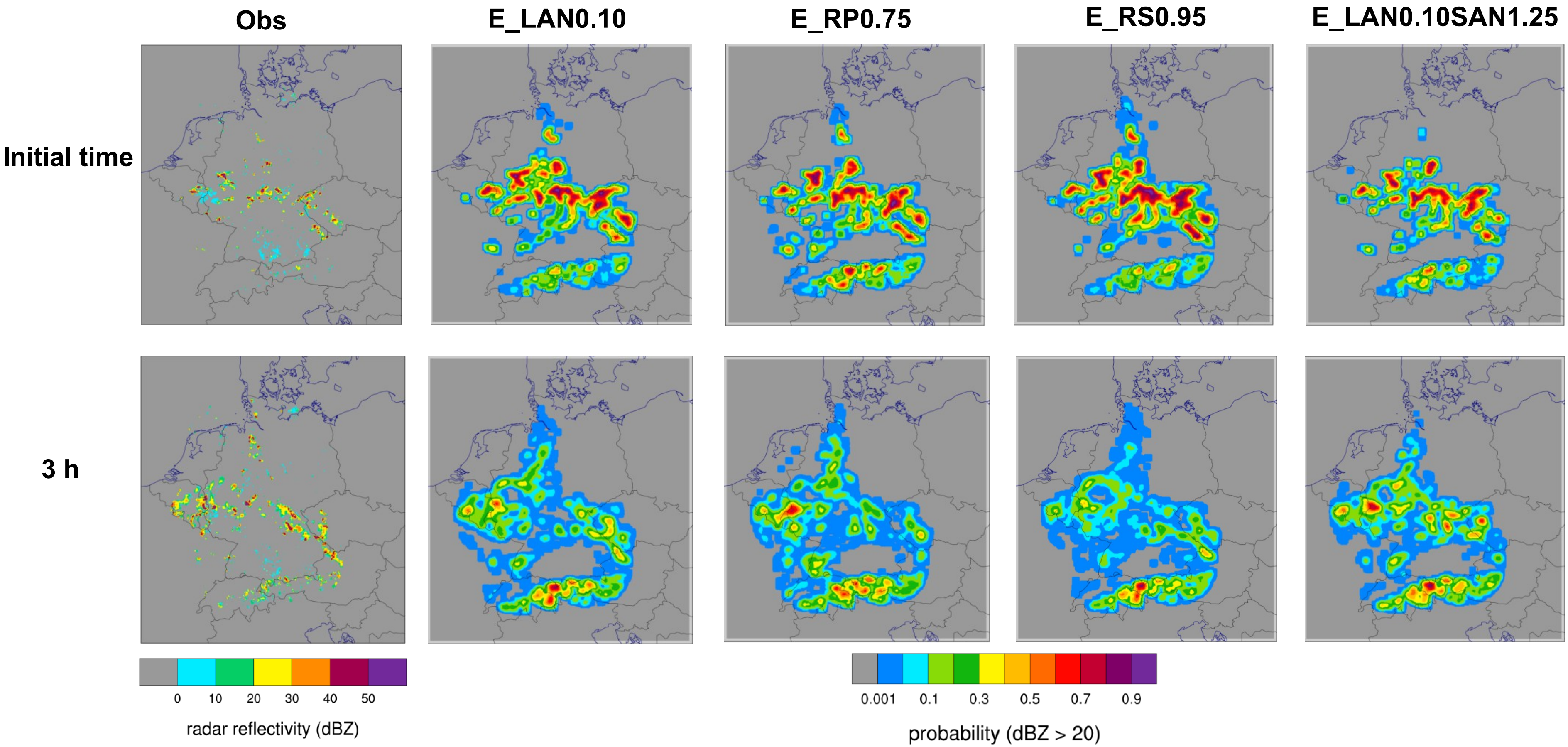
Reflectivity composite at initial time (14:00 30 May, 2016, strong forcing) & 3 h forecast



Probability: How much percent of ensemble members exceed 20 dBZ

Results

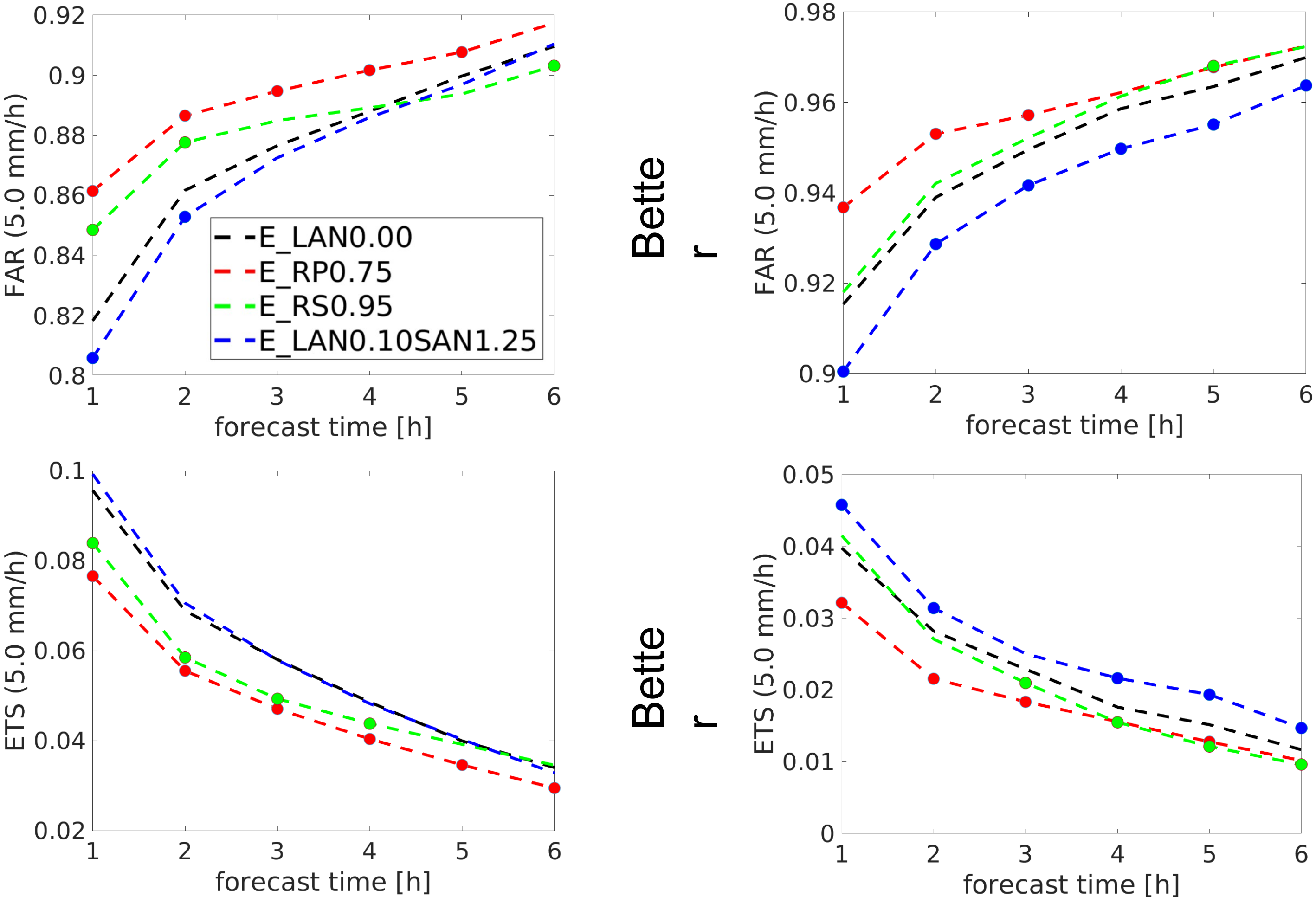
Reflectivity composite at initial time (12:00 06 June, 2016, weak forcing) & 3 h forecast



Probability: How much percent of ensemble members exceed 20 dBZ

yuefei.zeng@lmu.de

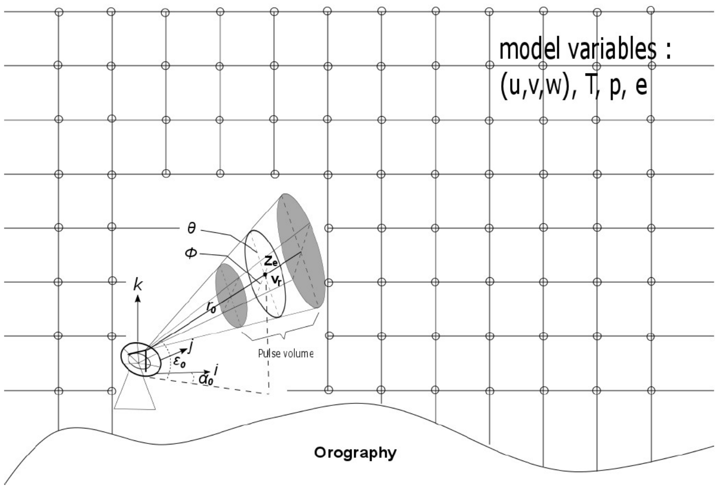
Results


Strong forcing

Weak forcing

E_LAN0.10 is **reference run**. Differences with statistical significance indicated with dots.

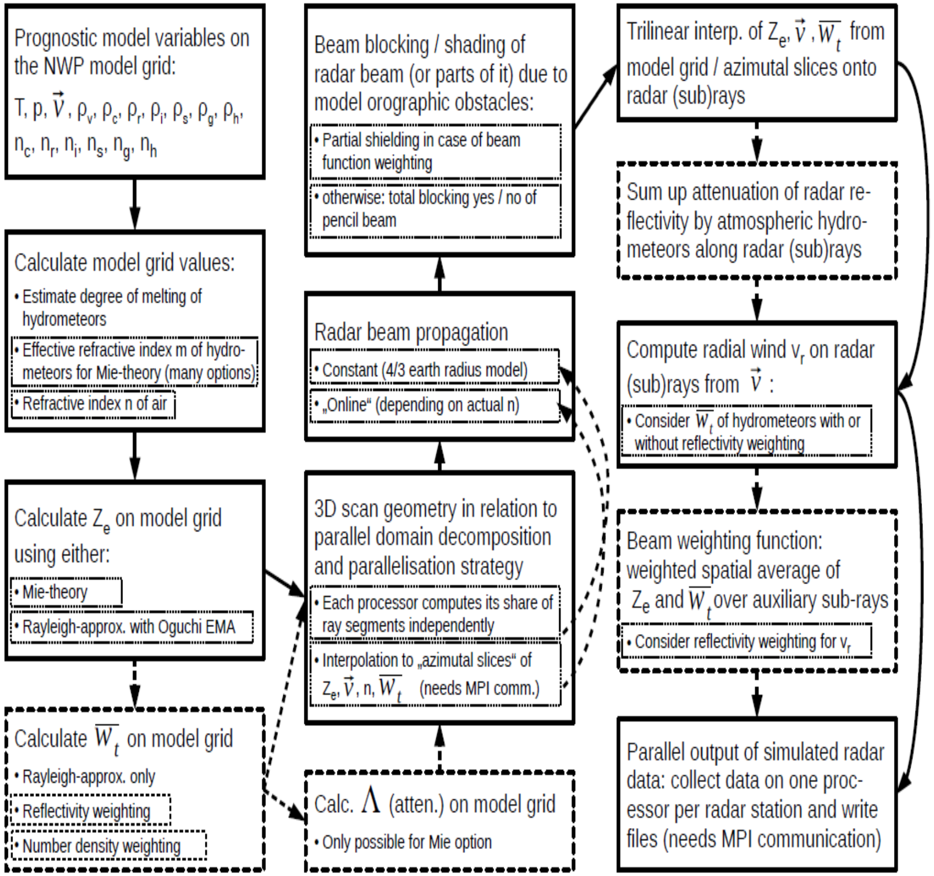
Development of radar observation operator



radar scan geometry ≠ model grid and radar observations ≠ model variables

Idea: $(u, v, w), T, p, e) \Rightarrow Z_e$ or v_r

Efficient Modular Volume RADar Operator (EMVORADO, pre-operational at DWD)



Prognostic model variables on the NWP model grid:

T, p, $\vec{v}$, ρ_v, ρ_c, ρ_r, ρ_i, ρ_s, ρ_g, ρ_h, n_c, n_r, n_i, n_s, n_g, n_h

Calculate model grid values:
- Estimate degree of melting of hydrometeors
- Effective refractive index m of hydrometeors for Mie-theory (many options)
- Refractive index n of air

Calculate Z_e on model grid using either:
- Mie-theory
- Rayleigh-approx. with Oguchi EMA

Calculate $\overline{W}_t$ on model grid
- Rayleigh-approx. only
- Reflectivity weighting
- Number density weighting

Beam blocking / shading of radar beam (or parts of it) due to model orographic obstacles:
- Partial shielding in case of beam function weighting
- otherwise: total blocking yes / no of pencil beam

Radar beam propagation
- Constant (4/3 earth radius model)
- „Online" (depending on actual n)

3D scan geometry in relation to parallel domain decomposition and parallelisation strategy
- Each processor computes its share of ray segments independently
- Interpolation to „azimutal slices" of Z_e, $\vec{v}$, n, $\overline{W}_t$ (needs MPI comm.)

Calc. Λ (atten.) on model grid
- Only possible for Mie option

Trilinear interp. of Z_e, $\vec{v}$, $\overline{W}_t$ from model grid / azimutal slices onto radar (sub)rays

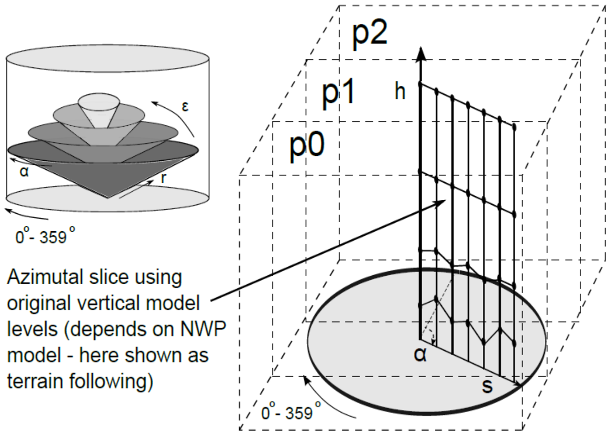
Sum up attenuation of radar reflectivity by atmospheric hydrometeors along radar (sub)rays

Compute radial wind v_r on radar (sub)rays from $\vec{v}$:
- Consider $\overline{W}_t$ of hydrometeors with or without reflectivity weighting

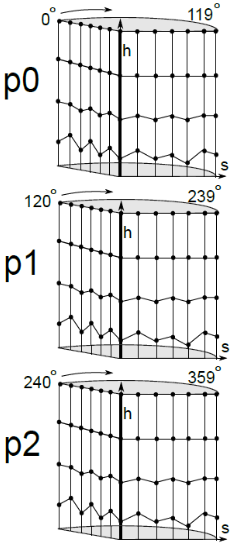
Beam weighting function: weighted spatial average of Z_e and $\overline{W}_t$ over auxiliary sub-rays
- Consider reflectivity weighting for v_r

Parallel output of simulated radar data: collect data on one processor per radar station and write files (needs MPI communication)

Zeng et al. 2014, JTECH;
Zeng et al. 2016, QJRMS



Azimutal slice using original vertical model levels (depends on NWP model - here shown as terrain following)



yuefei.zeng@lmu.de

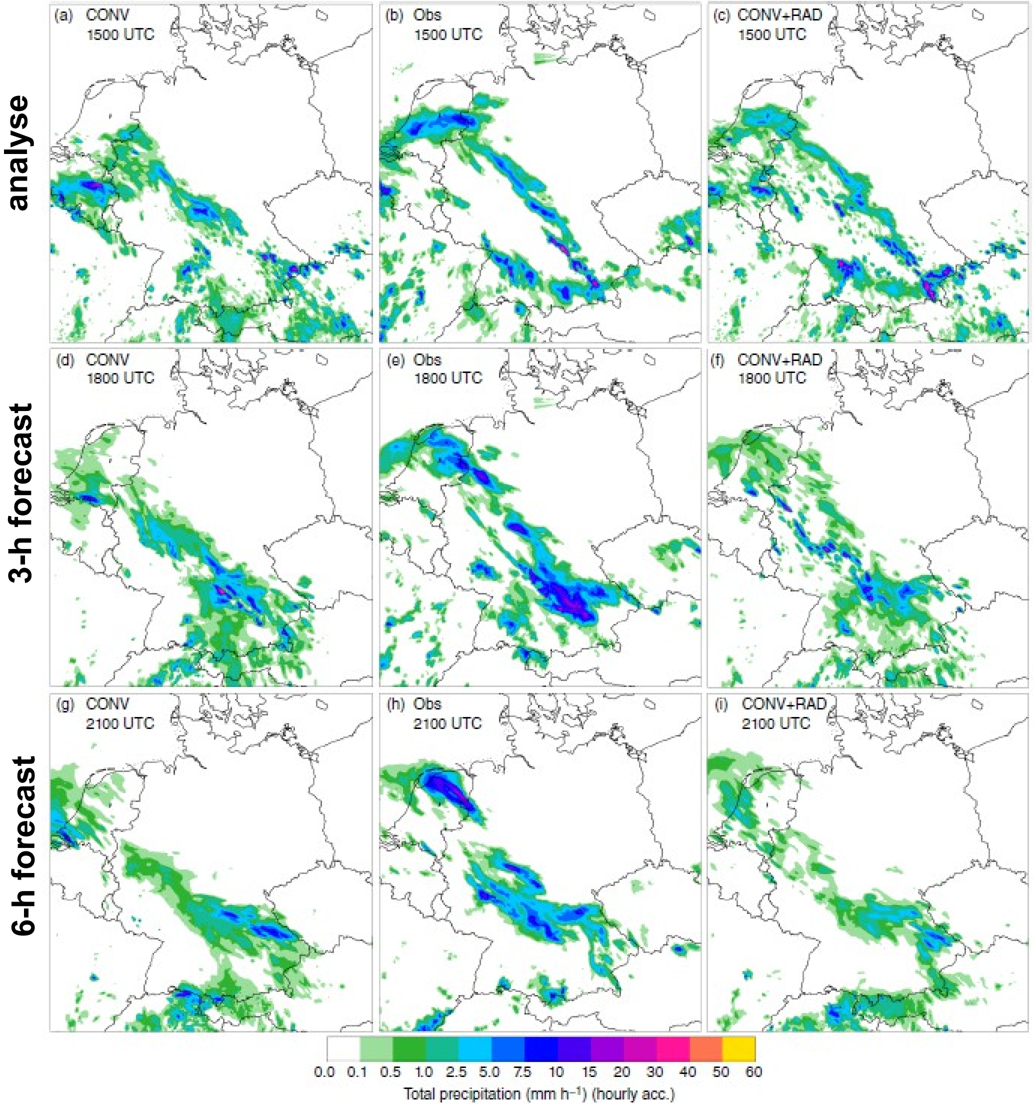
First experiments of radar reflectivity data assimilation

(Bick et al. 2016, QJRMS)

Set-up:

Exp	Conventional obs.	Radar reflectivity
CONV	✓	✗
CONV+RAD	✓	✓

Conventional obs.: SYNOP, AIREP, TEMP , PROF



yuefei.zeng@lmu.de

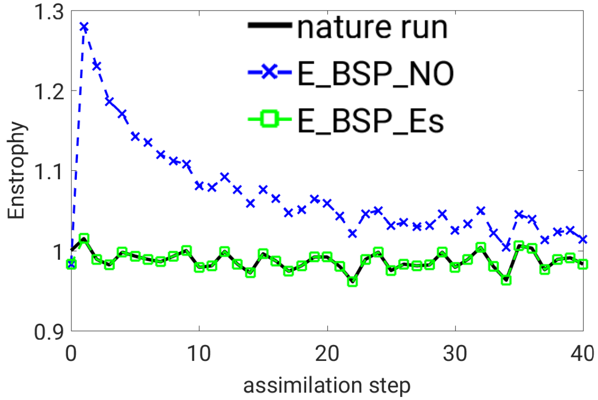
Conservation properties of data assimilation schemes

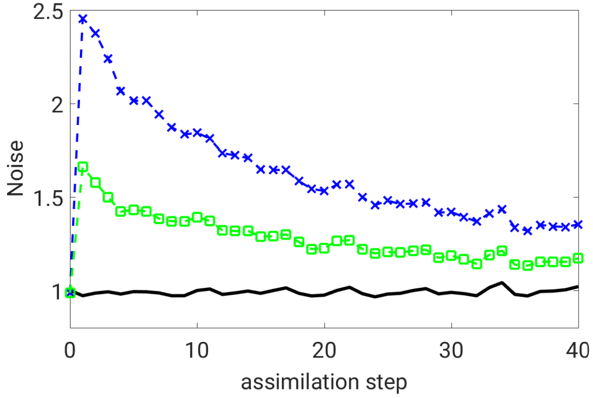
Idealized experiments with shallow water model

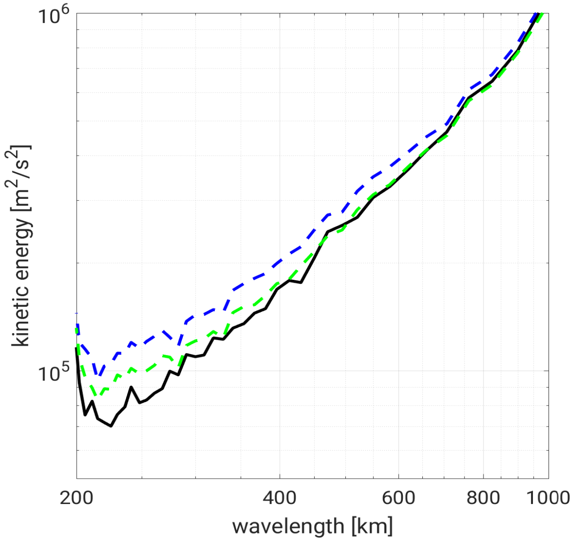
Set-up:

Exp	constraint of enstrophy conservation
E_BSP_NO	✗
E_BSP_Es	✓

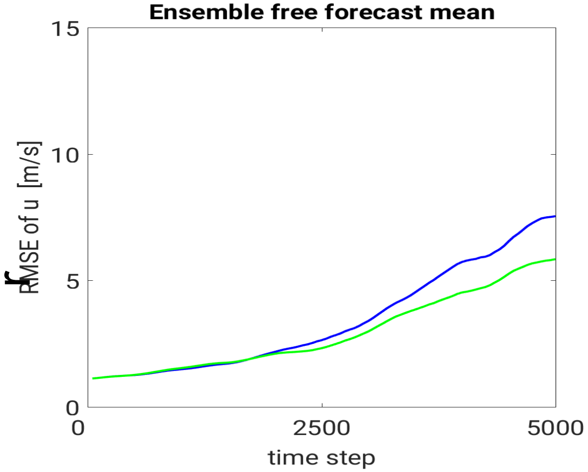
Zeng and Janjic 2016, QJRMS; Zeng et al. 2017, QJRMS







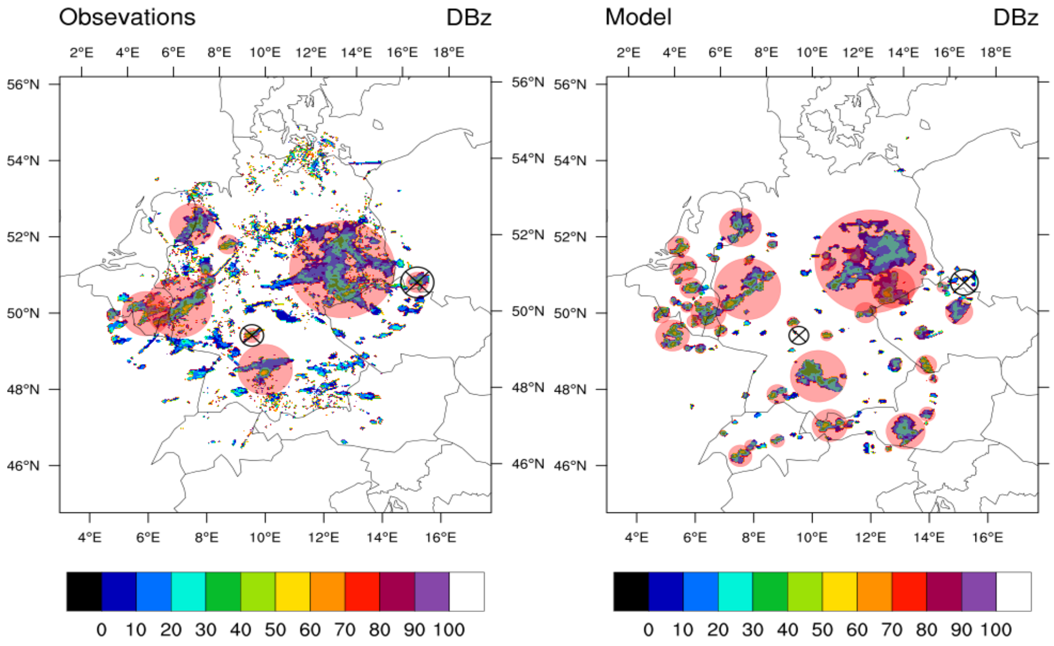
Better



yuefei.zeng@lmu.de

Alternatives for representing subgrid-scale model error

- New warm-bubble technique developed at DWD (Zeng et al. 2019b, MWR):



- **P**hysically based **S**tochastic **P**erturbations scheme (**PSP**, Kober and Craig 2016)

$$\left(\frac{\partial \Phi}{\partial t}\right)_{\text{total}} = \left(\frac{\partial \Phi}{\partial t}\right)_{\text{param}} + \alpha_{\text{tuning}}\,\eta\,\frac{1}{\tau_{\text{eddy}}}\frac{l_{\text{eddy}}}{\Delta x_{\text{eff}}}\sqrt{\overline{\Phi'^2}}$$

$\Phi \in T, q_v, w$; $\tau_{\text{eddy}} = 10$ minutes; $\Delta x_{\text{eff}} = 5\Delta x$; $\alpha_{\text{tuning}} = 7.2$; $l_{\text{eddy}} = 1$ km

$\sqrt{\overline{\Phi'^2}}$ is the subgrid standard deviation; η is a two-dimensional random field

Reference

Baldauf, M., A. Seifert, J. Förstner, D. Majewski, R. M., and T. Reinhardt (2011), Operational convective-scale numerical weather prediction with the COSMO model: description and sensitivities. Mon. Wea. Rev., 139, 3887--3905.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh (2007), Efficient data assimilation for Spatiotemporal Chaos: a Local Ensemble Transform Kalman Filter, Physica D: Nonlinear Phenomena, 230, 112--126

Keil, C., F. Heinlein, and G. C. Craig (2014), The convective adjustment time-scale as indicator of predictability of convective precipitation, Quart. J. Roy. Meteor. Soc., 140, 480--490.

Parrish, D. F., and J. C. John C. Derber (1992), The National Meteorological Center's Spectral Statistical-Interpolation Analysis System, Mon. Wea. Rev, 120, 1747--1763

Robert, N., and H. Lean (2008), Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events, Mon. Weather. Rev., 136, 78--96.

Whitaker, J. S., and T. M. Hamill (2012), Evaluating methods to account for system errors in ensemble data assimilation, Mon. Wea. Rev., 140(9), 3078--3089.

Wilks, D. S. (2006), Statistical Methods in the Atmospheric Sciences, Academic Press, New York.

Zängl, G., D. Reinert, P. Rpodas, and M. Baldauf (2015), The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core, Quar. J. Roy, Meteorol. Soc., 141, 563--579.

Zhang, F., C. Snyder, and J. Sun (2004), Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman Filter, Mon. Wea. Rev., 132(5), 1238--1253.

Zeng, Y., U. Blahak, and D. Jerger (2016), An efficient modular volume-scanning radar forward operator for NWP models: description and coupling to the COSMO model, Quart. J. Roy. Meteor. Soc., 142, 3234--3256.

Zeng, Y., T. Janjic, A. de Lozar, U. Blahak, H. Reich, C. Keil, and A. Seifert (2018a),Representation of model error in convective-scale data assimilation: additive noise, relaxation methods and combinations, J. Adv. Model. Earth Syst., DOI:10.1029/2018ms001375

Zeng, Y., T. Janjic, M. Sommer, A. de Lozar, U. Blahak and A. Seifert (2018b), Representation of model error in convective-scale data assimilation: additive noise based on model truncation error, J. Adv. Model. Earth Syst., under review