Assimilation of 3D radar reflectivities with an ensemble Kalman filter on the convective scale

RMetS
Royal Meteorological Society

Assimilation of 3D radar reflectivities with an ensemble Kalman filter on the convective scale

T. Bick,[a,b,c*] C. Simmer,[a,b,d] S. Trömel,[a,b] K. Wapler,[a,c] H.-J. Hendricks Franssen,[d,e] K. Stephan,[c] U. Blahak,[c] C. Schraff,[c] H. Reich,[c] Y. Zeng[c,f] and R. Potthast,[c,g]

[a]*Atmospheric Dynamics and Predictability Branch, Hans Ertel Centre for Weather Research, Bonn, Germany*
[b]*Meteorologisches Institut, Universität Bonn, Germany*
[c]*Deutscher Wetterdienst, Offenbach, Germany*
[d]*Centre for High-Performance Scientific Computing in Terrestrial Systems, Jülich, Germany*
[e]*Agrosphere (IBG-3), Institute of Bio- and Geosciences, Forschungszentrum Jülich, Germany*
[f]*Meteorologisches Institut der Ludwig-Maximilians-Universität, München, Germany*
[g]*Department of Mathematics and Statistics, University of Reading, UK*

*Correspondence to: T. Bick, Meteorologisches Institut, Universität Bonn, Auf dem Hügel 20, 53121 Bonn, Germany. E-mail: thbick@uni-bonn.de

An ensemble data assimilation system for 3D radar reflectivity data is introduced for the convection-permitting numerical weather prediction model of the COnsortium for Small-scale MOdelling (COSMO) based on the Kilometre-scale ENsemble Data Assimilation system (KENDA), developed by Deutscher Wetterdienst and its partners. KENDA provides a state-of-the-art ensemble data assimilation system on the convective scale for operational data assimilation and forecasting based on the Local Ensemble Transform Kalman Filter (LETKF). In this study, the Efficient Modular VOlume RADar Operator is applied for the assimilation of radar reflectivity data to improve short-term predictions of precipitation. Both deterministic and ensemble forecasts have been carried out. A case-study shows that the assimilation of 3D radar reflectivity data clearly improves precipitation location in the analysis and significantly improves forecasts for lead times up to 4 h, as quantified by the Brier Score and the Continuous Ranked Probability Score. The influence of different update rates on the noise in terms of surface pressure tendencies and on the forecast quality in general is investigated. The results suggest that, while high update rates produce better analyses, forecasts with lead times of above 1 h benefit from less frequent updates. For a period of seven consecutive days, assimilation of radar reflectivity based on the LETKF is compared to that of DWD's current operational radar assimilation scheme based on latent heat nudging (LHN). It is found that the LETKF competes with LHN, although it is still in an experimental phase.

Key Words: radar data assimilation; ensemble Kalman Filter; LETKF; convection-permitting model

Received 11 September 2015; Revised 30 December 2015; Accepted 21 January 2016; Published online in Wiley Online Library 15 March 2016

1. Introduction

Predicting convective events is challenging due to the atmosphere's chaotic and nonlinear behaviour. For predictions up to 6 h, nowcasting based on advection schemes mostly outperforms model forecasts due to uncertainties in the initial state and spin-up effects (Lin *et al.*, 2005). Improvements in short-term predictive skill can thus be achieved by reducing the uncertainty in initial conditions through data assimilation (Talagrand, 1997), and by incorporating more detailed process descriptions and life-cycle effects in observation-based nowcasting (Sun *et al.*, 2014; Wapler *et al.*, 2015).

During the past decades, nowcasting has usually denoted an observation-based initial state description with a forecast component based on extrapolation. Nowadays, nowcasting also refers to blending of extrapolation, statistical techniques, high-resolution data assimilation and rapid cycling numerical weather prediction (NWP), as for example suggested by Lilly (1990) or

Sun *et al.* (2014). The exploitation of 3D radar data for data assimilation presented in this article is an important step towards an integrated forecasting system.

Many widely used data assimilation schemes based on variational methods (3D-Var, 4D-Var) are already successfully applied to convective-scale assimilation of radar data both in idealized and real-data frameworks (Sun and Crook, 1997, 1998; Caya *et al.*, 2005; Schwitalla and Wulfmeyer, 2014). In recent years, ensemble-based assimilation methods have gained popularity because they allow us to estimate and use flow-dependent covariances, i.e. prediction uncertainties, from the ensemble instead of climatological covariances with static structures as in variational methods (Nichols, 2010; Freitag and Potthast, 2013).

Ensemble data assimilation constitutes a Monte Carlo approach: the ensemble forecasts represent a Monte Carlo sample of the prior probability distribution, and a Monte Carlo sample of the posterior distribution is obtained by assimilating observations. In the case of Gaussian probability distribution functions, the Ensemble Kalman Filter (EnKF) can be directly derived from Bayes' theorem. The Gaussian assumption makes EnKFs feasible for large-scale applications like operational NWP. But the Gaussian assumption is also a strong limitation, since atmospheric dynamics are highly nonlinear. However, the EnKF has proven successful for operational NWP (e.g. at the Canadian Meteorological Centre; Houtekamer and Mitchell, 2005).

The EnKF originally was derived from Evensen (1994), and the idea was applied to NWP first by Houtekamer and Mitchell (1998). Burgers *et al.* (1998) showed that the derivation of Evensen (1994) leads to an underestimation of the analysis-error covariance and developed an approach based on the perturbation of observations. A deterministic alternative to the perturbation of observations is for example given by the Ensemble Transform Kalman Filter (ETKF; Bishop *et al.*, 2001).

Due to limited computing power, it is usually not feasible to run an ensemble of more than 50 members in operational NWP applications. Since the model state space has a much higher dimension, the ensemble covariance provides only a low rank approximation of the real covariance structure. To overcome the limited ensemble size, localization is used. Localization is a key feature of the Local Ensemble Transform Kalman Filter (LETKF) by Hunt *et al.* (2007), which restricts the radius of influence of the observations. Analyses are calculated independently for each grid point which allows easy parallelization. Furthermore, different linear combinations of the ensemble members can be chosen for different regions. Thus, the global analysis is not restricted to the low-dimensional ensemble space, but can originate from a higher dimensional space (Hunt *et al.*, 2007). Localization is an important aspect in current research (e.g. Miyoshi *et al.*, 2007; Greybush *et al.*, 2011; Janjić *et al.*, 2011; Perianez *et al.*, 2014; Kirchgessner *et al.*, 2014). Recently, non-Gaussian methods (particle filters) have been analyzed, e.g. by van Leeuwen (2009) and Ades and van Leeuwen (2012), and a particle filter variant has successfully been applied to short-range NWP by Milan *et al.* (2014).

Radar observations are promising to improve predicting convective events in NWP because they capture the 3D spatial and temporal evolution of these systems. In current operational NWP systems, radar data are usually integrated based on two-dimensional composites of near-surface reflectivities or derived surface rain rates. The operational exploitation of the full three-dimensional information of the measurements is still in its infancy.

In a very first application, Snyder and Zhang (2003) revealed the potential of the EnKF for assimilating radar observations at convective scales in an Observing System Simulation Experiment (OSSE). The authors assimilated simulated radial velocities and proved the ability of the EnKF to transfer information onto the unobserved model variables. Further OSSE studies confirmed the results of Snyder and Zhang (2003), e.g. Caya *et al.* (2005), Tong and Xue (2005), Xue *et al.* (2006), Gao and Xue (2008), Sobash and Stensrud (2013), Lange and Craig (2014). In most of these studies, the radar variables were simulated directly on the model grid points. A few real-data studies focused on the impact of radar data on the analysis only, e.g. Dowell *et al.* (2004), Dowell *et al.* (2011), Snook *et al.* (2011), Jung *et al.* (2012), or excluded radar reflectivity in the assimilation, e.g. Zhang *et al.* (2009), Dowell and Wicker (2009), Chang *et al.* (2014). Research on the impact of radar reflectivity on ensemble forecasts in an imperfect model framework is still limited and so far mostly very short-range forecasts (up to 3 h) have been investigated (Aksoy *et al.*, 2010; Dong *et al.*, 2011; Dong and Xue, 2013; Snook *et al.*, 2015).

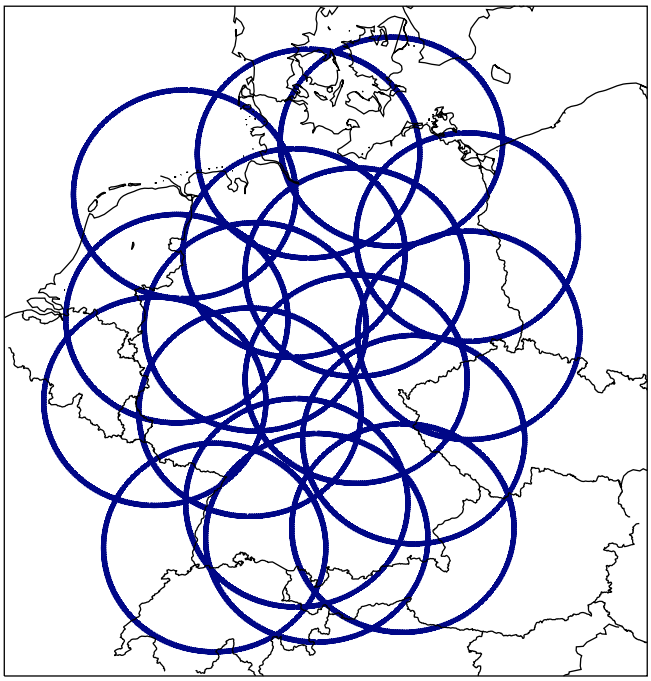

Figure 1. The DWD radar network provides a very dense coverage over Germany. For each station, the range is shown by a circle of 180 km radius.

In most studies, observations from only a single radar site are assimilated usually over US terrain where, as pointed out by Dowell *et al.* (2004) and Tong and Xue (2005), radar sites are far apart. Deutscher Wetterdienst (German Meteorological Service, DWD) operates a dense network consisting of 17 C-band radar sites covering Germany and parts of its bordering countries (Figure 1). The dense network assures a complete coverage over Germany. Due to large overlapping areas, dense vertical information is also available. 3D radar measurements are available with a temporal resolution of 5 min in standard operational mode.

In this study, 3D radar reflectivity measurements are assimilated in addition to conventional observations, as performed by Dong *et al.* (2011) in an OSSE framework and by Snook *et al.* (2015).

The Kilometre-Scale Ensemble Data Assimilation system (KENDA; Schraff *et al.*, 2016) is applied, which implements the LETKF following Hunt *et al.* (2007) for the NWP model COSMO-DE (Doms and Schättler, 2002; Baldauf *et al.*, 2011). Radar reflectivities are simulated by the Efficient MOdular RAdar Operator (EMVORADO; Blahak, 2008b; Blahak *et al.*, 2011; Zeng, 2013; Jerger, 2013; Zeng *et al.*, 2014) based on the COSMO-DE model fields. Three experiments address in particular the impact of the assimilation of radar reflectivities on the skill of precipitation forecasts.

- First, 3D radar reflectivity data are assimilated into COSMO-DE via KENDA and the skill of a subsequent ensemble forecast is evaluated.
- Second, the influence of the data assimilation update rate on the analysis and forecast is studied. Assimilation intervals of 5, 15, 30 and 60 min are analyzed. The noise introduced into the model runs is quantified by the domain-averaged surface pressure tendency during the assimilation cycle.

- A third experiment compares the forecasts with results obtained with the operational latent heat nudging (LHN; Stephan *et al.*, 2008), over a case-study period of seven consecutive days.

Thus, the novelty of this work consists of the assimilation of radar reflectivities with an ensemble Kalman filter over a central European domain, the investigation of the influence of update rates in a non-idealized setting, and the comparison to LHN. Furthermore, this study is not restricted to a few case-studies but examines a longer period.

Section 2 describes the model, the data assimilation algorithm and its implementation and software design. The radar forward operator is presented in section 3. Section 4 includes a description of the experimental set-up of the three experiments and also introduces the scores used for the evaluation of the numerical simulations. The results are presented and evaluated in section 5. Section 6 concludes with a summary and discussion.

2. NWP model and data assimilation scheme

2.1. The COSMO model

The COSMO model is a limited-area, non-hydrostatic, and fully compressible model for NWP and regional climate simulations (Doms and Schättler, 2002; Bachner *et al.*, 2008; Baldauf *et al.*, 2011), which has been developed by the multi-national COnsortium for Small-scale MOdelling (COSMO). COSMO-DE is a specific model configuration which is run operationally at DWD at 2.8 km horizontal resolution over central Europe (Figure 2) with a numerical grid consisting of 421 × 461 columns resolved into 50 terrain-following hybrid layers. In this study, an ensemble of lateral boundary conditions is provided by the global model ICON (Zängl *et al.*, 2015), which is also run in an LETKF framework. For solving the equations for compressible flow in a moist atmosphere, the time-splitting Runge–Kutta approach of Wicker and Skamarock (2002) is used in the COSMO model. Due to the high horizontal resolution, the model is able to resolve deep convection explicitly. Shallow convection is parametrized following the non-precipitating part of the Tiedtke scheme (Tiedtke, 1989). In order to meet the requirements of the explicit simulation of deep convection, the microphysics scheme is a Lin-type one-moment bulk microphysics scheme that includes cloud droplets, cloud ice, rain, snow and graupel (Lin *et al.*, 1983; Seifert and Beheng, 2001; Reinhardt and Seifert, 2006). Turbulence parametrization is based on the prognostic turbulent kinetic energy (TKE) equation according to Raschendorfer (2001). Radiative transfer is modelled according to Ritter and Geleyn (1992). The lower boundary conditions are generated by the multi-layer soil model TERRA (e.g. Grasselt *et al.*, 2008); a more detailed description is given in Doms *et al.* (2011).

2.2. Ensemble Kalman filter

This section briefly describes the Ensemble Kalman Filter (EnKF) method used in the present study. A more detailed description is given in Evensen (1994) and Hunt *et al.* (2007); an interpretation in a wider context is given in Freitag and Potthast (2013).

In the LETKF, the background-error covariance matrix $\mathbf{P}^{\mathrm{b}}$ (forecast uncertainty or first-guess uncertainty) is estimated by the background ensemble via

$$\mathbf{P}^{\mathrm{b}} = (N-1)^{-1}\mathbf{X}^{\mathrm{b}}(\mathbf{X}^{\mathrm{b}})^{\mathrm{T}}, \tag{1}$$

with columns of the matrix $\mathbf{X}^{\mathrm{b}}$ containing the background ensemble perturbations

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

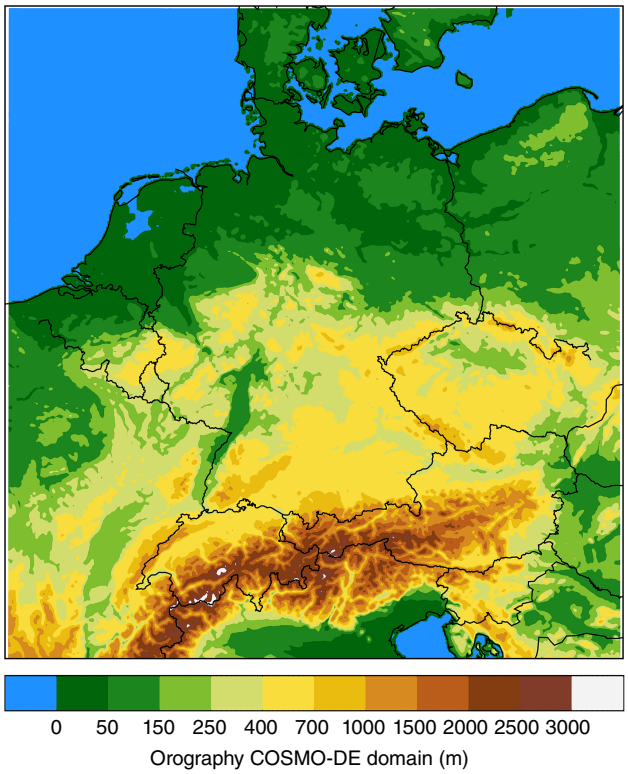

Figure 2. The orography (m) of the COSMO-DE domain in use operationally at DWD.

i.e. the deviations from the ensemble mean $\overline{\mathbf{x}}^{\mathrm{b}}$ of the N ensemble members $\mathbf{x}^{\mathrm{b}(i)}$, $i = 1, \ldots, N$. The LETKF analysis ensemble is constructed as a linear combination of the background perturbations:

$$\mathbf{x}^{\mathrm{a}(i)} = \overline{\mathbf{x}}^{\mathrm{b}} + \mathbf{X}^{\mathrm{b}}\mathbf{w}^{\mathrm{a}(i)}, \qquad i = 1, \ldots, N, \tag{3}$$

where $\mathbf{w}^{\mathrm{a}(i)}$ is a coefficient vector of size N. The mean $\overline{\mathbf{w}}^{\mathrm{a}}$ is determined by minimizing the cost function

$$\begin{aligned} J(\mathbf{w}) =& (N-1)(\mathbf{w}^{\mathrm{T}}\mathbf{w}) \\ &+ (\mathbf{y}^{\mathrm{o}} - \overline{\mathbf{y}}^{\mathrm{b}} - \mathbf{Y}^{\mathrm{b}}\mathbf{w})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}^{\mathrm{o}} - \overline{\mathbf{y}}^{\mathrm{b}} - \mathbf{Y}^{\mathrm{b}}\mathbf{w}), \end{aligned} \tag{4}$$

where $\mathbf{R}$ is the observation-error covariance matrix, and $\mathbf{y}^{\mathrm{o}}$ the observation vector. The nonlinear observation operator H is linearly approximated by

$$H(\overline{\mathbf{x}}^{\mathrm{b}} + \mathbf{X}^{\mathrm{b}}\mathbf{w}) \approx \overline{\mathbf{y}}^{\mathrm{b}} + \mathbf{Y}^{\mathrm{b}}\mathbf{w}. \tag{5}$$

$\mathbf{Y}^{\mathrm{b}}$ denotes the ensemble perturbation matrix in observation space. Its columns are given by the deviations of the background fields in observation space from the ensemble mean in observation space

$$\mathbf{Y}^{\mathrm{b}} = \left[\mathbf{y}^{\mathrm{b}(1)} - \overline{\mathbf{y}}^{\mathrm{b}}, \ldots, \mathbf{y}^{\mathrm{b}(N)} - \overline{\mathbf{y}}^{\mathrm{b}}\right], \tag{6}$$

where $\mathbf{y}^{\mathrm{b}(i)} = H(\mathbf{x}^{\mathrm{b}(i)})$ and $\overline{\mathbf{y}}^{\mathrm{b}}$ their mean. The minimum of J in Eq. (4) is found by:

$$\overline{\mathbf{w}}^{\mathrm{a}} = \mathbf{P}^{\mathrm{a}}(\mathbf{Y}^{\mathrm{b}})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}^{\mathrm{o}} - \overline{\mathbf{y}}^{\mathrm{b}}), \tag{7}$$

where the analysis-error covariance estimation in ensemble space is given by

$$\mathbf{P}^{\mathrm{a}} = \{(N-1)\mathbf{I} + (\mathbf{Y}^{\mathrm{b}})^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Y}^{\mathrm{b}}\}^{-1}. \tag{8}$$

The analysis ensemble members in ensemble space are then sampled by

$$\mathbf{w}^{\mathrm{a}(i)} = \overline{\mathbf{w}}^{\mathrm{a}} + \mathbf{W}^{\mathrm{a}(i)}, \tag{9}$$

where $\mathbf{W}^{a(i)}$ is the ith column of the matrix $\mathbf{W}^a$, which is the symmetric square root of $\mathbf{P}^a$:

$$\mathbf{W}^a = [(N-1)\mathbf{P}^a]^{1/2}. \quad (10)$$

The vectors $\mathbf{w}^{a(i)}$ determine the analysis ensemble as described in Eq. (3). Since each member $\mathbf{x}^{a(i)}$ of the analysis ensemble is a linear combination of the background ensemble perturbations, the analysis increments are restricted to the $(N-1)$-dimensional subspace spanned by the ensemble perturbations, and sampling errors can lead to spurious long-distance correlations in $\mathbf{P}^b$. To overcome this issue, analyses are performed independently for each model grid point taking only nearby observations into account; this is known as observation localization. The observation influence is commonly weighted by the Gaspari–Cohn correlation function (Gaspari and Cohn, 1999). As demonstrated by Perianez *et al.* (2014), the optimal localization radius depends on the observation density and quality. Since conventional and radar observations differ in both, different localization radii for the different observation types are chosen in this study. The Gaspari–Cohn localization length-scale is set to 80 km corresponding to a radius of main influence within 160 km for conventional observations, and 16 km corresponding to a radius of main influence within 32 km for radar observations.

2.3. The KENDA system

The KENDA system (Schraff *et al.*, 2016) implements the LETKF described above for COSMO-DE as a fully four-dimensional LETKF. During the COSMO forward integration, observation operators are applied at all observation time steps, i.e. a model equivalent is simulated for each available observation within the first-guess interval. Thus, KENDA is able to weight the ensemble members according to their trajectory over this first-guess window. In its standard configuration, KENDA runs in a 40+1 mode, meaning 40 ensemble members are contributing to the calculation of the weights in Eq. (7), which updates an additional independent deterministic run via

$$\mathbf{w}^{b(det)} = \mathbf{P}^a(\mathbf{Y}^b)^T\mathbf{R}^{-1}(\mathbf{y}^o - \mathbf{y}^{b(det)}), \quad (11)$$

where $\mathbf{y}^{b(det)} = H(\mathbf{x}^{b(det)})$. The deterministic run is used instead of picking one of the ensemble members for the following reasons: the deterministic run serves as a realistic version of the ensemble mean and is a better estimate than most of the members since the mechanism of the LETKF is to artificially create a certain amount of ensemble spread around the ensemble mean. Also, KENDA has been designed to be able to run the deterministic run at higher resolution than the ensemble (not employed in our tests), which takes care of the limited computing resources and maintains the benefit of ensemble runs.

KENDA addresses several issues relevant for convective-scale data assimilation, which we briefly summarize here (details in Schraff *et al.*, 2016).

2.3.1. Coarsening of the analysis grid

The KENDA suite allows us to compute the analysis weights on a coarsened grid (Yang *et al.*, 2009). The coarse analysis weights are interpolated onto the high-resolution grid before calculating the analysis ensemble in model space, i.e. the analysis increments are determined from the full-resolution background ensemble and therefore still capture small-scale features. Yang *et al.* (2009) show that this method efficiently decreases the computational costs and maintains a high accuracy of the analysis. Furthermore, the authors conclude that the weight interpolation may even damp unwanted imbalances, leading e.g. to gravity waves. In this study, we use a coarsening factor of 3, i.e. the analysis grid uses approximately 11% of the full model grid points.

2.3.2. Inflation and relaxation

The LETKF formalism as presented in section 2.2 does not include the model error. One way of accounting for model error is to inflate the ensemble and thus increase the ensemble spread. KENDA allows for multiplicative covariance inflation (Anderson and Anderson, 1999), relaxation to prior perturbations (RTPP; Zhang *et al.*, 2004) and relaxation to prior spread (RTPS; Whitaker and Hamill, 2012). The different approaches are analyzed and discussed in Harnisch and Keil (2015). In this study, RTPS is used, which relaxes the analysis ensemble spread σ^a towards the prior ensemble spread σ^b via

$$\sigma^a \leftarrow (1-\alpha)\,\sigma^a + \alpha\,\sigma^b. \quad (12)$$

The factor α is in this study set to $\alpha = 0.95$ giving a high weight to the prior spread following the suggestion of Whitaker and Hamill (2012). KENDA allows us to perform inflation adaptively. However, assimilating both sparse conventional observations and dense radar observations with adaptive inflation or localization is not straightforward, and would justify a separate study which exceeds the scope of this work. We therefore refrain from using the adaptive scheme in KENDA throughout this work.

2.3.3. Latent heat nudging

Latent heat nudging (LHN) denotes a nudging scheme for the assimilation of radar-derived precipitation rates (Stephan *et al.*, 2008; Milan *et al.*, 2008; Schraff *et al.*, 2016). The scheme modifies the thermodynamic state based on the relationship between precipitation rates near the surface and latent heat release. LHN is part of the COSMO package and can be continuously applied during the model forward integration. LHN has been used operationally as the (deterministic) data assimilation system for precipitation at DWD since April 2007. To compare the KENDA-based assimilation of radar reflectivities with LHN, LHN is integrated into KENDA by applying LHN to each ensemble member and the deterministic run during the model forward integration.

2.4. Cycling

Ensemble-based data assimilation needs a cycling between model runs and assimilation steps. DWD has developed the basic cycling environment (BACY), which mimics the operational data assimilation cycle. The BACY shell scripts invoke both the COSMO runs and the KENDA analysis within the assimilation loop. Due to this design, COSMO is newly initialized after each assimilation step. To avoid an entirely cold start of COSMO, the following variables are passed between KENDA and COSMO: the wind components u, v, w, temperature T, relative humidity q_v, pressure perturbation pp, and hydrometeor contents q_c, q_i, q_r, q_s, q_g (cloud drops, cloud ice, rain, snow and graupel).

For implementation on the convective scale, high update rates (rapid update cycles; RUCs) are important for bridging the gap towards nowcasting. The BACY system was adapted to allow testing of RUC including a flexible choice of the update time window Δt.

3. 3D volume radar forward operator

The EMVORADO* operator (Blahak, 2008b; Blahak *et al.*, 2011; Zeng, 2013; Jerger, 2013; Zeng *et al.*, 2014) simulates the 3D measurement process of the equivalent radar reflectivity factor ('reflectivity') and the radial wind of volume-scanning weather radars based on the COSMO model fields during runtime. Each

*Efficient Modular VOlume RADar Operator. This name has been newly defined and has not been previously used in the cited literature.

radar station (Figure 1) is processed individually. For this study only the reflectivity data are used.

EMVORADO runs efficiently on parallel and vector supercomputers (Blahak *et al.*, 2011; Zeng, 2013) to meet the demands of operational applications. The modular design of EMVORADO allows us to choose among different options with different numerical costs to find the 'best' compromise between run time and accuracy depending on application. Considered physical aspects (some are optional) are (details are in the cited literature):

- beam propagation, beam bending, beam blockage by (model) orography;
- Rayleigh- or Mie-scattering theory;
- partially melted particles ('bright band');
- attenuation by hydrometeors;
- beam weighting function averaging in the pulse volume, effective beam weighting function of a scanning radar;
- reflectivity weighting and hydrometeor fallspeed for radial wind.

3.1. Operator characteristics

First, reflectivity Z_e and (optionally) extinction coefficient Λ are computed for each hydrometeor category and summed up over all categories. Then, these grid point values are interpolated/averaged to the polar radar coordinates, taking into account beam bending/blockage and optionally the beam weighting function and attenuation.

The model hydrometeor information is converted into reflectivity Z_e and extinction coefficient Λ according to

$$Z_e = \frac{\lambda_0^4}{\pi^5 \left|K_{w,0}\right|^2} \int_0^\infty \sigma_b(D)\, N(D)\, \mathrm{d}D, \tag{13}$$

$$\Lambda = \int_0^\infty \sigma_{ext}(D)\, N(D)\, \mathrm{d}D, \tag{14}$$

with σ_b the backscatter coefficient and σ_{ext} the extinction coefficient of a particle with diameter D. Both depend on the chosen scattering theory and the complex refractive index m_r of the particle material as a function of D, bulk density ρ_{bulk} and radar wavelength λ_0. ρ_{bulk} is computed from the assumed mass–size relations (model microphysics) for a mass-equivalent sphere having diameter D. $N(D)$ is the particle size distribution derived from the model hydrometeor contents with shapes consistent with the microphysics (exponential or gamma distribution, depending on hydrometeor type). No subgrid-scale variability is considered. $\left|K_{w,0}\right|^2$ is a reference value of the dielectric factor of water at 0°C, usually 0.93. Mie scattering has been chosen for this study. Normally this would be much more costly than Rayleigh but, by using efficient look-up tables (Jerger, 2013), this option is computationally as efficient as the Rayleigh approximation.

Rain drops are assumed as water spheres. Cloud ice and graupel are modelled as largest-diameter-equivalent one-layered spheres composed of an ice/air or ice/water/air mixture material, the latter in the case of partially melted particles. Here, the melt water is assumed to be homogeneously distributed within the particle (soaked). Snowflakes, following an idea of Fabry and Szyrmer (1999), are assumed as two-layered spheres (Mie solution of Kerker, 1969) with a denser core and a less dense shell, with the core radius being half the shell radius. As for graupel and ice, meltwater is assumed to soak partially melted snow flakes. The m_r of partially melted graupel and cloud ice is computed according to a variant of the Maxwell Garnett mixing rule (Maxwell Garnett, 1904) for a twofold two-component mixture where spheroidal air inclusions are assumed to be suspended into an ice/air mixture (matrix), with ice inclusions assumed to be suspended in water. For snow, this model is used also for the core, whereas for the shell the air is assumed to be the matrix.

The degree of melting for all species is parametrized as a function of temperature and particle size (Blahak, 2008b). For graupel, wet growth is considered down to temperatures of −10 °C.

The full 3D radar operator equation for one pulse-volume-averaged Z_e value at range r_0, azimuth α_0 and elevation angle ϵ_0 is given by Eq. (15), where ℓ_n^{-2} is the path-integrated attenuation from the radar to location (r, α, ϵ), c the speed of light, and τ the pulse duration. f_e^4 is the approximate two-way effective beam weighting function of a scanning radar (Blahak, 2008a). The range weighting function is approximated by a simple box function.

In EMVORADO, volume integrals are computed by Gauss–Legendre quadrature. However, for this study beam weighting is not used, thus the beam collapses to its central ray ('pencil-beam') and the integrals in Eq. (15) vanish. Beam bending is computed by the 4/3 Earth radius approximation as a good compromise between accuracy and efficiency for applications in ensemble data assimilation (Zeng, 2013).

Measurements are simulated for each radar of the DWD radar network for plan position indicators (PPI) at ten elevation angles ranging from 0.5° to 25°. A full volume scan is available every 5 min with a range resolution of 1 km and an azimuthal resolution of 1° and cover a radius of 180 km.

3.2. Observation resolution

Close to the radar, observations are much denser than the model grid. However, assimilating observations that are denser than the analysis grid resolution might degrade the model state in the assimilation step (Liu and Rabier, 2002). Averaging several radar bins to so-called superobservations (superobbing), reduces the density of both the observed and the simulated data. In EMVORADO, large amounts of data have to be handled. Thus, for efficiency, superobbing is done by the forward operator both for the model and the observations. Superobbing is applied to each radar station and elevation individually. In overlapping regions, observations from each radar are treated as independent data. Superobbing is implemented in a quasi-Cartesian way: A 2D horizontal Cartesian grid with a desired resolution is constructed. For each elevation, the radar bin centres (range, azimuth) are projected onto the Cartesian plane in order to find the closest radar bin for each Cartesian grid point. This radar bin is the centre of the superobservation, i.e. the superobservations are constructed in radar coordinates. Next, several radar bins surrounding the centre bin are averaged, where the size of the averaging area ('pie-wedge'-shaped area) is defined by a range and azimuth interval. The number of radar bins that go into a superobservation decreases with range. The width of the range interval is defined by $L_x\sqrt{2}$, where L_x is the resolution of the underlying Cartesian grid, and the number of azimuths at range r_0 is given by 2 arctan $\left\{(L_x\sqrt{2}/2)/r_0\right\}$. Thus, for $L_x = 10$ km, close to the radar station, approximately 900 radar bins are averaged, whereas beyond a range of 120 km, fewer than 100 bins are averaged. This procedure is repeated for all elevations. With this strategy, a relatively homogeneous horizontal data distribution is achieved and at the same time the same data structures can be reused in the superobbing code as for the ordinary radar data, preserving runtime efficiency. The data reduction also significantly decreases the computational costs in the LETKF step.

An example of the superobbing is shown in Figure 3 which compares the measurements from one radar in its original resolution with superobbing at 10 km horizontal resolution. Since the analysis grid is coarsened by a factor of 3 (cf. section 2.3), we choose a superobbing resolution of 10 km to ensure that model and observations have approximately the same horizontal resolution.

3.3. No-precipitation information

Data assimilation has also been found valuable to suppress spurious precipitation by including no-precipitation information

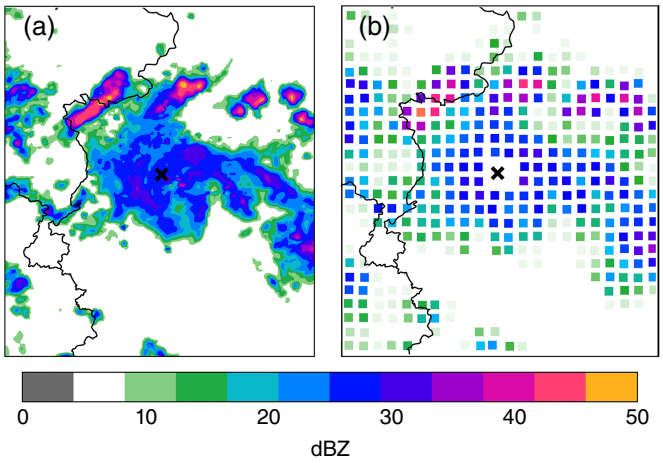

Figure 3. An example for the superobbing is shown based on a measurement from the radar station in Essen, Germany on a domain of approximately 220 km by 260 km. Panel (a) shows a contour plot of the original polar PPI data, and (b) shows the superobbed measurement (10 km). The location of the station is shown by the black cross. Each averaging area, given in polar coordinates, approximates a quadratic area. This area is only well defined if it is assured that the radar station is not contained in any of these areas. This leads to small gaps around the station.

(Tong and Xue, 2005). Therefore, reflectivity information is also assimilated in clear-air regions. However, the measured dBZ values in logarithmic units range from −30 dBZ up to 60 dBZ, and a large subset of this range refers to non-precipitating signals. Thus, deviations of no-precipitation observations and simulations might become very large. In the assimilation step, such differences would be translated into large analysis increments without physical relevance. To mitigate this arbitrariness, all Z_e values below a certain value are thresholded, which keeps the weight of the non-precipitation information in the assimilation process within reasonable limits. This weight is largely determined by the absolute value of the threshold. Based upon experimentation, we selected a threshold of 5 dBZ; this threshold has previously been used by other authors, e.g. Aksoy *et al.* (2009). This threshold is applied before superobbing.

$$\langle Z_e^{(R)} \rangle (r_0, \alpha_0, \epsilon_0) = \frac{\displaystyle\int_{r_0-c\tau/4}^{r_0+c\tau/4} \int_{\alpha_0-\pi}^{\alpha_0+\pi} \int_{\epsilon_0-\pi/2}^{\epsilon_0+\pi/2} Z_e(r,\alpha,\epsilon) \overbrace{\exp\left(-2\int_0^r \Lambda(r',\alpha,\epsilon)\,\mathrm{d}r'\right)}^{\ell_n^{-2}(r,\alpha,\epsilon)} \frac{f_e^4(\alpha,\epsilon)}{r^2} \cos\epsilon\,\mathrm{d}\epsilon\,\mathrm{d}\alpha\,\mathrm{d}r}{\displaystyle\int_{r_0-c\tau/4}^{r_0+c\tau/4} \int_{\alpha_0-\pi}^{\alpha_0+\pi} \int_{\epsilon_0-\pi/2}^{\epsilon_0+\pi/2} \frac{f_e^4(\alpha,\epsilon)}{r^2} \cos\epsilon\,\mathrm{d}\epsilon\,\mathrm{d}\alpha\,\mathrm{d}r} \tag{15}$$

3.4. *Observation errors*

In this study, the observation-error covariance matrix **R** (cf. Eqs (4), (7) and (8)) is assumed to be diagonal, i.e. observation errors are assumed to be spatially uncorrelated. We acknowledge that this assumption is only a rough first approximation. The standard deviation of each observation bin is set to 10 dBZ. Although this value appears to be somewhat larger than that used in other studies (e.g. 5 dBZ in Tong and Xue, 2005), this value has been found appropriate based on experimentation.

4. Experimental set-up

4.1. *Observation datasets*

The conventional observations assimilated by KENDA comprise aircraft measurements (AMDAR) of temperature and horizontal wind, surface station measurements (SYNOP) of 10 m horizontal wind and surface pressure, radiosonde data (TEMP) of temperature, horizontal wind and humidity, and wind profiler measurements of horizontal wind. Table 1 gives the number of observations of each data type and variable used in the assimilation including the number of radar bins after superobbing.

A clutter filter removes ground clutter from the radar observations but moving features such as wind turbines or migratory birds are not removed. However, the latter are of small intensity and thus effectively removed by thresholding at 5 dBZ. Although individual radar bins were collected during a 5 min window, it is assumed all observations are valid at the 5 min intervals. Exact observation times are not considered.

The LETKF-based radar data assimilation uses 3D reflectivity data, while LHN assimilates a 2D mosaic of radar-derived surface precipitation rates. This mosaic is also used for verification and therefore described in more detail. The mosaic merges the so-called precipitation scans (terrain-following scans every 5 min), of the 17 German radar sites, which are not included in the 3D volume data used by KENDA via EMVORADO. The measured reflectivity from the precipitation scans is transformed to a precipitation rate using weather-dependent $Z - R$ relations of the form $Z = a R^b$. A typical choice for a and b is given by Marshall and Palmer (1948). However, these parameters are based on a particular drop-size distribution and not applicable for e.g. hail. Thus, five different classes distinguish between stratiform and convective cases as well as hail events with the class based on the measured reflectivity. In regions of overlap, the maximum of the derived precipitation rates is taken.

In the operational set-up, the mosaic used within LHN uses data from the German radar sites as well as data from stations in neighbouring countries. However, the EMVORADO operator so far only simulates measurements from the German radar stations. To allow for a fair comparison between KENDA and LHN, only the German sites are assimilated by LHN in this study.

4.2. *Description of the experiments*

The impact of 3D radar reflectivity assimilation on the analysis and precipitation forecasts has been investigated in three data assimilation experiments. The first experiment evaluates the improvement by assimilating radar data on top of conventional data. In the second experiment the impact of the cycling frequency is investigated. The third experiment evaluates a longer time period and compares the assimilation of 3D reflectivity data with the assimilation of 2D-derived surface precipitation rates via LHN.

4.2.1. *Experiment 1*

The first experiment studies the influence of the assimilation of radar reflectivity during 26 May 2014. Between a slowly propagating trough stretching from Ireland to the Bay of Biscay and an extensive high over eastern Europe, a convergence zone extended from westnorthwestern to southeastern Germany. Along this line, precipitation developed over the day which in places persisted for several hours and locally led to high precipitation amounts (Figure 4(b, e, h)).

- The ensemble is initialized at 0000 UTC and propagated for 12 h without data assimilation to ensure the ensemble

Table 1. Number of conventional observations (AMDAR, SYNOP, TEMP and wind profiler) used in the assimilation and number of used radar observations after superobbing. The numbers are based on average values determined during the week 22–29 May 2014 and denote the number of observations approximately collected over 1 h. The number of available AMDAR and TEMP data depends on the time of the day: there are fewer AMDAR measurements during the night, and radiosondes are only launched four times a day. For each observation type, the temporal resolution is also given (column 3).

Observation type	Assimilated variables	Temporal resolution	Number of observations per hour
AMDAR	Temperature	1 min	0–520
	Horizontal wind		0–550
SYNOP	10 m horizontal wind	1 h	240
	Surface pressure		600
TEMP	Temperature	6 h	0–270
	Horizontal wind		0–300
	Relative humidity		0–170
Wind profiler	Horizontal wind	30 min	500
Radar	Reflectivity	5 min	1 300 000

approximates the forecast error in a realistic manner following Zhang *et al.* 2009.

- In the ensuing assimilation cycle, the observations are assimilated over a time period of 3 h with hourly updates. In the set-up named CONV, only conventional observations are assimilated, while in CONV+RAD, radar reflectivities are assimilated in addition to conventional observations.
- The analysis step is followed by a 6 h ensemble and deterministic forecast for the period 1500–2100 UTC. The deterministic forecast is initialized from the analysis of the deterministic KENDA run (Eq. (11)).

4.2.2. Experiment 2

Given the CONV+RAD set-up from experiment 1, experiment 2 is designed to to find the optimal update interval, again for 26 May 2014.

- Again, prior to data assimilation, the ensemble is run for 12 h from 0000 to 1200 UTC
- The assimilation cycle is run from 1200 to 1500 UTC, now with observations assimilated every 5, 15, 30 or 60 min. In each assimilation step, all data collected during the previous time window is assimilated by the 4D-LETKF of KENDA. The set-ups are called CONV+RAD_5, CONV+RAD_15, CONV+RAD_30, and CONV+RAD_60.
- The analysis cycle is again followed by a 6 h forecast for the ensemble and the deterministic run(s).

4.2.3. Experiment 3

The experiment encompasses the week 22–29 May 2014 with forecasts initialized every 6 h. The weather during this week was dominated by a deepening quasi-stationary trough across western Europe and a ridge which extended from the central Mediterranean towards the Baltic Sea and Belarus/Russia. With a strong southsouthwesterly flow from the Mediterranean Sea, warm moist and potentially unstable air masses were advected towards central Europe. At the eastern edge of the deep trough, cyclogenesis was induced repeatedly. Due to dynamic forcing ahead and along the frontal systems which moved over central Europe, precipitation systems developed. Towards the end of the period the pressure gradient in central Europe weakened. The period included large frontal systems as well as single cells and multi-cells. On average, precipitation during this time period had a clear diurnal cycle.

- Three set-ups are studied: CONV, CONV+RAD and CONV+LHN. In the latter, conventional observations are assimilated with the LETKF as in CONV and LHN is applied to every ensemble member and the deterministic run.
- The assimilation cycle is initialized on 21 May 2014 at 1200 UTC. For all set-ups, observations are assimilated hourly until 0000 UTC on 29 May 2014.
- 24 h deterministic forecasts are initialized every 6 h, beginning at 0000 UTC on 22 May, until 0000 UTC on 29 May.

A summary of the set-ups used in the experiments is provided in Table 2

4.3. Verification scores

4.3.1. Experiments 1 and 2

For verification of the ensembles, the Brier Score (BS; Wilks, 2006) and the Continuous Ranked Probability Score (CRPS; Matheson and Winkler, 1976) are computed. Verification is performed both for reflectivity in dBZ and for radar-derived precipitation in $\mathrm{mm\,h^{-1}}$. For reflectivity, the superobbed data of the lowest PPI scan are used. In regions with station overlap, the data of the closest station are used.

The Brier score measures the accuracy of probabilistic forecasts in terms of the mean squared error between forecast and observation. To this goal, observations and model output are transferred into binary fields based on the exceedance of a threshold. The probabilistic forecast at a particular grid point k is given by the fraction of the ensemble members exceeding the threshold, denoted $p(x_k)$. The observation is either 1 if the event occurred or 0 otherwise. The Brier score is thus defined as

$$BS = \frac{1}{n}\sum_{k=1}^{n}\{p(x_k) - y_k\}^2. \tag{16}$$

The Brier skill score (BSS) is obtained via

$$BSS = 1 - \frac{BS}{BS_{\mathrm{ref}}}. \tag{17}$$

CONV is taken as the reference, leading to

$$BSS = 1 - \frac{BS(CONV + RAD)}{BS(CONV)}. \tag{18}$$

Therefore, positive BSS values imply an improvement by CONV+RAD_5, CONV+RAD_15, CONV+RAD_30, and CONV+RAD_60.

For the CRPS, its alternative form (Gneiting and Raftery, 2007)

$$CRPS(F, y) = \mathrm{E}_F|X - y| - \frac{1}{2}\mathrm{E}_F|X - X'|, \tag{19}$$

is used. Here, X and X' are independent copies of a random variable with distribution F. For a discrete ensemble, the CRPS is calculated according to Grimit *et al.* (2006) via

$$CRPS = \frac{1}{N}\sum_{i=1}^{N}|x_i - y| - \frac{1}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}|x_i - x_j|, \tag{20}$$

and averaged over all grid points. Analogously to the BSS, a skill score based on the CRPS is obtained via

$$CRPSS = 1 - \frac{CRPS(CONV + RAD)}{CRPS(CONV)}. \tag{21}$$

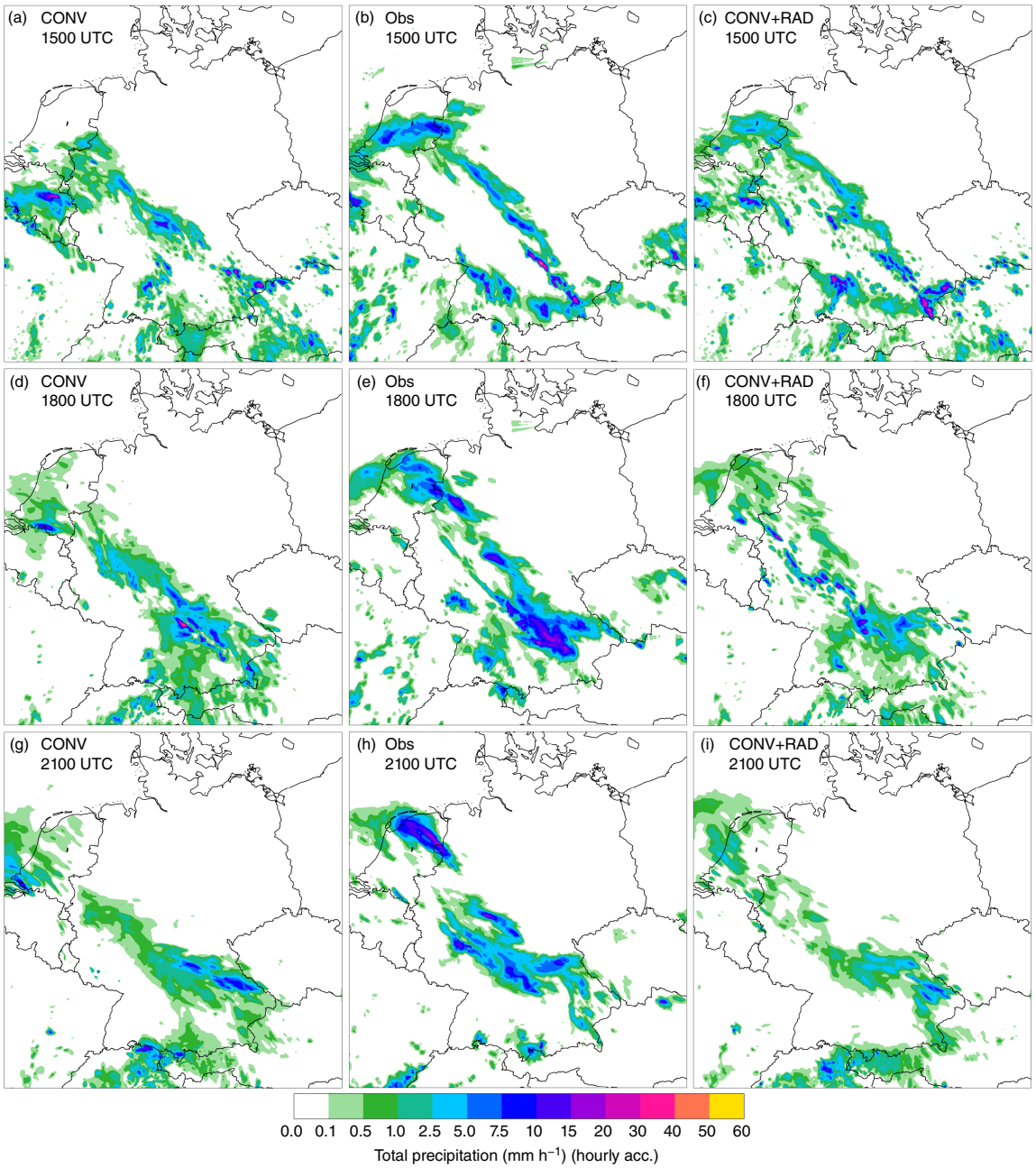

Figure 4. Hourly precipitation accumulation for experiment 1 (26 May 2014). Panels (b, e, h) show the radar-derived precipitation, (a, d, g) the model-generated precipitation of set-up CONV, and (c, f, i) the model-generated precipitation of set-up CONV+RAD. (a–c) are valid at 1500 UTC, (d–f) at 1800 UTC and (g–i) at 2100 UTC.

Table 2. Summary of the experimental set-ups.

	Radar data assimilated by	Update interval (min)
CONV	Not used	60
CONV+RAD	LETKF (3D)	60
CONV+RAD_60	LETKF (3D)	60
CONV+RAD_30	LETKF (3D)	30
CONV+RAD_15	LETKF (3D)	15
CONV+RAD_5	LETKF (3D)	5
CONV+LHN	LHN (2D)	60

4.3.2. *Experiment 3*

The deterministic forecasts in experiment 3 are compared to radar-derived precipitation by the Frequency BIas (FBI; Wilks, 2006) and by the Fraction Skill Score (FSS; Roberts and Lean, 2008). The FSS is also a skill score of the form

$$FSS = 1 - \frac{MSE}{MSE_{\text{ref}}}. \tag{22}$$

where MSE denotes the mean squared error and is calculated by the following steps: first, observations and model output are transferred into binary fields based on the exceedance of a threshold. For each grid point, neighbouring points within a certain distance s are averaged in the model and the observation field. Since experiment 3 performs verification based on accumulated precipitation, averaging takes place on the native model grid and neighbourhood areas are determined by adjacent model grid points that lie within a certain distance based on s. Based on these fractions, the MSE is defined as:

$$MSE = \frac{1}{n}\sum_{k=1}^{n}\left\{y(s)_k - x(s)_k\right\}^2. \tag{23}$$

The reference MSE is calculated as the 'worst possible MSE' that can be calculated from the underlying field and is obtained via

$$MSE_{\text{ref}} = \frac{1}{n}\sum_{k=1}^{n} y(s)_k^2 + \frac{1}{n}\sum_{k=1}^{n} x(s)_k^2. \tag{24}$$

Table 3. 2 × 2 contingency table based on yes/no outcomes of observation and forecast with notations for (a) hits, (b) false alarms, (c) misses and (d) correct negatives.

		Observation	
		Yes	No
Forecast	Yes	a	b
	No	c	d

The FBI is based on the contingency table (Table 3) which consists of frequencies of non-probabilistic *yes/no* forecasts and observations. In terms of precipitation, this binary outcome is defined by the exceedance of a threshold. The FBI is defined as the ratio of *yes* forecasts to *yes* observations,

$$FBI = \frac{a+b}{a+c}, \tag{25}$$

which is 1 for unbiased forecasts. Values below 1 indicate that the event was forecast less often than observed (underforecasting), and vice versa (overforecasting).

The deterministic forecasts are also compared against SYNOP station measurements. For this purpose, the Root Mean Squared Error (RMSE) is calculated according to

$$RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(x_k - y_k)^2}. \tag{26}$$

The RMSE ranges from 0 to ∞ with perfect score $RMSE = 0$.

4.3.3. *Experiments 1, 2 and 3*

A common method to assess the uncertainty in the previously defined scores is the bootstrap method (Efron and Tibshirani, 1993). From the original dataset, m samples are drawn with replacement. Each sample is of the same size as the original dataset. For these so-called bootstrap samples, the statistic of interest (e.g. the BSS) is calculated. Then, percentiles of the m realizations of the statistic of interest are used to estimate the uncertainty. In this study, the bootstrap samples consist of 1000 realizations. In the case of experiment 1 and 2, bootstrapping is applied to the ensemble, and in experiment 3, bootstrapping is applied to the set of multiple deterministic forecasts. The error bars shown in section 5 are obtained from the 2.5 and 97.5% percentiles of the bootstrap samples.

5. Results

5.1. *Experiment 1*

The first experiment studies the impact of radar reflectivity and conventional data assimilated via KENDA compared with the assimilation of conventional data only. Figure 4 displays hourly accumulated precipitation from observations and from the model (deterministic run). Panels (a)–(c) show CONV, the reference, and CONV+RAD at 1500 UTC compared to the observation, i.e. accumulated precipitation from 1400 to 1500 UTC immediately before the last assimilation step. By visual comparison, CONV+RAD is clearly in better agreement with the observation (especially over the Netherlands or southern Germany). Panels (d)–(f) display hourly accumulated precipitation valid at 1800 UTC (i.e. after 3 h forecast). CONV+RAD is still better than CONV; e.g. the cells over the western part of Germany seem to better match the observation. After 6 h of forecast (Figure 4(g–i)), it is difficult to judge by visual comparison, whether one set-up is superior. Both forecasts underestimate the intensity of the event.

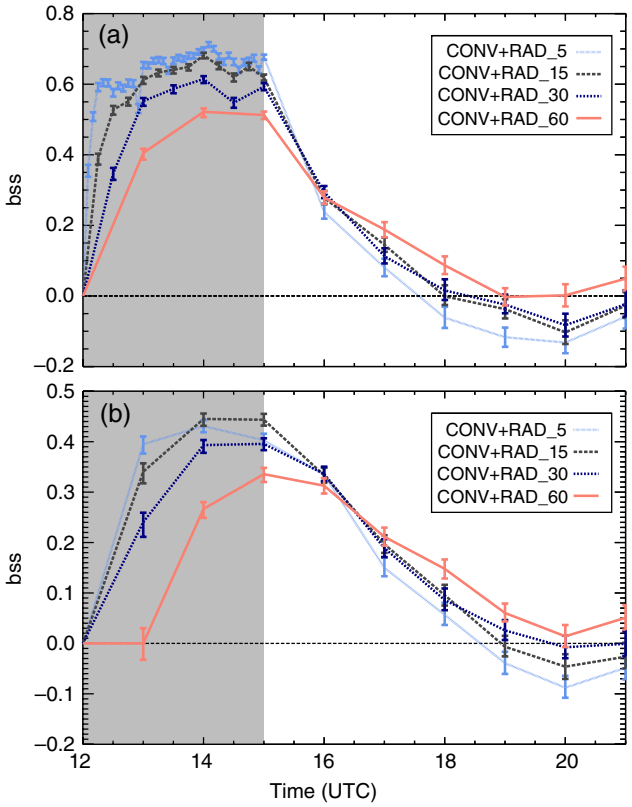

Figure 5. The Brier Skill Score (BSS) for set-ups CONV+RAD_5 (light blue fine dashes), CONV+RAD_15 (dark grey dashed), CONV+RAD_30 (dark blue dotted), and CONV+RAD_60 (red solid), where CONV is taken as the reference. The BSS is plotted against time, where the assimilation window is highlighted by the grey shaded area. The error bars are obtained via bootstrapping (2.5 and 97.5%iles). In (a), the BSS is based on instantaneous reflectivity measurements with a threshold of 20 dBZ. Technically, a dBZ model equivalent is not directly written by the analysis. Here, the analysis is approximated by a 5 min forecast during the assimilation window. For graphical display, these approximations are shown at their respective analysis times (i.e. every 5, 15, 30 or 60 min). In (b), the BSS is based on hourly accumulated precipitation with a threshold of 0.5 mm h^{-1}.

For an objective comparison, the BSS and CRPSS are computed. Due to technical constraints, precipitation is not contained in the analysis, thus the model needs to be run for a short period of time (we use 5 min) to generate the dBZ model equivalent.

The BSS of CONV+RAD with CONV as reference (solid red line in Figure 5) exhibits a strong increase during the assimilation window, indicating a strong positive impact of the radar reflectivity assimilation. This is visible both in terms of reflectivities (Figure 5(a)) and accumulated precipitation (Figure 5(b)). After the assimilation window, skill decreases but remains significantly positive for the first three forecast hours based on reflectivities. For accumulated precipitation, the results indicate an even longer positive impact.

Similar to the BSS, the CRPSS of CONV+RAD with CONV as reference (solid red line in Figure 6) shows a significant increase of skill during the assimilation window again visible in terms of reflectivities (Figure 6(a)) and accumulated precipitation (Figure 6(b)). In both Figures 5 and 6, a positive impact is visible for up to 4 h.

5.2. *Experiment 2*

Experiment 2 studies the impact of different update rates on ensemble spread, noise in the model state, and quality of analysis and forecast in terms of the BSS and CRPSS.

Figure 7 compares the number of ensemble members exceeding 0.5 mm h^{-1} of accumulated precipitation for CONV+RAD_5/15/30/60 (Figure 7(b, c, e, f)) at the end of the assimilation window (1500 UTC), with the corresponding observation (Figure 7(a, d)).

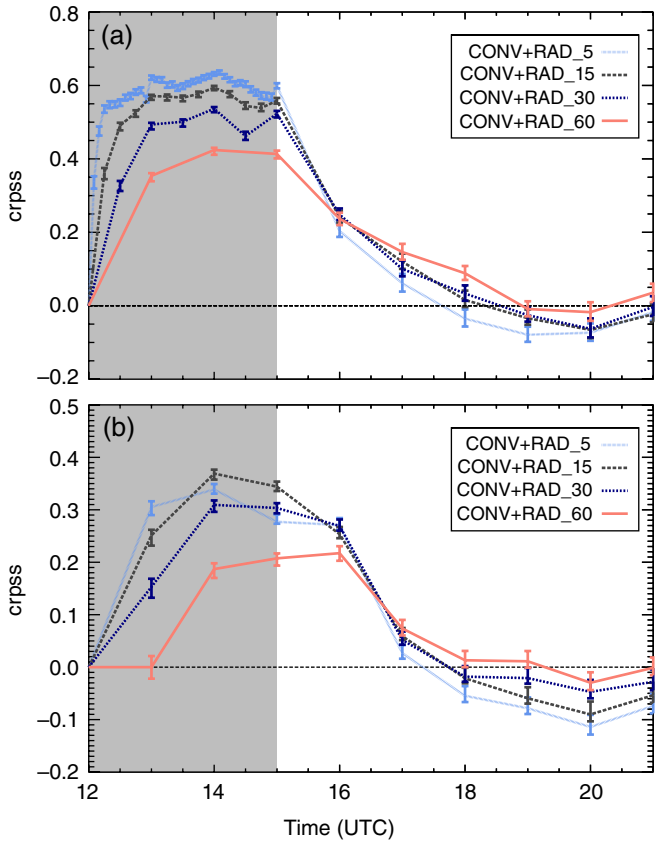

Figure 6. As Figure 5, but for the Continuous Ranked Probability Skill Score (CRPSS).

For all update frequencies, the assimilation picks up the observations very well, i.e. KENDA is able to move the model into a state where it fits the precipitation distribution to a high degree of accuracy. CONV+RAD_5/15/30 lead to very similar high numbers of ensemble members which reproduce observed precipitation rates above the threshold. Especially for CONV+RAD_5 and CONV+RAD_15, the red areas practically coincide with the observation. The ensemble members differ only outside the radar-observed precipitation area. In CONV+RAD_60, the ensemble is further away from the observation and has more spread. This is not necessarily a disadvantage; the analysis increments are a linear combination of the ensemble perturbations, i.e. ensemble spread is required in order to successfully assimilate the observations. However, in all four set-ups, the ensemble members have similar precipitation system structures indicated by the large red areas.

During assimilation, the BSS (Figure 5 against set-up CONV) clearly improves for all assimilation frequencies. More frequent updates lead to faster improvements. For reflectivities (Figure 5(a)), there is a hierarchy among the four set-ups towards the end of the assimilation window, where the most frequent updates are able to draw the model states closer to the observations. All four set-ups are able to improve the analyses to a high degree compared to CONV. The same holds for accumulated precipitation (Figure 5(b)). Towards the end of the assimilation window, CONV+RAD_5 seems to perform slightly worse than CONV+RAD_15. CONV+RAD_60 again yields the worst analysis among the four CONV+RAD set-ups. During the forecast window, the 1 h forecasts of all four set-ups yield almost identical scores both visible in Figure 5(a) and (b). For forecast horizons of 3 h or longer, the less frequent update set-ups are better able to maintain the information gained during the assimilation than the more frequent update set-ups. The most frequent updates even seem to harm the forecast compared to CONV, as indicated by the negative BSS.

A similar behaviour is also shown by the CRPSS (Figure 6). More frequent updates lead to faster improvements during the assimilation window (Figure 6(a)). Among the four set-ups, the analysis of CONV+RAD_60 is worst, but the forecast quality decreases quicker the higher the update rate.

Data assimilation always introduces numerical noise into the model state. This noise arises if structures in the analysis increments, especially small-scale structures, are incompatible with the actual dynamics of the model and if the model is not able to dissolve mass–momentum imbalances (Stauffer and Seaman, 1990). A common measure to quantify these imbalances is the first time derivative of surface pressure, averaged over the model domain (Stauffer and Seaman, 1990; Chen and Huang, 2006; Reen, 2007):

$$S_t = \frac{1}{n\,m}\sum_{i=1}^{n}\sum_{j=1}^{m}\left|\frac{\partial p_s}{\partial t}\right|_{ij}, \qquad (27)$$

where p_s is the surface pressure and summation is done over the entire model domain. In Figure 8, the evolution of S_t of the deterministic runs of CONV+RAD_5, CONV+RAD_15, CONV+RAD_30, and CONV+RAD_60 during the assimilation window is shown.

Updating the atmospheric model state less frequently causes a more intense shock and produces higher noise than updating the state more frequently. This behaviour is reasonable since, at less frequent updates, the model may be further away from the observation and the analysis introduces larger changes. However, only in CONV+RAD_60, does the noise in the model state decay almost to zero before the next assimilation, i.e. the model state is able to adjust to the changes introduced by data assimilation. Especially for CONV+RAD_5 and CONV+RAD_15, the model state is not able to completely recover from the analysis and remains at a higher noise level throughout the assimilation window. This persistent noise suggests that the analyses of the frequently updated set-ups are physically less consistent which may cause poorer forecast quality.

The data assimilation cycling script BACY (cf. section 2.4) alternately calls the model and the data assimilation module via shell scripts. This means that COSMO-DE and KENDA communicate via their respective output files and, after each LETKF-step, COSMO is newly initialized based on a restricted set of model fields. This is in contrast to a restart set-up where restart files save information about the whole model state. Thus noise also occurs without assimilating any observations. However, this noise (i.e. its influence on the surface pressure tendency) is approximately a factor of 10 smaller than that observed in the CONV+RAD set-up (not shown), thus the main contribution to the noise in Figure 8 comes from the assimilation of observations.

Since CONV+RAD_60 yields best results for lead times beyond 1 h, an update interval of 60 min is chosen for Experiment 3 discussed in the next section. Experiment 3 does not focus on analysis quality but on the quality of precipitation forecasts with lead times up to 24 h.

5.3. *Experiment 3*

The third experiment compares the forecast quality of the assimilation of radar reflectivities plus conventional observations based on the LETKF via KENDA against the assimilation of conventional data only and against the assimilation of conventional data via KENDA plus the assimilation of radar-derived precipitation rates by LHN.

For this experiment, 29 deterministic forecasts initialized through one week are evaluated. The FSS is calculated with a threshold of 0.5 mm h^{-1} and a neighbourhood area of five grid points in each direction, i.e. a neighbourhood window diameter of 11 grid points. During the first four forecast hours, a significant improvement of the radar data assimilation via CONV+RAD over the forecasts based on CONV is visible in the FSS indicated by the positive values of FSS differences in Figure 9(a). The results

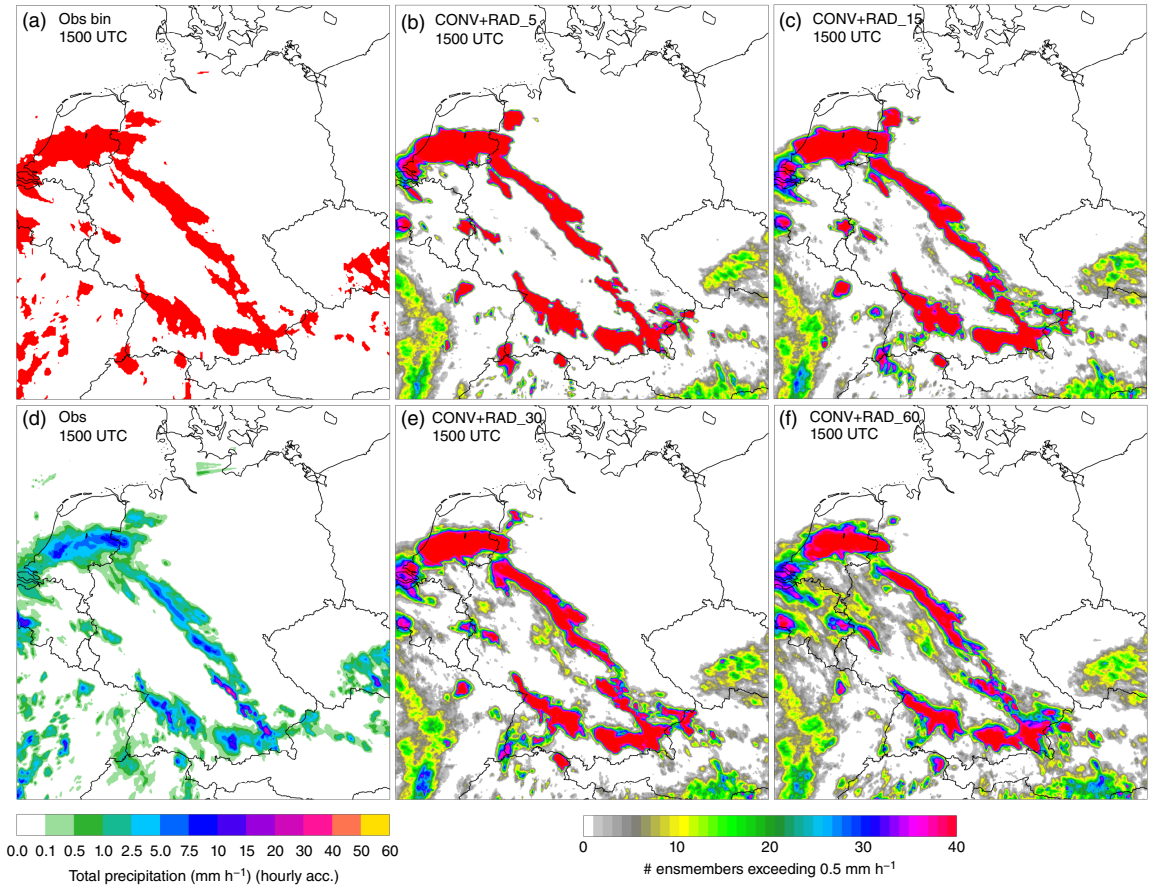

Figure 7. First-guess ensemble and observation of experiment 2 (26 May 2014) valid at 1500 UTC. In (a), the observation is transformed to a binary field: grid points exceeding 0.5 mm h^{-1} are shown in red. For the ensemble, the number of ensemble members exceeding 0.5 mm h^{-1} is counted for all four update frequencies: (b) CONV+RAD_5, (c) CONV+RAD_15, (e) CONV+RAD_30, and (f) CONV+RAD_60. As a reference, the observed total precipitation is shown in (d).

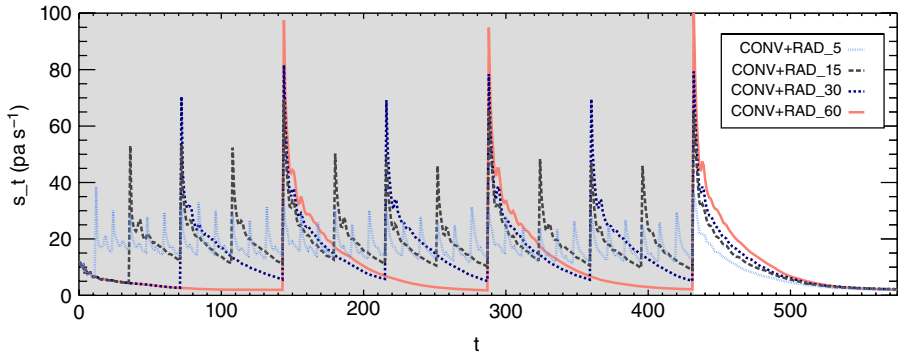

Figure 8. Domain-averaged surface pressure tendencies for each integration time step (1 time step ≈ 25 s, i.e. 3 h ≈ 432 time steps) during the assimilation window (grey shading) for CONV+RAD_5 (light blue fine dashes), CONV+RAD_15 (dark grey dashed), CONV+RAD_30 (dark blue dotted), and CONV+RAD_60 (red solid).

suggest a slight positive impact during forecast hours 6 to 10, though significance cannot be proven here. For longer lead times, the effect is neutral.

In Figure 9(b), the forecasts based on CONV+RAD are compared against the forecasts of the set-up CONV+LHN in terms of FSS. In the first forecast hours, a slight advantage of CONV+RAD on forecast quality over CONV+LHN is visible. Throughout the entire forecast lead time, the CONV+RAD set-up is able to compete with the CONV+LHN set-up as measured by FSS.

Figure 10 compares the FBI for CONV, CONV+RAD, and CONV+LHN and indicates an underestimation of events for all set-ups, but is most pronounced for CONV+RAD.

Measurements from SYNOP stations of 2 m temperature, 10 m horizontal wind and 2 m relative humidity are used to calculate the RMSE for forecasts based on all three set-ups (Figure 11). As expected, the RMSE for all variables increases with forecast time with slight fluctuations from hour to hour. Since the RMSE is based on forecasts with different initialization times during the day, these fluctuations can be attributed to the diurnal cycle of the RMSE. Errors are typically larger at noon (due to convective activity being highest) when the forecast started early in the day (not shown). The differences in RMSE of all variables are small compared to the associated uncertainty. CONV+RAD does not seem to degrade the forecast quality of 2 m temperature, 10 m horizontal wind and 2 m relative humidity.

6. Summary and discussion

In this study, 3D radar reflectivity measurements from the German radar network have been assimilated into the convection-permitting COSMO model with the ensemble Kalman filter KENDA. To achieve this, the radar forward operator EMVORADO has been integrated into the model and KENDA. Three data assimilation experiments addressed the

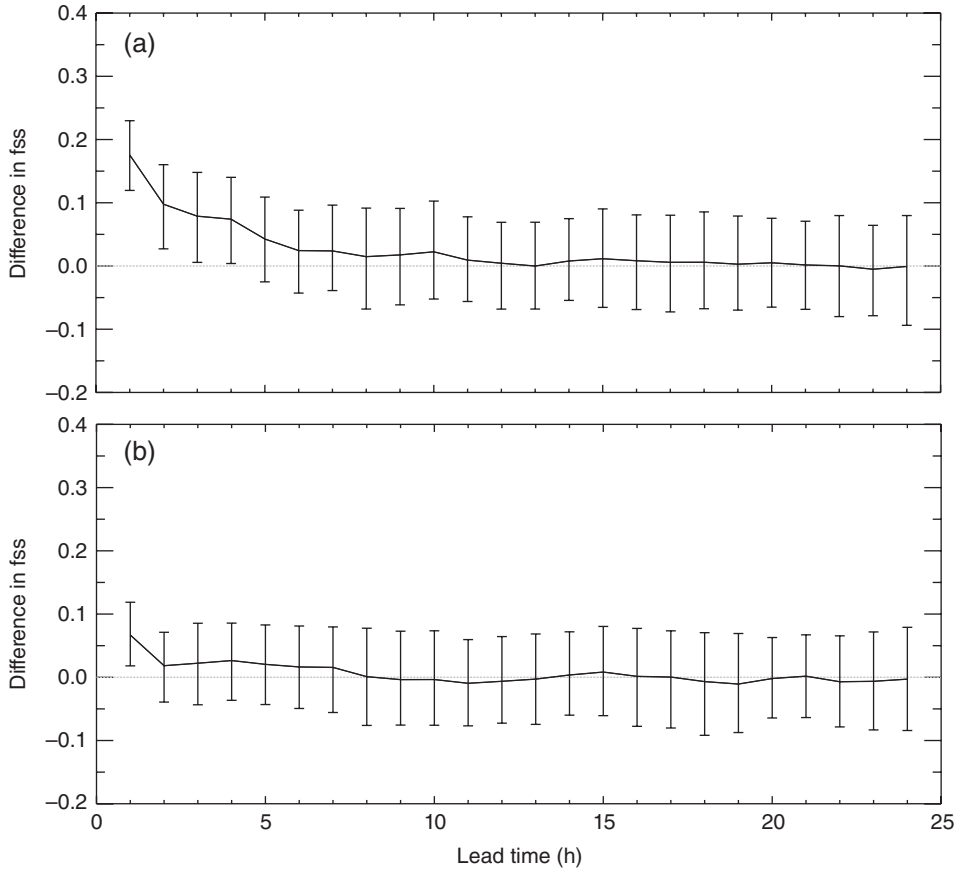

Figure 9. The difference in Fraction Skill Score (FSS) against forecast lead time over the 29 deterministic forecasts: (a) CONV+RAD minus CONV and (b) CONV+RAD minus CONV+LHN. The FSS is calculated for a neighbourhood of five grid points (in each horizontal direction) and a threshold of 0.5 mm h^{-1}. The error bars are obtained via bootstrapping (2.5 and 97.5%iles).

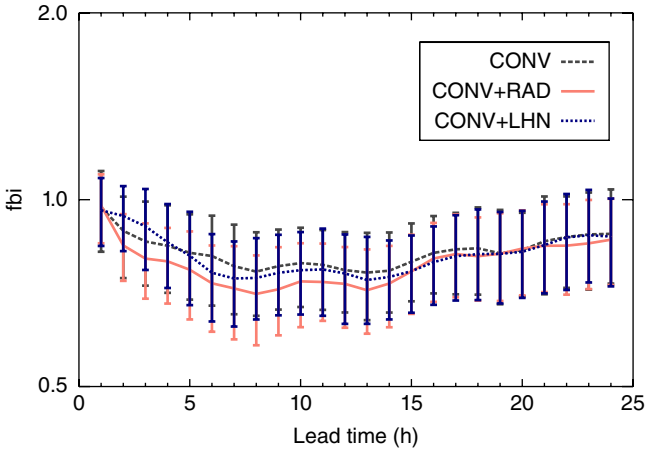

Figure 10. The FBI based on the forecasts of CONV (grey dashed), CONV+RAD (red solid) and CONV+LHN (blue dotted). Scores are calculated for a threshold of 0.5 mm h^{-1}. The error bars are obtained via bootstrapping (2.5 and 97.5%iles).

influence of the assimilation of radar reflectivity on the quality of analysis and forecasts of precipitation and the influence of the assimilation frequency. Finally, the impact of assimilating 3D radar reflectivities in the ensemble Kalman filter was compared to results obtained when reflectivity assimilation was replaced by LHN based on radar-derived surface precipitation fields.

3D radar data have a clear positive impact when assimilated in addition to conventional observations (Experiments 1 and 3). Experiment 3 also shows that the positive impact of our experimental set-up is at least comparable to LHN, which is the operational data assimilation scheme for radar-derived precipitation at DWD. During the first forecast hours, the KENDA-based assimilation of reflectivities performs even slightly better than LHN, though this needs to be confirmed by longer periods and other seasons for stronger statistical significance.

The comparison with SYNOP station data shows (within the limits of the available case-studies) that the assimilation of radar reflectivities does not degrade the forecast of other prognostic model variables such as temperature, wind, and relative humidity.

Despite the distribution of radar reflectivity data being non-Gaussian (which violates basic Kalman filter assumptions), the quality of both analysis and forecast of precipitation are strongly improved. It is remarkable that, even at this early stage of the development, the 3D radar reflectivity data assimilation via the LETKF already challenges the well-tested and well-tuned LHN scheme.

Experiment 2 tested the impact of various update frequencies (5, 15, 30, or 60 min) and revealed that the model is dragged closer to the observation with high update frequencies, while the forecast quality for lead times beyond 1 h suffered. This well-known phenomenon results from two competeting goals in data assimilation, namely the achievement of a physically consistent state and having a state close to the observations. Interestingly, previous radar data assimilation studies favoured more frequent update rates for very short-term predictions. However, our study suggests that less frequent updates, although producing less accurate analyses, lead to more balanced model states which eventually lead to better forecasts for lead times above 1 h. These results are in line with tests in an idealized set-up described in Lange and Craig (2014).

The results strongly encourage the assimilation of 3D radar reflectivity data for operational short-range NWP and for bridging the gap between nowcasting and NWP. However, further investigations are necessary to corroborate or improve our results on rapid update cycling, because more frequent assimilation improves the forecast in the first forecast hour. Also the superobbing strategy may be improved, e.g. by superobbing

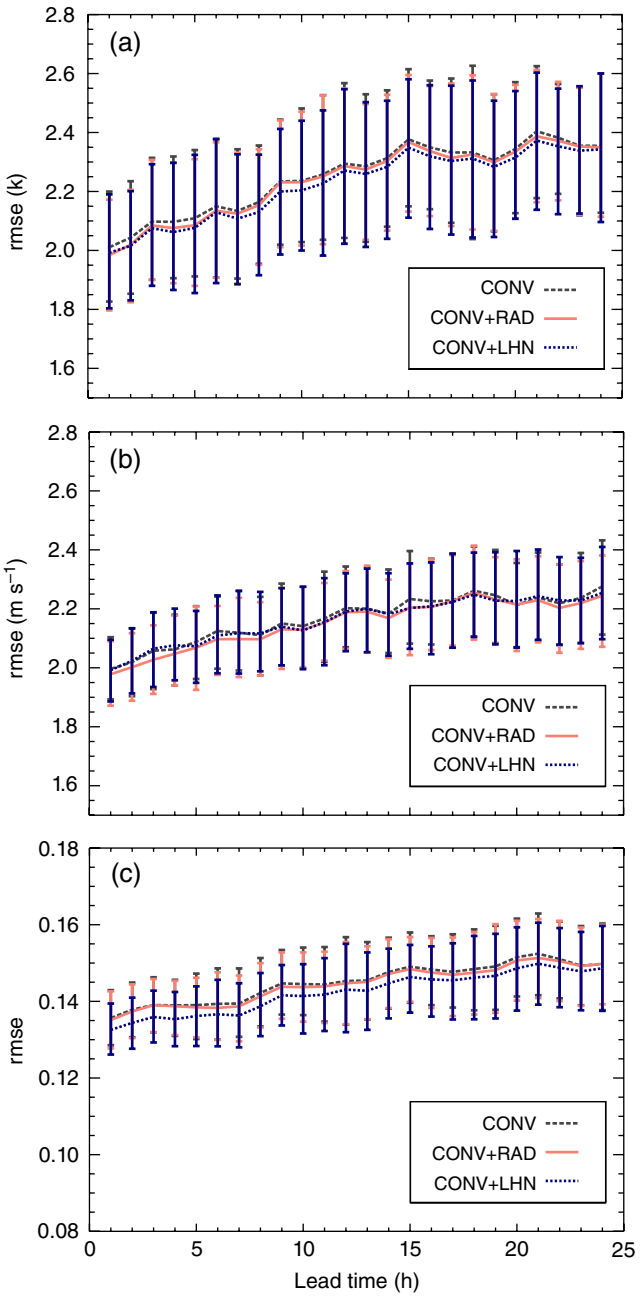

Figure 11. The RMSE of forecasts against observations at ~840 SYNOP stations for (a) 2 m temperature, (b) *u*-component of 10 m wind, and (c) 2 m relative humidity for CONV (grey dashed), CONV+RAD (red solid) and CONV+LHN (blue dotted). The error bars are obtained via bootstrapping (2.5 and 97.5%iles).

in the vertical or by a better treatment of horizontally overlapping areas within the radar network. The appropriate weights of different observation types – conventional and radar – in the assimilation step needs further investigation. Due to their different densities, refinements in localization in general, and adaptive localization in particular, might lead to further improvements. Further optimization steps are necessary to achieve sufficient speed for operational use. Finally, the assimilation of radial winds within KENDA, as already prepared by Zeng (2013), needs to be extensively tested and combined with reflectivity assimilation.

Acknowledgements

This study was performed within the framework of the Atmospheric Dynamics and Predictability Branch of the Hans Ertel Centre for Weather Research (HErZ; Weissmann *et al.*, 2014; Simmer *et al.*, 2015) funded by the BMVI (Federal Ministry of Transport and Digital Infrastructures). We also acknowledge substantial financial support of Theresa Bick by DWD, the Transregional Collaborative Research Center 32 (TR32) on Patterns in Soil-Vegetation-Atmosphere Systems – Monitoring, Modeling and Data Assimilation, and the Centre for High-Performance Scientific Computing in Terrestrial Systems of ABC/J Geoverbund. We also like to thank the Associate Editor Peter Houtekamer and two anonymous reviewers for their valuable comments.

References

Ades M, van Leeuwen PJ. 2012. An exploration of the equivalent weights particle filter. *Q. J. R. Meteorol. Soc.* **28**: 2–28, doi: 10.1002/qj.1995.

Aksoy A, Dowell DC, Snyder C. 2009. A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analyses. *Mon. Weather Rev.* **137**: 1805–1824, doi: 10.1175/2008MWR2691.1.

Aksoy A, Dowell DC, Snyder C. 2010. A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part II: Short-range ensemble forecasts. *Mon. Weather Rev.* **138**: 1273–1292, doi: 10.1175/2009MWR3086.1.

Anderson JL, Anderson SL. 1999. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Weather Rev.* **127**: 2741–2758, doi: 10.1175/1520-0493(1999)127<2741:amciot>2.0.CO;2.

Bachner S, Kapala A, Simmer C. 2008. Evaluation of daily precipitation characteristics in the CLM and their sensitivity to parameterizations. *Meteorol. Z.* **17**: 407–419, doi: 10.1127/0941-2948/2008/0300.

Baldauf M, Seifert A, Förstner J, Majewski D, Raschendorfer M, Reinhardt T. 2011. Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities. *Mon. Weather Rev.* **139**: 3887–3905, doi: 10.1175/MWR-D-10-05013.1.

Bishop CH, Etherton BJ, Majumdar SJ. 2001. Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Weather Rev.* **129**: 420–436, doi: 10.1175/1520-0493(2001)129<0420:aswtet>2.0.CO;2.

Blahak U. 2008a. An approximation to the effective beam weighting function for scanning meteorological radars with axisymmetric antenna pattern. *J. Atmos. Oceanic Technol.* **25**: 1182–1196, doi: 10.1175/2007JTECHA1010.1.

Blahak U. 2008b. 'RADAR_MIE_LM and RADAR_MIELIB – calculation of radar reflectivity from model output'. Internal report. Institute of Meteorology and Climate Research, KIT: Karlsruhe, Germany. urlulrich.blahak@dwd.de (accessed 7 February 2016).

Blahak U, Zeng Y, Epperlein D. 2011. 'Radar forward operator for data assimilation and model verification for the COSMO model'. In Extendend abstract in *Proceedings of 35th Conference on Radar Meteorology*, 26–30 September 2011. Pittsburgh, PA. https://ams.confex.com/ams/35Radar/webprogram/Handout/Paper192019/ext_abstract_blahak.pdf (accessed 6 February 2016).

Burgers G, van Leeuwen PJ, Evensen G. 1998. Analysis scheme in the ensemble Kalman filter. *Mon. Weather Rev.* **126**: 1719–1724, doi: 10.1175/1520-0493(1998)126<1719:asitek>2.0.CO;2.

Caya A, Sun J, Snyder C. 2005. A comparison between the 4D-VAR and the ensemble Kalman filter techniques for radar data assimilation. *Mon. Weather Rev.* **133**: 3081–3094, doi: 10.1175/MWR3021.1.

Chang W, Chung KS, Fillion L, Baek SJ. 2014. Radar data assimilation in the Canadian high-resolution ensemble Kalman filter system: Performance and verification with real summer cases. *Mon. Weather Rev.* **142**: 2118–2138, doi: 10.1175/MWR-D-13-00291.1.

Chen M, Huang XY. 2006. Digital filter initialization for MM5. *Mon. Weather Rev.* **134**: 1222–1236, doi: 10.1175/MWR3117.1.

Doms G, Schättler U. 2002. A description of the non-hydrostatic regional model LM, Part I: Dynamics and numerics. *COSMO Newslett.* **2**: 225–235.

Doms G, Förstner J, Heise E, Herzog H-J, Mironov D, Raschendorfer M, Reinhardt T, Ritter B, Schrodin R, Schulz J-P, Vogel G. 2011. 'A description of the nonhydrostatic regional COSMO model. Part II: Physical parameterization'. http://www.cosmo-model.org/content/model/documentation/core/cosmoPhysParamtr.pdf (accessed 6 February 2016).

Dong J, Xue M. 2013. Assimilation of radial velocity and reflectivity data from coastal WSR-88D radars using an ensemble Kalman filter for the analysis and forecast of landfalling hurricane *Ike* (2008). *Q. J. R. Meteorol. Soc.* **139**: 467–487, doi: 10.1002/qj.1970.

Dong J, Xue M, Droegemeier K. 2011. The analysis and impact of simulated high-resolution surface observations in addition to radar data for convective storms with an ensemble Kalman filter. *Meteorol. Atmos. Phys.* **112**: 41–61, doi: 10.1007/s00703-011-0130-3.

Dowell DC, Wicker LJ. 2009. Additive noise for storm-scale ensemble data assimilation. *J. Atmos. Oceanic Technol.* **26**: 911–927, doi: 10.1175/2008JTECHA1156.1.

Dowell DC, Zhang F, Wicker LJ, Snyder C, Crook NA. 2004. Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: Ensemble Kalman filter experiments. *Mon. Weather Rev.* **132**: 1982–2005, doi: 10.1175/1520-0493(2004)132<1982:watrit>2.0.CO;2.

Dowell DC, Wicker LJ, Snyder C. 2011. Ensemble Kalman filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Mon. Weather Rev.* **139**: 272–294, doi: 10.1175/2010MWR3438.1.

Efron B, Tibshirani RJ. 1993. *An Introduction to the Bootstrap*. Chapman & Hall/CRC: Abingdon, UK.

Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **99**: 10143–10162, doi: 10.1029/94JC00572.

Fabry F, Szyrmer W. 1999. Modeling of the melting layer. Part II: Electromagnetic. *J. Atmos. Sci.* **56**: 3593–3600.

Freitag MA, Potthast R. 2013. Synergy of inverse problems and data assimilation techniques. In *Large-Scale Inverse Problems*, Cullen M, Freitag MA, Kindermann S, Scheichl R. (eds.): 1–54. de Gruyter: Berlin, doi: 10.1515/9783110282269.1.

Gao J, Xue M. 2008. An efficient dual-resolution approach for ensemble data assimilation and tests with simulated Doppler radar data. *Mon. Weather Rev.* **136**: 945–963, doi: 10.1175/2007MWR2120.1.

Gaspari G, Cohn SE. 1999. Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.* **125**: 723–757, doi: 10.1002/qj.49712555417.

Gneiting T, Raftery AE. 2007. Strictly proper scoring rules, prediction, and estimation. *J. Am. Stat. Assoc.* **102**: 359–378, doi: 10.1198/016214506000001437.

Grasselt R, Schüttemeyer D, Warrach-Sagi K, Ament F, Simmer C. 2008. Validation of TERRA-ML with discharge measurements. *Meteorol. Z.* **17**: 763–773, doi: 10.1127/0941-2948/2008/0334.

Greybush SJ, Kalnay E, Miyoshi T, Ide K, Hunt BR. 2011. Balance and ensemble Kalman filter localization techniques. *Mon. Weather Rev.* **139**: 511–522, doi: 10.1175/2010MWR3328.1.

Grimit EP, Gneiting T, Berrocal VJ, Johnson NA. 2006. The continuous ranked probability score for circular variables and its application to mesoscale forecast ensemble verification. *Q. J. R. Meteorol. Soc.* **132**: 2925–2942, doi: 10.1256/qj.05.235.

Harnisch F, Keil C. 2015. Initial conditions for convective-scale ensemble forecasting provided by ensemble data assimilation. *Mon. Weather Rev.* **143**: 1583–1600, doi: 10.1175/MWR-D-14-00209.1.

Houtekamer P, Mitchell HL. 2005. Ensemble Kalman filtering. *Q. J. R. Meteorol. Soc.* **131**: 3269–3289, doi: 10.1256/qj.05.135.

Houtekamer PL, Mitchell HL. 1998. Data assimilation using an ensemble Kalman filter technique. *Mon. Weather Rev.* **126**: 796–811, doi: 10.1175/1520-0493(1998)126<0796:dauaek>2.0.CO;2.

Hunt B, Kostelich EJ, Szunyogh I. 2007. Efficient data assimilation for spatiotemporal chaos: A local transform Kalman filter. *Physica D* **230**: 112–126, doi: 10.1016/j.physd.2006.11.008.

Janjić T, Nerger L, Albertella A, Schroeter J, Skachko S. 2011. On domain localization in ensemble-based Kalman filter algorithms. *Mon. Weather Rev.* **139**: 2046–2060, doi: 10.1175/2011MWR3552.1.

Jerger D. 2013. 'Radar forward operator for verification of cloud-resolving simulations within the COSMO model', Dissertation. IMK-TRO, Department of Physics, KIT: Karlsruhe, Germany. http://digbib.ubka.uni-karlsruhe.de/volltexte/1000038411 (accessed 6 February 2016).

Jung Y, Xue M, Tong M. 2012. Ensemble Kalman filter analyses of the 29–30 May 2004 Oklahoma tornadic thunderstorm using one- and two-moment bulk microphysics schemes, with verification against polarimetric radar data. *Mon. Weather Rev.* **140**: 1457–1475, doi: 10.1175/MWR-D-11-00032.1.

Kerker M. 1969. *The Scattering of Light and other Electromagnetic Radiation*. Academic Press: New York, NY.

Kirchgessner P, Nerger L, Bunse-Gerstner A. 2014. On the choice of an optimal localization radius in ensemble Kalman filter methods. *Mon. Weather Rev.* **142**: 2165–2175, doi: 10.1175/MWR-D-13-00246.1.

Lange H, Craig GC. 2014. The impact of data assimilation length scales on analysis and prediction of convective storms. *Mon. Weather Rev.* **142**: 3781–3808, doi: 10.1175/MWR-D-13-00304.1.

van Leeuwen PJ. 2009. Particle filtering in geophysical systems. *Mon. Weather Rev.* **137**: 4089–4114, doi: 10.1175/2009MWR2835.1.

Lilly DK. 1990. Numerical prediction of thunderstorms – has its time come? *Q. J. R. Meteorol. Soc.* **116**: 779–798, doi: 10.1002/qj.49711649402.

Lin YL, Farley RD, Orville HD. 1983. Bulk parameterization of the snow field in a cloud model. *J. Clim. Appl. Meteorol.* **22**: 1065–1092, doi: 10.1175/1520-0450(1983)0221065:bpotsf2.0.CO;2.

Lin C, Vasi S, Kilambi A, Turner B, Zawadzki I. 2005. Precipitation forecast skill of numerical weather prediction models and radar nowcasts. *Geophys. Res. Lett.* **32**: L14801, doi: 10.1029/2005GL023451.

Liu ZQ, Rabier F. 2002. The interaction between model resolution, observation resolution and observation density in data assimilation: A one-dimensional study. *Q. J. R. Meteorol. Soc.* **128**: 1367–1386. doi: 10.1256/003590002320373337.

Marshall JS, Palmer WMK. 1948. The distribution of raindrops with size. *J. Meteorol.* **5**: 165–166, doi: 10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2.

Matheson JE, Winkler RL. 1976. Scoring rules for continuous probability distributions. *Manage. Sci.* **22**: 1087–1096.

Maxwell Garnett JC. 1904. Colours in metal glasses and in metallic films. *Philos. Trans. R. Soc. London Ser. A* **203**: 385–420.

Milan M, Venema V, Schüttemeyer D, Simmer C. 2008. Assimilation of radar and satellite data in mesoscale models: A physical initialization scheme. *Meteorol. Z.* **17**: 887–902, doi: 10.1127/0941-2948/2008/0340.

Milan M, Schüttemeyer D, Bick T, Simmer C. 2014. A sequential ensemble prediction system at convection-permitting scales. *Meteorol. Atmos. Phys.* **123**: 17–31, doi: 10.1007/s00703-013-0291-3.

Miyoshi T, Yamane S, Enomoto T. 2007. Localizing the error covariance by physical distances within a local ensemble transform Kalman filter (LETKF). *SOLA* **3**: 89–92, doi: 10.2151/sola.2007-023.

Nichols NK. 2010. Mathematical concepts of data assimilation. In *Data Assimilation – Making Sense of Observations*, Lahoz W, Khattatov B, Menard R. (eds.): 13–39. Springer: Berlin.

Perianez A, Reich H, Potthast R. 2014. Optimal localization for ensemble Kalman filter systems. *J. Meteorol. Soc. Jpn.* **92**: 585–597, doi: 10.2151/jmsj.

Raschendorfer M. 2001. The new turbulence parameterisation of LM. *COSMO Newslett.* **1**: 89–97.

Reen BP. 2007. 'Data assimilation strategies and land-surface heterogeneity effects in the planetary boundary layer', Dissertation. Pennsylvania State University: State College, PA.

Reinhardt T, Seifert A. 2006. A three-category ice scheme for LMK. *COSMO Newslett.* **6**: 115–120.

Ritter B, Geleyn JF. 1992. A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. *Mon. Weather Rev.* **120**: 303–325, doi: 10.1175/1520-0493(1992)120<0303:acrsfn>2.0.CO;2.

Roberts NM, Lean HW. 2008. Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Weather Rev.* **136**: 78–97, doi: 10.1175/2007MWR2123.1.

Schraff C, Reich H, Rhodin A, Schomburg A, Stephan K, Perianez A, Potthast R. 2016. Kilometre-scale ensemble data assimilation for the COSMO model (KENDA). *Q. J. R. Meteorol. Soc.*, doi: 10.1002/qj.2748.

Schwitalla T, Wulfmeyer V. 2014. Radar data assimilation experiments using the IPM WRF rapid update cycle. *Meteorol. Z.* **23**: 79–102, doi: 10.1127/0941-2948/2014/0513.

Seifert A, Beheng KD. 2001. A double-moment parameterization for simulating autoconversion, accretion and selfcollection. *Atmos. Res.* **59**: 265–281, doi: 10.1016/S0169-8095(01)00126-0.

Simmer C, Adrian G, Jones S, Wirth V, Göber M, Hohenegger C, Janjić T, Keller J, Ohlwein C, Seifert A, Trömel S, Ulbrich T, Wapler K, Weissmann M, Keller J, Masbou M, Meilinger S, Riss N, Schomburg A, Vormann A Weingärtner C. 2015. HErZ – the German Hans-Ertel Centre for Weather Research. *Bull. Am. Meteorol. Soc.*, doi: 10.1175/BAMS-D-13-00227.1.

Snook N, Xue M, Jung Y. 2011. Analysis of a tornadic mesoscale convective vortex based on ensemble Kalman filter assimilation of CASA X-band and WSR-88D radar data. *Mon. Weather Rev.* **139**: 3446–3468, doi: 10.1175/MWR-D-10-05053.1.

Snook N, Xue M, Jung Y. 2015. Multi-scale EnKF assimilation of radar and conventional observations and ensemble forecasting for a tornadic mesoscale convective system. *Mon. Weather Rev.* **143**: 1035–1057, doi: 10.1175/MWR-D-13-00262.1.

Snyder C, Zhang F. 2003. Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Weather Rev.* **131**: 1663–1677, doi: 10.1175//2555.1.

Sobash RA, Stensrud D. 2013. The impact of covariance localization for radar data on ENKF analyses of a developing MCS: Observing system simulation experiments. *Mon. Weather Rev.* **141**: 3691–3709, doi: 10.1175/MWR-D-12-00203.1.

Stauffer DR, Seaman NL. 1990. Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data. *Mon. Weather Rev.* **118**: 1250–1277, doi: 10.1175/1520-0493(1990)118<1250:uofdda> 2.0.CO;2.

Stephan K, Klink S, Schraff C. 2008. Assimilation of radar-derived rain rates into the convective-scale model COSMO-DE at DWD. *Q. J. R. Meteorol. Soc.* **134**: 1315–1326, doi: 10.1002/qj.269.

Sun J, Crook NA. 1997. Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and Its adjoint. Part I: Model development and simulated data experiments. *J. Atmos. Sci.* **54**: 1642–1661, doi: 10.1175/1520-0469(1997)054<1642:damrfd>2.0.CO;2.

Sun J, Crook NA. 1998. Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part II: Retrieval experiments of an observed Florida convective storm. *J. Atmos. Sci.* **55**: 835–852, doi: 10.1175/1520-0469(1998)055<0835:damrfd>2.0.CO;2.

Sun J, Xue M, Wilson JW, Zawadzki I, Ballard SP, Onvlee-Hooimeyer J, Joe P, Barker DM, Li PW, Golding B, Xu M, Pinto J. 2014. Use of NWP for nowcasting convective precipitation: Recent progress and challenges. *Bull. Am. Meteorol. Soc.* **95**: 409–426, doi: 10.1175/BAMS-D-11-00263.1.

Talagrand O. 1997. Assimilation of observations, an introduction. *J. Meteorol. Soc. Jpn.* **75**: 191–209.

Tiedtke M. 1989. A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Weather Rev.* **117**: 1779–1800, doi: 10.1175/1520-0493(1989)117<1779:acmfsf>2.0.CO;2.

Tong M, Xue M. 2005. Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Mon. Weather Rev.* **133**: 1789–1807, doi: 10.1175/MWR2898.1.

Wapler K, Harnisch F, Pardowitz T, Senf F. 2015. Characterisation and predictability of a strong and a weak forcing severe convective event a multi-data approach. *Meteorol. Z.* **24**: 393–410, doi: 10.1127/metz/2015/0625.

Weissmann M, Göber M, Hohenegger C, Janji T, Ohlwein C, Seifert A, Trömel S, Ulbrich T, Wapler K, Bollmeyer C, Deneke H. 2014. Initial phase of the Hans-Ertel Centre for Weather Research: A virtual centre at the interface of basic and applied weather and climate research. *Meteorol. Z.* **23**: 193–208, doi: 10.1127/0941-2948/2014/0558.

Whitaker JS, Hamill TM. 2012. Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Weather Rev.* **140**: 3078–3089, doi: 10.1175/MWR-D-11-00276.1.

Wicker LJ, Skamarock WC. 2002. Time-splitting methods for elastic models using forward time schemes. *Mon. Weather Rev.* **130**: 2088–2097, doi: 10.1175/1520-0493(2002)130<2088:tsmfem>2.0.CO;2.

Wilks DS. 2006. *Statistical Methods in the Atmospheric Sciences* (2nd edn) Academic Press/Elsevier: Cambridge, MA.

Xue M, Tong M, Droegemeier KK. 2006. An OSSE framework based on the ensemble square root Kalman filter for evaluating the impact of data from radar networks on thunderstorm analysis and forecasting. *J. Atmos. Oceanic Technol.* **23**: 46–66, doi: 10.1175/JTECH1835.1.

Yang SC, Kalnay E, Hunt B, Bowler NE. 2009. Weight interpolation for efficient data assimilation with the local ensemble transform Kalman filter. *Q. J. R. Meteorol. Soc.* **135**: 251–262, doi: 10.1002/qj.353.

Zängl G, Reinert D, Rpodas P, Baldauf M. 2015. The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Q. J. R. Meteorol. Soc.* **141**: 563–579, doi: 10.1002/qj.2378.

Zeng Y. 2013. 'Efficient radar forward operator for operational data assimilation within the COSMO-model', Dissertation. IMK-TRO 60. KIT Scientific Publishing: Karlsruhe, Germany, doi: 10.5445/KSP/10000 36921.

Zeng Y, Blahak U, Neuper M, Epperlein D. 2014. Radar beam tracing methods based on atmospheric refractive index. *J. Atmos. Oceanic Technol.* **31**: 2650–2670, doi: 10.1175/JTECH-D-13-00152.1.

Zhang F, Snyder C, Sun J. 2004. Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Mon. Weather Rev.* **132**: 1238–1253, doi: 10.1175/1520-0493(2004)132<1238:ioieao>2.0.CO;2.

Zhang F, Weng Y, Sippel JA, Meng Z, Bishop CH. 2009. Cloud-resolving hurricane initialization and prediction through assimilation of Doppler radar observations with an ensemble Kalman filter. *Mon. Weather Rev.* **137**: 2105–2125, doi: 10.1175/2009MWR2645.1.