



Comparison of Methods Accounting for Subgrid-Scale Model Error in Convective-Scale Data Assimilation

YUEFEI ZENG AND TIJANA JANJIĆ

Meteorologisches Institut, Ludwig-Maximilians-Universität München, Munich, Germany

ALBERTO DE LOZAR

Deutscher Wetterdienst, Offenbach, Germany

STEPHAN RASP

Meteorologisches Institut, Ludwig-Maximilians-Universität München, Munich, Germany

ULRICH BLAHAK AND AXEL SEIFERT

Deutscher Wetterdienst, Offenbach, Germany

GEORGE C. CRAIG

Meteorologisches Institut, Ludwig-Maximilians-Universität München, Munich, Germany

(Manuscript received 8 March 2019, in final form 25 March 2020)

ABSTRACT

Different approaches for representing model error due to unresolved scales and processes are compared in convective-scale data assimilation, including the physically based stochastic perturbation (PSP) scheme for turbulence, an advanced warm bubble approach that automatically detects and triggers absent convective cells, and additive noise based on model truncation error. The analysis of kinetic energy spectrum guides the understanding of differences in precipitation forecasts. It is found that the PSP scheme results in more ensemble spread in assimilation cycles, but its effects on the root-mean-square error (RMSE) are neutral. This leads to positive impacts on precipitation forecasts that last up to three hours. The warm bubble technique does not create more spread, but is effective in reducing the RMSE, and improving precipitation forecasts for up to 3 h. The additive noise approach contributes greatly to ensemble spread, but it results in a larger RMSE during assimilation cycles. Nevertheless, it considerably improves the skill of precipitation forecasts up to 6 h. Combining the additive noise with either the PSP scheme or the warm bubble technique reduces the RMSE within cycles and improves the skill of the precipitation forecasts, with the latter being more beneficial.

1. Introduction

An ensemble Kalman filter (EnKF, Evensen 1994) uses the flow dependency from an ensemble of background states to construct the background error covariance. However, a number of factors can affect the

soundness of the background error covariance, such as the validity of Gaussian assumptions, unbalanced correlations among variables, sampling error due to the limited size of ensemble and model error due to poorly resolved physical processes and physics parameterizations at subgrid scale. In the context of convective-scale data assimilation, it is important that the background error covariance captures uncertainties on the smallest resolvable scales as well as effects on the resolved scales of subgrid-scale uncertainties. Deficient representation of model error will usually lead to overconfidence of the

Denotes content that is immediately available upon publication as open access.

Corresponding author: Yuefei Zeng, yuefei.zeng@lmu.de

DOI: 10.1175/MWR-D-19-0064.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy \(www.ametsoc.org/PUBSReuseLicenses\)](https://www.ametsoc.org/PUBSReuseLicenses).

ensemble, eventually deteriorating the performance of convective-scale data assimilation and consequently the quality of the subsequent forecasts.

To account for model error, different forms of methods are available. For instance, relaxation methods such as the relaxation to prior perturbations (RTPP; [Zhang et al. 2004](#)) and relaxation to prior spread (RTPS; [Whitaker and Hamill 2012](#)) are usually used to take unknown sources of model error into account. However, these methods are ad hoc and may be associated with problems like smoothing of background errors or unbalanced model states as shown in [Zeng et al. \(2018\)](#). [Zeng et al. \(2019\)](#) supplemented the large-scale additive noise (LAN) derived from global climatological atmospheric background error covariance with the small-scale additive noise (SAN) using random samples of model truncation error. This combination of additive noise inherently provides information on large/synoptic-scale model uncertainties arising from the lateral boundary conditions and small/unresolved-scale error due to the limited resolution. [Zeng et al. \(2019\)](#) showed that use of this additive noise considerably improved 6-h precipitation forecasts, especially under weak forcing synoptic conditions. However, this method is not flow dependent.

In addition to these techniques, there are other methods initially used in probabilistic forecasting, such as stochastically perturbed parameterization tendencies (SPPT; [Buizza et al. 1999](#)), the multiphysics approach ([Murphy et al. 2004](#)), the stochastic kinetic energy backscatter (SKEB; [Shutts 2005](#)) and the physically based stochastic perturbation scheme (PSP; [Kober and Craig 2016](#); [Rasp et al. 2018](#)) for representing unresolved boundary layer turbulence. Although the SPPT and SKEB are usually implemented in the global models, they have recently also proved useful for some convection-permitting models, such as the SPPT in the Applications of Research to Operations at Mesoscale (AROME; [Bouttier et al. 2012](#)) model or the SKEB in the Weather Research and Forecasting (WRF) Model ([Berner et al. 2015](#); [Duda et al. 2016](#)) model. Furthermore, some methods mentioned above are applied more and more in data assimilation. For instance, the SPPT was incorporated into an ensemble transform data assimilation system by [Reynolds et al. \(2008\)](#) and a desired increase in initial variance in the tropics was achieved. [Fujita et al. \(2007\)](#) and [Meng and Zhang \(2008\)](#) showed in their mesoscale data assimilation experiments that better background error covariance and mean can be obtained by using multiphysics schemes. The SKEB scheme was used in ensemble data assimilation experiments by [Leutbecher et al. \(2007\)](#), however, it led to still underdispersive analysis ensembles.

There are numerous studies that compare the performance of different methods as well as various combinations of them. [Houtekamer et al. \(2009\)](#) showed that the additive isotropic model error perturbations are very effective to improve the innovation statistics and the multiphysics approach has relatively small but clearly positive impacts. Therefore, both are utilized in the operational EnKF for medium-range forecasts of the ensemble prediction system at the Meteorological Service of Canada, however, the additional use of the SKEB and SPPT could not further improve the results. [Whitaker and Hamill \(2012\)](#) compared the combination of the RTPS and additive noise (based on random samples of model truncation error) with the combination of the RTPS and SKEB in their idealized global case study, and concluded that the latter combination could hardly improve the former one. [Hamrud et al. \(2015\)](#) showed with the ECMWF Ensemble Prediction System that extending the baseline system consisting of the RTPS only with additive noise (based on a climatology of forecast differences (48 minus 24 h) valid at the same time) or the SPPT can both further improve ensemble mean forecast skill scores, but the combination of the RTPS and the additive noise is adopted as the standard configuration because it leads to more improvements. The additional use of the SPPT in the standard configuration cannot lead to further improvements. [Ha et al. \(2015\)](#) combined adaptive multiplicative inflation ([Anderson 2009](#)) with the SKEB and with the multiphysics approach, respectively, using the WRF-DART ([Anderson et al. 2009](#)) system, and showed the first combination consistently outperforms the second one in innovation statistics and mesoscale forecasts. [Bowler et al. \(2017\)](#) use a quasi-operational 4D-EnVar system at Met Office and combined the RTPP or the RTPS with the control run that included random parameters, the SKEB and the additive noise (based on random draws of samples for analysis increment [Piccolo and Cullen 2016](#)), and they found that the RTPP was effective in maintaining spread but it also produced too large-scale and too balanced perturbations, and the RTPS required an unusual relaxation factor greater than one to generate sufficient spread. The combination of the RTPP and RTPS turned out to be the best.

Another technique that does not fall into any categories mentioned above but proves to be useful for storm-scale data assimilation is to create convective plumes by triggering warm bubbles. The technique is well-known from idealized convection studies of [Weisman and Klemp \(1982\)](#) to forcibly trigger convection if the environmental atmospheric conditions are favorable but the “natural” trigger is missing.

This technique was first used in data assimilation by [Dowell et al. \(2004\)](#), in which warm bubbles were added at random locations close to the storm in each initial ensemble member. It was compared with the additive noise based on random smoothed perturbations in [Dowell and Wicker \(2009\)](#), in which the latter one reduced the amount of spurious convection during the data assimilation. But it should be noted that in that study the storm location was a priori known and warm bubbles were triggered only at the initial time and not used during the data assimilation to mitigate model error.

In this work, a more advanced warm bubble approach automatically detecting and triggering absent convective cells in a simulation is introduced. The convection-permitting numerical weather prediction (NWP) models are able to mimic realistic convective dynamics, however, they often fail to capture processes that trigger convection if they occur below the simulation resolution. During the data assimilation cycles, if a convective cell is missing in all ensemble members, the Gaussian-based data assimilation scheme cannot recover it in the analysis either. As possible remedy, the warm bubble technique will be applied during assimilation cycles, aiming to initiate convective cells that are observed by radars but missing in model runs: If a member misses one convective cell, a warm bubble is added to the planetary boundary layer at that location. If the member has enough instability, the rising warm bubble will trigger a cell during the next 10–15 min in a nonlinear and (thermo-) dynamically consistent way up to the level neutral buoyancy. Different members will have different instabilities and will react differently to the bubble.

It can be noticed that there are quite a number of different methods available for the representation of model error for various sources and scales, and the performance of different combinations has been explored. However, few studies focus on the comparison of methods accounting for subgrid-scale model error in the context of convective-scale data assimilation. In this article, we will compare the PSP, the SAN, the new warm bubble technique and their combinations. The PSP scheme adds the small-scale instability to the boundary layer to account for unresolved physical processes. The amplitude of the perturbations corresponds to different types of synoptic forcing. [Kober and Craig \(2016\)](#) showed that the PSP increased overall the turbulent variability and had neutral effects on the root-mean-square error (RMSE) of surface wind and temperature forecasts. Therefore, one may expect that the application of the PSP in the ensemble data assimilation should have more impact on increasing the ensemble spread than decreasing the RMSE of atmospheric states. The new

warm bubble technique attempts to adjust model states based on the real-time observations; therefore, its application in ensemble data assimilation could lead to reduction of the RMSE of atmospheric state but it adds only positive perturbations and could not effectively increase the overall ensemble spread since its perturbations are very local, and it targets only very rare high-impacts events. Besides, both the PSP and the warm bubble technique are designed to increase kinetic energy of analyses at smaller scales and to favor initiation of convection. But these two methods trigger convection in different manners. Therefore, we expect that their impact on the kinetic energy spectra and precipitation forecasts are different. Moreover, as shown in [Zeng et al. \(2019\)](#), the SAN added dual signed perturbations and resulted in more ensemble spread. However, the SAN is a flow-independent statistical scheme that might not be able to reduce the RMSE of atmospheric states. Based on the potential effects on the spread and RMSE, one may anticipate that the combination of the SAN and the warm bubble technique complement each other more than the combination of the SAN and the PSP. To investigate those hypotheses, several studies are carried out during a convective period in Germany with weak synoptic forcing, using the operational Kilometre-scale Ensemble Data Assimilation (KENDA) system of the Deutscher Wetterdienst (DWD) ([Schraff et al. 2016](#)). The performance within data assimilation cycles and subsequently the skill of forecasts are investigated.

The paper is organized as follows. [Section 2](#) briefly introduces the methods and describes the warm bubble technique. [Section 3](#) gives the experimental setup and [section 4](#) presents the results of different studies by investigating the spread, RMSE, kinetic energy spectrum and precipitation forecasts skill scores. [Section 5](#) summarizes the obtained results.

2. Approaches for representation of model error

In this section, we introduce the LAN and three methods that aim at representing subgrid-scale model error, including the SAN, the PSP scheme, and the new warm bubble technique. For the last method, a detailed description is given.

a. Additive noise

To represent model error on multiple scales ([Zeng et al. 2019](#)), two types of additive noise methods are used in this study.

1) LARGE-SCALE ADDITIVE NOISE (LAN)

As in [Zeng et al. \(2018\)](#), the LAN, derived from the climatological atmospheric background error covariance

B matrix of the global EnVar data assimilation system for the ICON (ICOsahedral Nonhydrostatic, Zängl et al. 2015) model, is applied. The **B** matrix is constructed vertically by using the NMC method and horizontally by compactly supported functions. The LAN should represent large-scale uncertainties arising from the global driving model ICON.

The LAN is executed at the analysis step [i.e., a random sample $\boldsymbol{\eta}_l^{(i)}$]. Based on the **B** matrix is added to the analysis ensemble member $\mathbf{x}^{a(i)}$ as follows:

$$\mathbf{x}^{a(i)} \leftarrow \mathbf{x}^{a(i)} + \alpha_l \boldsymbol{\eta}_l^{(i)}. \quad (1)$$

where α_l is a tunable parameter and set to 0.1 to mimic model error of 1-h forecasts. The LAN is based on 1-yr NMC-statistics that are originally designed to represent the background error of 3-h forecasts. More details about the implementation of the LAN can be found in Rhodin et al. (2013) and Zeng et al. (2018).

2) SMALL-SCALE ADDITIVE NOISE (SAN)

As in Zeng et al. (2019), the SAN, based on random draws from a set of samples for model truncation error (Hamill and Whitaker 2005), is used for convective-scale data assimilation. To create the set of samples, the COSMO-DE model is run with a high resolution of 1.4 km for a 2-week period in 2014 over Germany, governed by a low-pressure cyclone and associated with severe thunderstorms. The hourly outputs of model forecast runs at resolution of 1.4 km are then interpolated (via the program INT2LM, Schättler 2014) onto the coarse 2.8 km grid and serve as initial conditions for 1-h COSMO-DE forecast runs at resolution of 2.8 km. Therefore, each difference between high- and low-resolution forecast runs at a same forecast valid time provides a sample for model truncation error. As the LAN, the SAN adds a random sample $\boldsymbol{\eta}_s^{(i)}$ from the set to the analysis ensemble member $\mathbf{x}^{a(i)}$:

$$\mathbf{x}^{a(i)} \leftarrow \mathbf{x}^{a(i)} + \alpha_s \boldsymbol{\eta}_s^{(i)}. \quad (2)$$

Coefficient $\alpha_s = 1.25$ has been tuned in Zeng et al. (2019) under weak and strong forcing synoptic situations. Its value is similar to 1.20 used by Hamill and Whitaker (2005) for their application. The coefficient α_s is slightly larger than one since these samples of the model truncation error still cannot capture all small-scale processes. To ensure that additive noise effects the analysis ensemble covariance without changing the ensemble mean, the mean $\bar{\boldsymbol{\eta}}_s = 1/N_e \sum_{i=1}^{N_e} \boldsymbol{\eta}_s^{(i)}$ is subtracted from each randomly chosen sample (Whitaker et al. 2008), where N_e is the ensemble size (the same is also done for

the LAN). More details about the description of the SAN can be found in Zeng et al. (2019).

b. Stochastic boundary layer perturbations

The physically based stochastic perturbation scheme (PSP) for turbulence was first introduced by Kober and Craig (2016) and it has been conceptually revised by Rasp et al. (2018). The PSP scheme injects variability into the boundary layer based on the unresolved turbulence with the aim to improve the coupling between subgrid turbulence and resolved convection in kilometer-scale models. It is formulated as an additive perturbation scheme to the parameterized tendencies of each variable $\Phi \in \{\text{temperature } T, \text{ relative humidity } q_v, \text{ vertical velocity } w\}$:

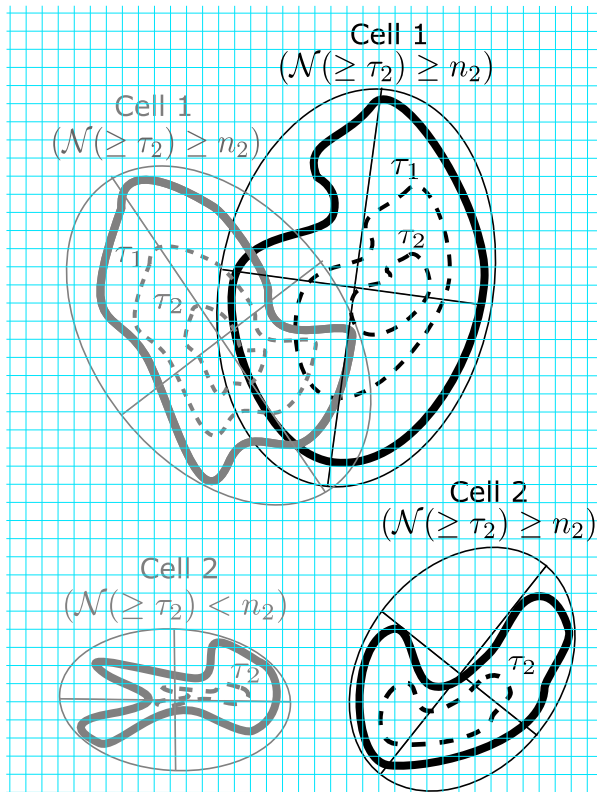
$$\left(\frac{\partial \Phi}{\partial t}\right)_{\text{total}} = \left(\frac{\partial \Phi}{\partial t}\right)_{\text{param}} + \alpha_p \eta \frac{1}{\tau_{\text{eddy}}} \frac{l_{\text{eddy}}}{\Delta x_{\text{eff}}} \sqrt{\Phi^2}, \quad (3)$$

where $\tau_{\text{eddy}} = 10$ min, $l_{\text{eddy}} = 1$ km are typical time and length scales for convective boundary layer eddies, and $\Delta x_{\text{eff}} = 5\Delta x$ is the effective model resolution (Bierdel et al. 2012); $\sqrt{\Phi^2}$ is the subgrid standard deviation computed in COSMO's turbulence scheme, which is a 1.5-order local closure (Raschendorfer 2001; Mellor and Yamada 1982). η is a two-dimensional random field with horizontal correlation length $5\Delta x$. A new random field is generated every 10 min, during which it is held constant. The T , q_v , and w perturbations are correlated by using the same random field η (i.e., they have the same sign). This is physically reasonable since the upward branch in a hypothetical subgrid eddy is warm and moist. The amplitudes of the perturbations are determined from the subgrid variances computed in the boundary layer scheme. These variances are also correlated (Mellor and Yamada 1982). Finally, α_p is set to 7.2.

The effect of the PSP scheme is twofold: first, it triggers more and earlier convection in situations with weak synoptic forcing (Keil et al. 2019). Second, it causes a quick error growth on convective scales if used in an ensemble forecasting setting.

c. Warm bubble

In this work, we present a new warm bubble technique that is more advanced than the one in Dowell et al. (2004) and Dowell and Wicker (2009) because it can automatically determine where to initiate warm bubbles in the model, based on the comparison of radar observations with the model counterpart simulated by the radar forward operator EMVORADO (Zeng et al. 2014, 2016). This technique consists of detection and triggering algorithms and is meant to compensate for well-known underrepresented trigger mechanisms in the



Cell 1: Convective cell
 Cell 2: Non-convective cell
 Cell 1: Convective cell
 Cell 2: Missing cell

FIG. 1. Sketch of the detection algorithm: Observed cells are denoted with black thick lines. Simulated cells are denoted with gray thick lines. Dashed black (gray) lines are contour lines of reflectivity for observations (simulations). Ellipses with axes in black (gray) thin lines are determined by principal component analysis. In each individual cell, $\mathcal{N}(\geq \tau_2)$ means the number of grid points with reflectivities exceeding the threshold value τ_2 . The observed cell 1 is identified as a convective cell and it overlaps with the simulated cell 1, therefore, it is not a missing cell. The observed cell 2 is also identified as a convective cell but there is no overlap with any simulated cells, so it is a missing cell. The simulated cell 2 does not satisfy $\mathcal{N}(\geq \tau_2) \geq n_2$, it is thus not considered as a convective cell. Blue thin lines in the background indicate the model grid.

model itself (model errors), caused by, for example, too coarse grid resolution to resolve small-scale orographic lift or subgrid-scale perturbations by shallow convection.

The detection algorithm works on the radar reflectivity composite, obtained by interpolation of radar scans at elevation of 0.5° onto the COSMO-DE grid. In case of overlapping by several radars, the maximal reflectivity value is taken. The composite covers the entirety of Germany and part of the neighboring countries. The detection algorithm searches for continuous regions of reflectivities above the threshold value τ_1 by checking the neighboring grid points. A continuous region is defined as a convective feature if two conditions are

TABLE 1. Parameters and their values used in the detection algorithm for model and observation.

Detection parameter	Model	Observation
τ_1 (dBZ)	25	25
τ_2 (dBZ)	30	30
n_1 (grid point)	15	15
n_2 (grid point)	4	4
m_{en}	1.0	1.0
m_{add} (km)	10	10

fulfilled: first, it contains at least n_1 grid points in the region; second, at least n_2 grid points exceed the threshold value τ_2 (note that $\tau_2 > \tau_1$). The second condition indicates existence of a cell core, for instance, $\tau_2 = 30$ dBZ is a typical value for mid-European supercells. Once a convective feature is detected, we use the principal component analysis to find the best ellipse that matches the region. The resulting ellipses can then be optionally enlarged by multiplying the axis lengths with a factor m_{en} or by adding extra length m_{add} to the axes. This option has been introduced to avoid bubbles being triggered too close to existing convection in the model. A sketch of the detection algorithm is given in Fig. 1. All parameters of the detection algorithm used in this study are summarized in Table 1. In the present work, the same parameter values are used for both observations and model to detect convective cells. Note that it is usually enough to find missing convective cells in composite of a single radar elevation for efficiency reasons. Figure 2 illustrates an actual case, in which six missing convective cells are found.

Once the missing cells are detected, the goal of the triggering algorithm is to initiate convection in those regions in the model. It first searches for ellipses that are identified as convective cells by the detection algorithm in observations but do not overlap with any ellipses in the model. At those locations, the temperature will be instantaneously increased (while keeping relative humidity constant). The perturbed volume has a fixed ellipsoid shape as illustrated in Fig. 3. In the horizontal, the volume center is set as each ellipse center with horizontal axis lengths D_x and D_y . In the vertical, the volume center is H_z above ground level. D_z is the vertical axis length. The maximum temperature disturbance D_T is at the center and it decreases toward the ellipsoid border following a cosine function within $[0, \pi/2]$. The bubbles are allowed to freely develop and produce realistic convection if the preconvective environment is favorable. A simulation that a triggered warm bubble evolves to convection is demonstrated in Fig. 4. All triggering parameters used in this study are given in Table 2. It should be mentioned that those

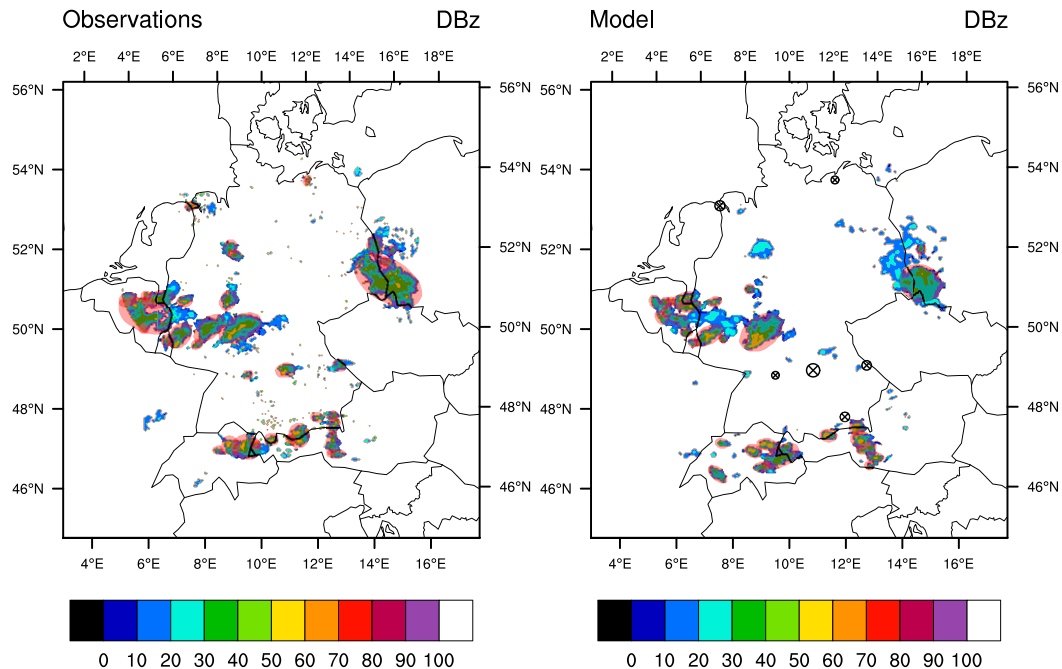


FIG. 2. An example of detecting missing convective cells in COSMO-DE domain. The red ellipses indicate well-defined convective features. (left) Observation and (right) model simulation of radar reflectivity; \otimes indicates the identified missing cells in model simulation. In this case, six missing convective cells are found.

parameters have been tuned, so that relatively balanced states can be achieved. We have also tried more aggressive configurations (e.g., larger $D_{x,y}$ and D_T), but they tend to generate too many spurious gravity waves above the tropopause, probably because the warm adjustment is not mass balanced. Although more convective clouds can be created, they also decay very fast. Consequently, good results are not possible within assimilation cycles, nor in subsequent forecasts. We have also tested including moisture perturbations in the bubble, but the results were very similar (not shown).

The detection/triggering algorithm is run independently for each ensemble member and every 15 min. This time allows for the early development of convective cells, so that warm bubbles will not be repeatedly triggered in the same position.

3. Experimental setup

We perform data assimilation experiments using the LETKF (Hunt et al. 2007) of the KENDA system at the DWD. Observations are assimilated hourly, which include conventional observations (e.g., TEMP, PROF, AIREP and SYNOP) as well as radar reflectivity and no-precipitation observations (i.e., ≤ 5 dBZ). Furthermore, radar reflectivity observations are temporally thinned (i.e., only the latest 5 min of radar observations before

the analysis time are assimilated). The size of the ensemble N_e is 40 and twenty members are used for 6-h ensemble forecasts that are initialized at 1000, 1100, ..., 1700, and 1800 UTC. A detailed description of the KENDA system and the implementation of the LETKF is given in Schraff et al. (2016), Bick et al. (2016), and Lange and Janjić (2016). More details about treatment (e.g., specification of observation error, localization and superobbing) of observations (especially radar reflectivity) are available in Zeng et al. (2018).

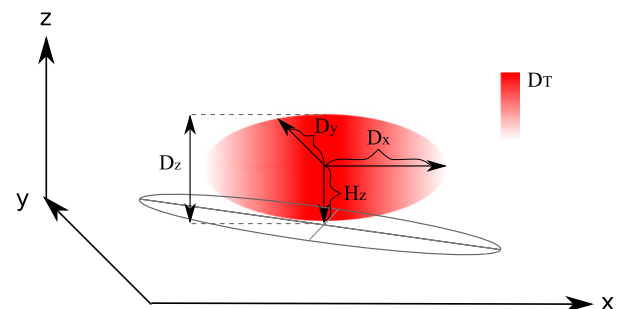


FIG. 3. Sketch of the triggering algorithm: the perturbed volume has a fixed ellipsoid shape. In the horizontal, the volume center is set as the ellipse center with horizontal axis lengths D_x and D_y . In the vertical, the volume center is H_z above ground level; D_z is the vertical axis length. The maximum temperature disturbance D_T is at the center and it decreases toward the ellipsoid border following a cosine function within $[0, \pi/2]$.

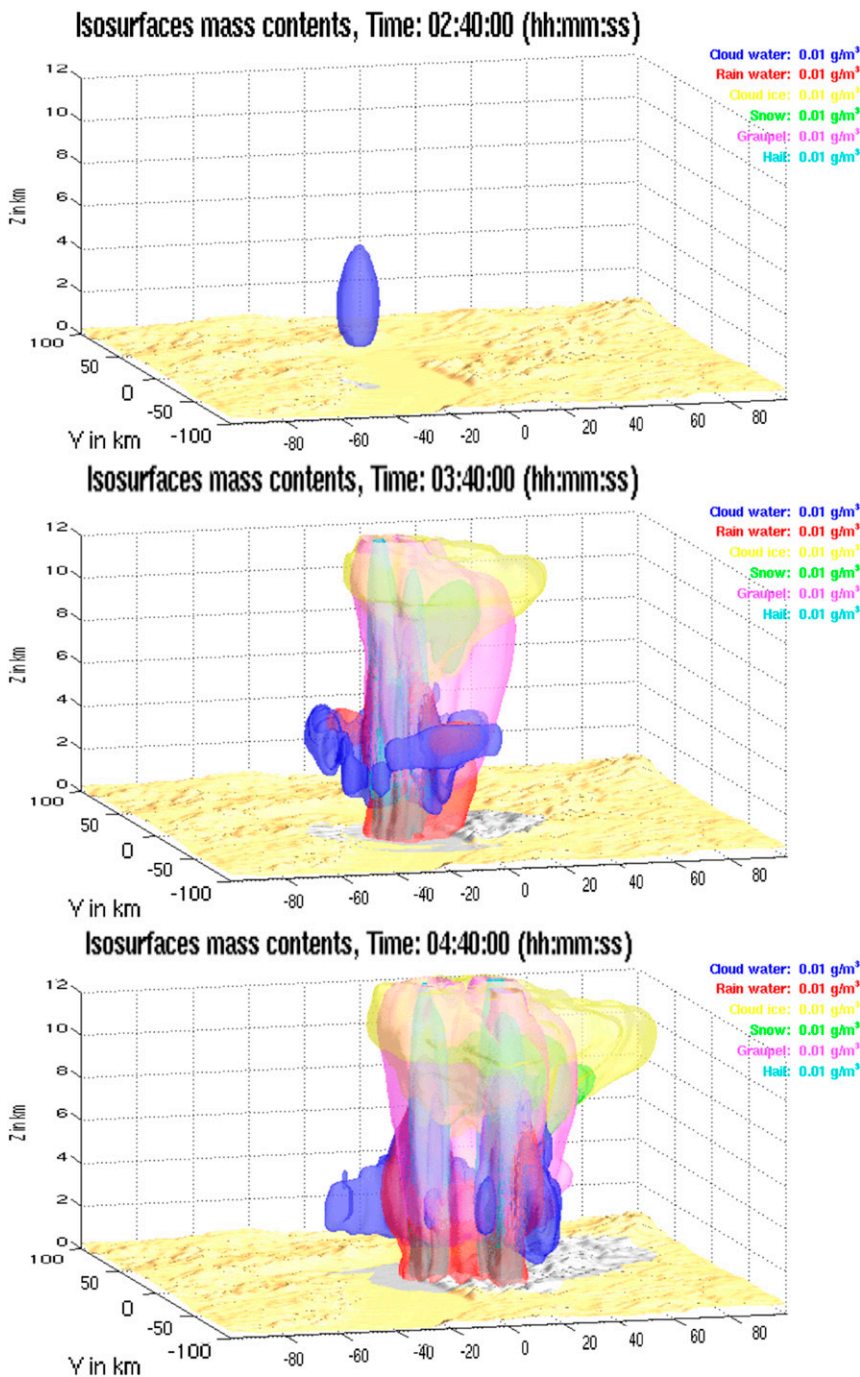


FIG. 4. Illustrative example of development of convection triggered by the warm bubble technique.

The COSMO model is convection-permitting, fully compressible and nonhydrostatic (Baldauf et al. 2011; Doms et al. 2011; Doms and Baldauf 2018). The number of grid points are $421 \times 461 \times 50$ and the horizontal grid spacing is 2.8 km. The one-moment microphysical scheme Lin et al. (1983) and Reinhardt and Seifert

(2006) is used. The ensemble prediction system (EPS) of the operational global ICON model provides lateral boundary conditions.

In the present work, we choose a week from 3 June to 10 June 2016. This period was under weak synoptic forcing conditions with many scattered convective cells

TABLE 2. Parameters and their values used in the triggering algorithm.

Bubble parameter	Model
$D_{x,y}$ (km)	10
H_z (km)	2
D_z (km)	4
D_T (K)	1.5

over Germany (Zeng et al. 2018). Experimental setups are given in Table 3. Three studies are carried out, in which the LAN will be always applied but different techniques for representation of subgrid-scale model error will be used in different studies. In the first study, the reference experiment E_BASE is compared to E_P, in which the PSP scheme is applied in addition to the LAN; In the second study, E_BASE is compared to E_B, in which the warm bubble technique is applied in addition to the LAN. One may worry that radar reflectivity observations are used twice for detection of missing cells as well as for assimilation, which may conflict with the assumption underpinning data assimilation that observation errors and background errors are independent from each other. However, there are several studies (e.g., Dowell and Wicker 2009; Sobash and Wicker 2015; Hu et al. 2019) in practice, which use the reflectivity observations both for additive noise and for assimilation. Furthermore, the assimilation window is 1 h in this work, the detection algorithm is run every 15 min, and only the latest 5 min of radar observations before the analysis time are assimilated; therefore, only a small portion of observations used for the detection algorithm are also assimilated and thus the assumption is very mildly violated. In the third study, E_BASE is compared to E_SAN that combines the LAN with the SAN. Detailed comparisons between E_BASE and E_SAN can be found in Zeng et al. (2019). Furthermore, two other experiments E_SANP and E_SANB with additional use of the PSP scheme or warm bubble technique are performed. Instead of E_BASE, E_SAN is chosen as the reference experiment in the third study.

4. Experimental results

We hypothesized in the introduction that the use of the PSP scheme in ensemble data assimilation should be effective in increasing the ensemble spread due to its perturbations over the full domain, but it might not be able to reduce the RMSE (see Study 1). The application of the warm bubble technique might help to reduce the RMSE due to the adjustment to real-time observations, but it might not be good at increasing the spread since the perturbations are fairly local (see Study 2). Both

TABLE 3. Experimental setup: ✓ means “on” and × means “off.”

EXP	LAN	PSP	Warm bubble	SAN
Study 1				
E_BASE	✓	×	×	×
E_P	✓	✓	×	×
Study 2				
E_BASE	✓	×	×	×
E_B	✓	×	✓	×
Study 3				
E_BASE	✓	×	×	×
E_SAN	✓	×	×	✓
E_SANP	✓	✓	×	✓
E_SANB	✓	×	✓	✓

methods are supposed to increase kinetic energy of analyses at smaller scales and trigger more convection. However, their impacts on the kinetic energy spectra and precipitation forecasts may be different and can be explored. In addition, the SAN results in larger spread but it might not produce smaller RMSE because its perturbations are random and lack flow dependency. Based on effects on the RMSE and spread, it seems that the combination of the SAN and the warm bubble technique might compensate each other more than the combination of the SAN and the PSP (see Study 3).

In the following, we will use several metrics to compare the performance of experiments. Besides the spread and RMSE for assimilation cycles, we use fractions skill score (FSS; Roberts and Lean 2008) and false alarm rate (FAR) for precipitation forecasts. More details about the calculation of FSS and FAR can be found in Zeng et al. (2018). For verification by the RMSE, FSS, and FAR, the differences relative to the reference run are computed. The bootstrap approach is applied to take uncertainties in verification scores into account. 10 000 bootstrap resampling is performed to check the statistical significance at 95% confidence intervals, using the bootci() function of Matlab R2019b with bias-corrected and accelerated bootstrap (BCa bootstrap, Efron 1987). In addition, a spectral analysis of horizontal and vertical kinetic energy based on the comparison to E_BASE is carried out for short-term forecasts. At scales smaller than 300 km, the horizontal kinetic energy is strongly related to the amount of convective precipitation as shown by Selz et al. (2019). One goal of the spectral analysis here is to examine if the improvements in precipitation forecasts correspond to the added kinetic energy. Furthermore, we divide the horizontal kinetic energy into divergent and rotational parts to show the impacts of different methods in more details.

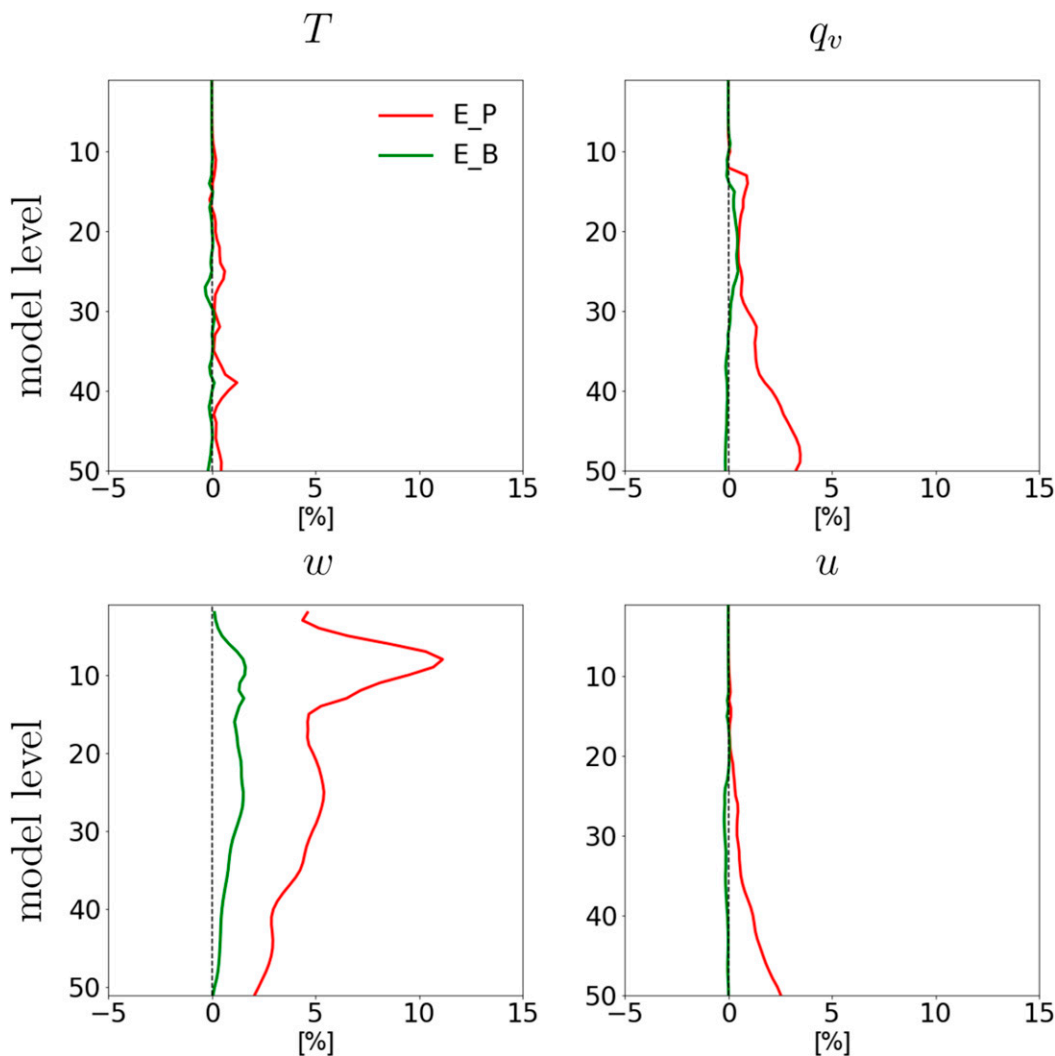


FIG. 5. Vertical profiles of relative differences (%) in background ensemble spread of E_P and E_B, relative to E_BASE, averaged over all data assimilation cycles, for model state variables (top left) T , (top right) q_v , (bottom left) w , and (bottom right) u . The vertical dashed line indicates no difference relative to E_BASE.

a. Study 1

Figure 5 illustrates the vertical profiles of relative differences (%) in the background ensemble spread of E_P for model states T , q_v , w , and u , compared to E_BASE. For T , E_BASE and E_P are almost indistinguishable. For q_v and u , E_P has slightly larger spread than E_BASE below model level 10 (~10 km); for w , E_P is considerably larger than E_BASE for the whole profile. This is the result of more convection being triggered by the PSP scheme, increasing variability of w in the troposphere.

Figure 6 also shows the vertical profiles of relative differences (%) in the background ensemble spread of E_P but for the radial wind. Additionally, the vertical profiles of relative differences in the RMSE of background ensemble,

verified against the radial wind, are also given (note that radial wind is not assimilated and thus it is an independent validation dataset). It is worth noting that the spread of E_BASE is adequate based on the spread skill ratio (Aksoy et al. 2009) as shown in Zeng et al. (2018). The spread of E_P is larger than that of E_BASE and the discrepancies decrease with the increasing height. The RMSE of E_P is slightly larger than that of E_BASE in the atmosphere lower than 2 km and slightly smaller around 4 km.

Table 4 gives the relative changes (%) of kinetic energy (in spectrum space) of analyses (initial states) in E_P compared to E_BASE for different scales and heights. Similar as Selz et al. (2019), we focus on the wavelength range between 14 and 1000 km, and for ease

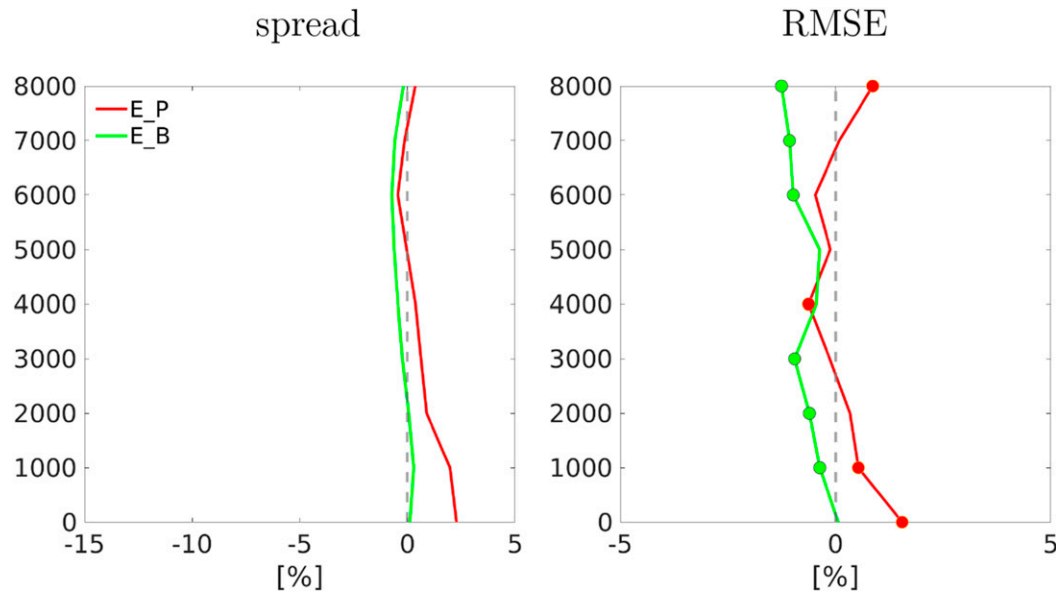


FIG. 6. Vertical profiles of relative differences (%) in (left) background ensemble spread and (right) RMSE of the radial wind in E_P and E_B, compared to E_BASE, averaged over all data assimilation cycles. For the RMSE, filled dots mean that differences are statistically significant at 95% confidence intervals, no dots mean no statistical significance.

of discussion, the wavelengths are divided into 14–100, 100–300, and 300–1000 km, which may roughly represent convective scale, mesoscale, and synoptic scale, respectively. The spectra of horizontal kinetic energy (KE) are calculated by the sums of energy spectra of divergent (DIV) and rotational (ROT) wind. In addition, the relative changes of vertical kinetic energy (KEW) are also given. It should be noted that the spectra of KE are steep and most of KE is gathered at larger scales but the spectra of KEW are fairly flat (not shown). At 10 km, the gains of DIV and ROT are largest at 14–100 km and decrease rapidly with increasing wavelengths, and the reduction of ROT at 300–1000 km even leads to smaller total KE. At 5 km, marginally smaller DIV and ROT are present at 100–300 km while marginally larger DIV and ROT at other wavelengths. Regarding KEW, the largest gains occur at 14–100 km at both heights and decline fast with larger scales. Table 5 gives the relative gains of kinetic energy of 1-h forecasts. In general, the gains in KE have decayed a bit while the gains in KEW have remarkably decreased for all scales and heights. The rapid decay of vertical velocity perturbations is a known issue with the PSP scheme that has been corrected in the most recent version Hirt et al. (2019).

Figure 7 compares E_BASE and E_P with the radar reflectivity observation composite (at elevation of 0.5° at the initial time (1000 UTC 6 June 2016) and in the 2-h forecast. The ensemble probability (the number of

ensemble members exceeding a threshold (here 20 dBZ) divided by the size of ensemble) is shown. At both times, the intensity and locations of observed reflectivities are well captured by both E_BASE and E_P, whereas the latter one leads to a general increase in the probability (e.g., at the marked locations), as would be expected from the analysis of Kober and Craig (2016), but this effect is small.

For the verification of 6-h ensemble forecasts against the observed radar-derived precipitation rate, Fig. 8 shows the ensemble mean areal coverage (%) of hourly accumulated precipitation rate exceeding the threshold values 1.0 and 5.0 mm h^{-1} over the radar precipitation scanning domain (see Fig. 2 of Hirt et al. 2019), aggregated hourly over all forecasts initiated at 1800 UTC. For 1.0 mm h^{-1} , it is noticed that both mean areal coverages of E_BASE and E_P underpredict the precipitation but the latter one results in slightly more events. Some ensemble members overestimate the precipitation at the last 3-h forecast lead times. Similar can also be seen for 5.0 mm h^{-1} except that overestimation by some members occur at third and fourth lead times. Figure 9 illustrates the FSS (for different scales, i.e., 14, 70, and 140 km) and FAR values for threshold value 1.0 and 5.0 mm h^{-1} , as a function of forecast lead time. With respect to the FSS, E_P is slightly better than E_BASE for 1.0 mm h^{-1} but the differences are not statistically significant for all scales; For 5.0 mm h^{-1} , E_P is considerably better than E_BASE with statistical significance

TABLE 4. Comparison of analyses of E_P, E_B, E_SAN, E_SANP, and E_SANB relative to those of E_BASE by relative change of kinetic energy spectra over all wavelengths, 14–100, 100–300, and 300–1000 km at two different heights (10 and 5 km): The spectrum of horizontal kinetic energy (KE) is calculated by the sum of energy spectra of divergent (DIV) and rotational (ROT) winds. As in Selz et al. (2019), DIV and ROT are computed by the two-dimensional discrete cosine transform (DCT; Denis et al. 2002) of the divergence and relative vorticity, respectively. The one-dimensional spectra then result from radial sums of the two-dimensional DCT spectra (Erico 1985). The spectra are computed over a region given by discarding 20/30 grid points in longitude-latitude on each side of the COSMO-DE domain to mitigate boundary effects. The spectra are averaged over all ensemble members and initial dates for 6-h forecasts. The relative change of DIV is calculated by $[DIV(E_2) - DIV(E_1)]/KE(E_1) \times 100\%$, where E_2 is one experiment from the first column and E_1 is the reference run E_BASE. The relative change of ROT is calculated by $[ROT(E_2) - ROT(E_1)]/KE(E_1) \times 100\%$; The relative change of KE is thus equal to the sums of relative changes of DIV and ROT. The relative change of vertical kinetic energy spectra (KEW) is calculated by $[KEW(E_2) - KEW(E_1)]/KEW(E_1) \times 100\%$. Therefore, the numbers given in the table are in unit of percentage.

Exp	10 km	5 km	10 km	5 km	10 km	5 km	Total						
							14–100 km		100–300 km		300–1000 km		
	DIV	ROT	DIV	ROT	DIV	ROT	KEW	KE	KEW	KE	KEW	KE	KEW
E_P	0.19	-0.39	2.43	2.13	4.56	7.15	0.37	0.41	0.78	6.37	0.19	-0.42	-0.23
	0.08	0.19	1.32	1.11	2.43	6.53	-0.58	-0.42	-1.00	3.56	0.08	0.19	0.27
E_B	0.36	-0.13	1.46	0.99	2.45	3.40	0.82	1.22	2.04	3.54	0.36	-0.16	0.20
	0.23	0.12	-0.26	-0.39	-0.65	3.26	0.82	-0.42	0.40	3.04	0.37	0.13	0.23
E_SAN	1.80	0.28	9.16	8.44	17.60	21.66	5.59	4.69	10.28	11.13	1.67	0.15	1.82
	2.80	0.69	12.96	11.70	24.66	28.74	6.57	4.67	11.25	4.02	2.47	0.34	2.81
E_SANP	3.91	-0.27	10.20	9.80	20.00	24.03	6.82	5.43	12.25	11.84	3.82	-0.45	3.37
	4.22	0.40	13.80	12.05	25.85	32.65	6.80	4.97	11.77	4.66	3.92	0.03	3.95
E_SANB	2.16	0.45	10.25	9.02	19.27	25.72	6.87	5.06	11.93	14.97	2.01	0.31	2.32
	3.39	0.30	13.25	12.19	25.44	32.34	6.95	4.89	11.84	4.53	3.08	-0.08	3.00

up to 4 h before they approach for all scales. It is interesting to note that although the areal coverage of precipitation exceeding a threshold is increased in E_P, the FAR is decreased. This suggests that the increased ensemble spread allows the data assimilation to better locate convective cells, even though the negative bias is only partially corrected.

To conclude, the additional application of the PSP scheme within the data assimilation cycles results in more background ensemble spread of specific model states (i.e., q_v , u , and w) but its influences on the RMSE seems to be neutral. Furthermore, it produces analysis model states with more horizontal and vertical kinetic energy at convective and mesoscales. The added horizontal energy weakens slowly in 1-h forecasts whereas the added vertical energy dissipates much faster. However, with respect to precipitation forecasts, its positive impacts linger about 3 h.

b. Study 2

For the comparison of background ensemble spread for model states, it can be seen in Fig. 5 that E_B produces slightly larger spread than E_BASE for w for the whole profile, probably due to more convection triggered by warm bubbles. However, no clear differences are visible for T , q_v , and u . Similarly, E_B does not result in more background ensemble spread than E_BASE for the radial wind as shown in Fig. 6. It should be due to the fact that the perturbations of the warm bubble technique are fairly localized. However, the RMSE of E_B is overall smaller. It seems that the introduction of warm bubbles has not only the potential to recover missed convective cells but also to improve the atmospheric state within the assimilation cycles.

Figure 7 also compares E_BASE and E_B with the radar reflectivity observation composite. It can be seen that E_B is better than E_BASE at the marked locations at the initial time and 2-h forecast.

With respect to the areal coverage of hourly accumulated precipitation in 6-h ensemble forecasts, it is shown in Fig. 8 that for the threshold value 1.0 mm h^{-1} E_B results in more events than E_BASE at the first and last 2-h forecast lead times. For 5.0 mm h^{-1} , E_B is slightly larger than E_BASE for all lead times. Furthermore, the FSS is given in Fig. 9. For the threshold value 1.0 mm h^{-1} and the scale 14 km, it is noticeable that E_B is better than E_BASE with statistical significance at the beginning and the forecast skills converge by 3 h. Similar results are found for scales 70 and 140 km although the differences are not statistically significant. For 5.0 mm h^{-1} and all scales, E_B is slightly better than E_BASE in the first 3 h and then slightly worse, although most of the differences are not statistically

TABLE 5. As in Table 4, but for 1-h forecasts.

Exp	Total	14–100 km					100–300 km					300–1000 km					
		DIV	ROT	KE	KEW	KEW	DIV	ROT	KE	KEW	KEW	DIV	ROT	KE	KEW	KEW	
E_P	10 km	0.14	-0.40	-0.26	2.03	1.98	1.95	3.93	2.22	0.26	0.26	0.52	0.14	0.14	-0.43	-0.29	-0.10
	5 km	-0.05	0.19	0.14	1.64	1.26	0.91	2.17	1.89	-0.18	-0.68	-0.86	-0.53	-0.07	0.19	0.12	-1.08
E_B	10 km	0.28	-0.15	0.13	2.12	1.25	1.07	2.32	2.08	0.80	1.12	1.92	3.94	0.27	-0.17	0.10	-0.05
	5 km	0.21	-0.02	0.19	1.62	-0.19	-0.28	-0.47	1.72	1.07	-0.17	0.90	2.18	0.20	-0.02	0.18	-1.92
E_SAN	10 km	1.28	0.16	1.44	1.44	5.05	4.87	9.92	1.18	4.91	4.09	9.00	3.28	1.17	0.05	1.22	6.41
	5 km	1.48	1.24	2.72	-5.96	5.40	4.83	10.23	-6.18	6.56	3.78	10.34	-4.14	1.21	1.08	2.29	-1.68
E_SANP	10 km	3.40	-0.20	3.20	2.34	6.06	6.14	12.20	2.31	6.23	4.81	11.04	1.84	3.32	-0.34	2.98	4.93
	5 km	2.93	1.48	4.41	-4.67	6.45	5.47	11.92	-4.63	6.83	4.27	11.10	-5.66	2.69	1.31	4.00	-2.59
E_SANB	10 km	1.64	0.38	2.02	5.04	6.27	5.87	12.14	4.83	6.18	4.44	10.62	8.62	1.51	0.26	1.77	5.88
	5 km	2.07	1.00	3.07	-2.95	5.66	5.41	11.07	-3.09	6.86	4.15	11.01	-0.92	1.82	0.79	2.61	-1.64

significant. The FAR values indicate slightly better location of convective cells by E_B.

In addition, some differences between the PSP scheme and the warm bubble technique are worth pointing out here. For instance, E_B also results in analyses with more KEW than E_BASE over all wavelength ranges, but the gains of KEW are generally smaller than those of E_P and more evenly distributed over horizontal scales (see Table 4), this is also a sign of localized perturbations (the transformation of the perturbations into frequency space contains components over a wide range of scales if they are very localized in physical space). For the 1-h forecasts (see Table 5), the gains of KEW almost decay for all scales but not as much as in E_P, which indicates that the warm bubble technique is more effective in creating convective systems with larger circulation even if its perturbations are of smaller amplitudes than those of the PSP, and the perturbations grow upscale and exhibit mesoscale influences in very short-term forecasts. Furthermore, Fig. 10 compares the variation of surface pressure tendency S_t within the assimilation cycles in E_BASE, E_P and E_B. In all experiments, S_t arrives at its peaks at each analysis step and decreases rapidly in each forecast step. However, it can be seen that S_t behaves similarly in E_BASE and E_B whereas it remains at a slightly higher value in forecast steps in E_P for cycles between 0800 to 1800 UTC. This may be related to the fact that the PSP scheme is designed to be more effective in triggering convection at those specific times. A side-effect would be to produce more imbalanced model states as indicated by S_t .

To sum up, the use of the warm bubble technique during assimilation cycling does not effectively increase background ensemble spread of most model states except w , but it reduces the RMSE. Furthermore, it creates more horizontal kinetic energy especially at convective and mesoscales for analysis model states and almost evenly more vertical kinetic energy over different scales. With respect to precipitation forecasts, it is advantageous in the first 3 h. Compared to the PSP, the perturbations of the warm bubble technique may be of smaller amplitude and very localized but they are less imbalanced and seem to be more efficient in making larger-scale convective systems.

c. Study 3

Figure 11 compares the vertical profiles of spread and RMSE of the background ensemble in E_BASE, E_SAN, E_SANP, and E_SANB, verified against radial wind. In regard to background ensemble spread, it can be seen that E_SAN produces much more spread than E_BASE. E_SANB is close to E_SAN and E_SANP has slightly more spread than E_SAN under 2 km. However,

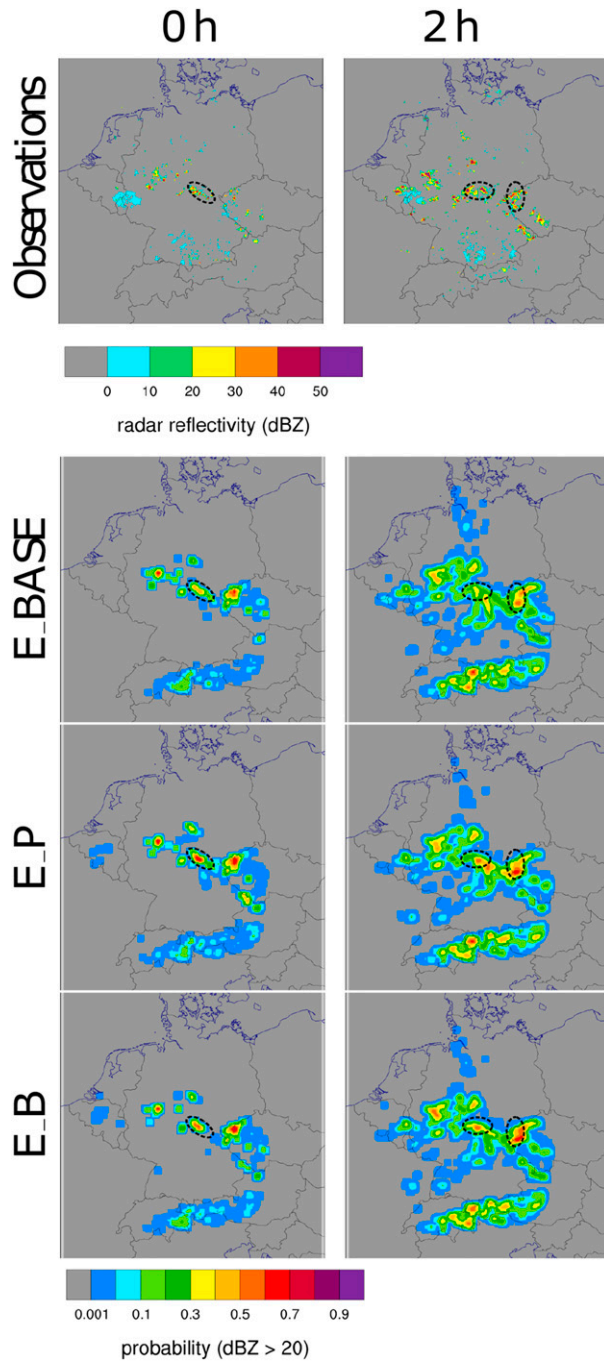


FIG. 7. Radar reflectivity composite at the elevation 0.5° at 1300 (initial time) and 1000 UTC 6 Jun 2016 (2-h forecast). (top row) Observations; (second row) ensemble probability of E_BASE for threshold 20 dBZ; (third and fourth rows) as in the second row, but for E_P and E_B.

in regard to background ensemble RMSE, E_BASE seems to be associated with the smallest RMSE and E_SAN with the largest. This seems to be a common downside of statistical schemes. For instance, as shown

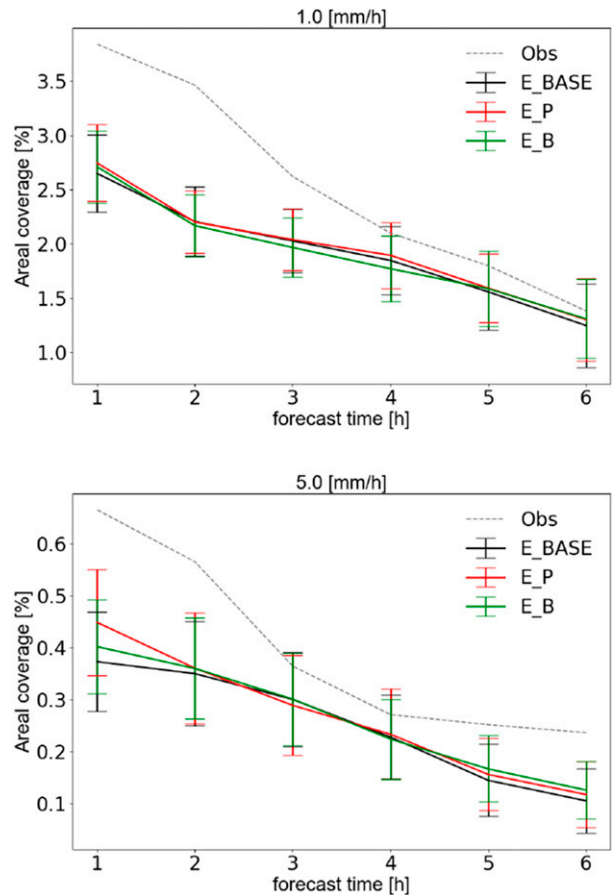


FIG. 8. Comparison of E_BASE, E_P, and E_B with precipitation observations by ensemble mean areal coverage (%) of hourly accumulated precipitation rate exceeding the threshold values (top) 1.0 and (bottom) 5.0 mm h^{-1} over the verification domain, aggregated hourly over all forecasts initiated at 1800 UTC. The error bar represents plus and minus one standard deviation of the ensemble.

in Piccolo et al. (2018), using samples of analysis increments to account for model error greatly increases the ensemble spread and reliability but also increases the RMSE for some fields. Nevertheless, with the help of the PSP scheme or the warm bubble technique, the RMSE can be reduced in E_SANP and E_SANB (especially in the latter one). This can be explained by the fact that the SAN is flow independent and thus does not represent model “errors of the day.” Combining it with physically based model uncertainty schemes such as the PSP scheme or with real-time observation-dependent warm bubble technique may compensate for some of the deficiencies.

As seen in Table 4, E_SAN results in analyses with considerably more total KE than E_BASE at both heights, the gains are largest at 14–100 km and decrease rapidly with increasing scales. The gains in total KEW are even more significant, dominated by contributions

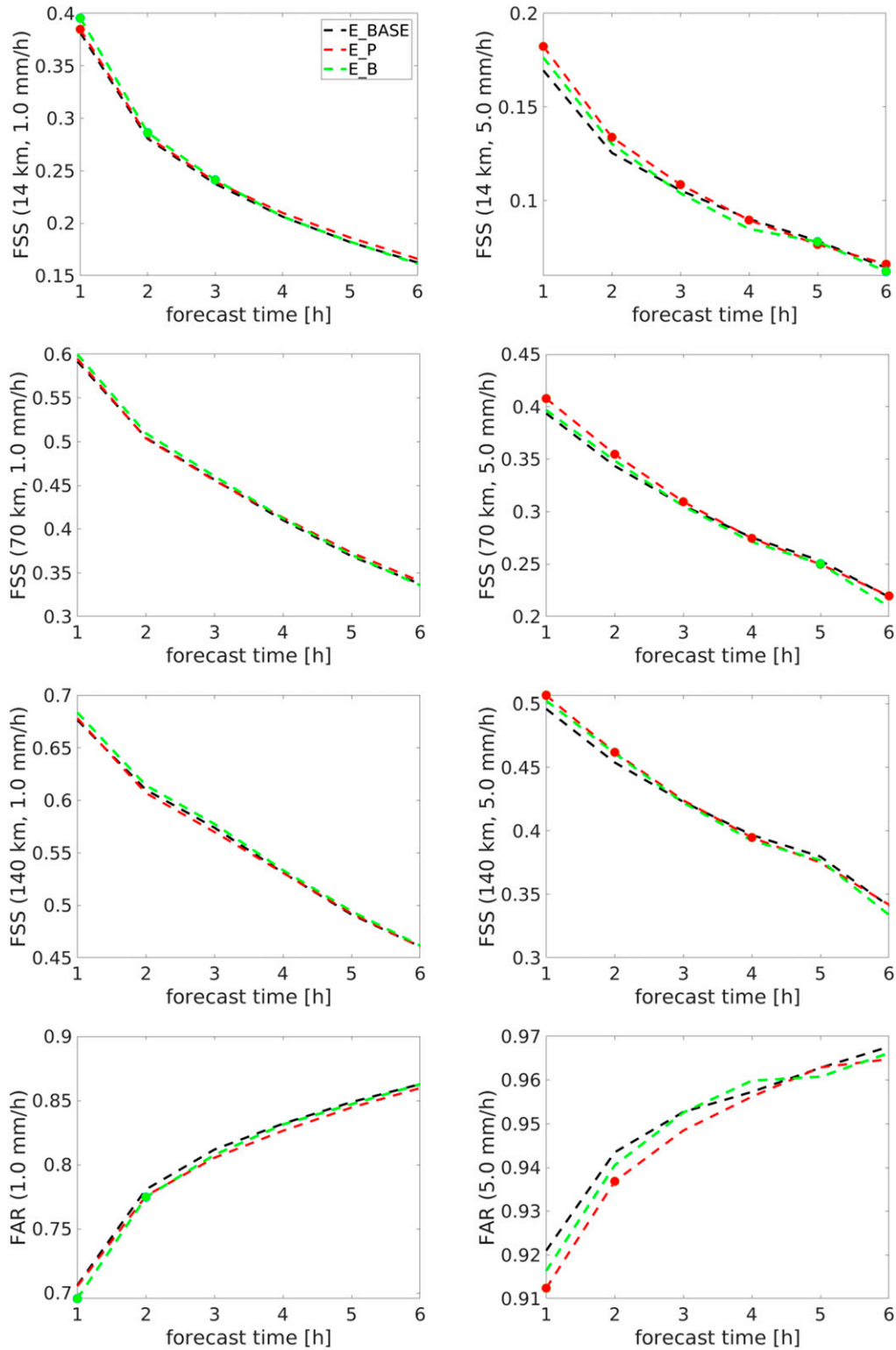


FIG. 9. Verification of 6-h ensemble forecasts against radar-derived precipitation rate for comparison of E_BASE and E_P in Study 1. (left) The first three panels illustrate the FSS values of experiments for the threshold value of 1.0 mm h^{-1} as a function of forecast lead time for different scales 14, 70, and 140 km, respectively. The last panel illustrates the FAR values for the threshold value of 1.0 mm h^{-1} . (right) As in (left),

Downloaded from <http://journals.ametsoc.org/mwr/article-pdf/148/6/2457/4947966/mwr148064.pdf> by guest on 29 August 2020

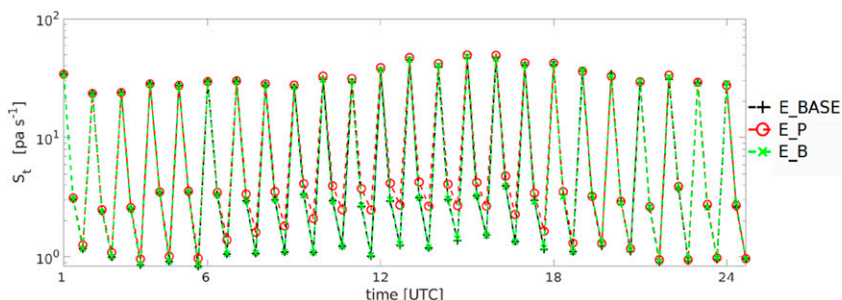


FIG. 10. Half-hourly evolution of surface pressure tendency S_t within the assimilation cycles in E_BASE, E_P, and E_B, averaged over all ensemble members for the period 0000 UTC 3 Jun–0000 UTC 4 Jun. Recall that the assimilation window is 1 h. Peaks occur at analysis steps.

from 14 to 100 km. For the 1-h forecasts, it can be seen in Table 5 that the gains in KE at 14–100 km have greatly decreased at both heights but still remain large while the gains at 100–300 and 300–1000 km are relatively steady. The gains in KEW have excessively reduced at both heights, especially at 14–100 km, and the reduction even leads to smaller KEW than E_BASE over all scales at 5 km height. In addition, the kinetic energy spectra of analyses and 1-h forecasts of E_SANP and E_SANB exhibit the added effects of the PSP scheme and the warm bubble, both of which amplify KE and KEW at all scales and heights.

Figure 12 compares E_BASE, E_SAN, E_SANP, and E_SANB with the radar reflectivity observation composite at the initial time (1400 UTC 6 June 2016) and in the 2-h forecast. At the initial time, E_SAN is superior to E_BASE at the highlighted location, so are E_SANP and E_SANB. For the 2-h forecast, it is not clear whether E_BASE or E_SAN is better at the marked location, but E_SANP and E_SANB are better than E_BASE.

Figure 13 depicts the verification of precipitation rates in 6-h ensemble forecasts based on the areal coverage. For the threshold value 1.0 mm h^{-1} , E_SANP produces most events, followed by E_SANB, E_SAN, and E_BASE. Different from Fig. 8 for Studies 1 and 2, not just some ensemble members but also the ensemble mean approached and even exceeds the observations at the late forecast lead times. For 5.0 mm h^{-1} , E_SAN results in more events than E_BASE at the first and last 2-h forecast lead times. Both E_SANP and E_SANB

increase the areal coverages which are larger than that of E_BASE except at the third forecast lead time. Figure 14 shows the verification of precipitation forecasts based on the FSS. Recall that E_SAN is the reference run instead of E_BASE. Based on the FSS values for the threshold value 1.0 mm h^{-1} and all scales, E_SAN is slightly better than E_BASE (some differences are statistically significant). E_SANP is slightly better than E_SAN in the first 2 or 3 h and then slightly worse. E_SANB is better than E_SAN up to 6 h but the differences become smaller with increasing lead time. For 5.0 mm h^{-1} , similar differences can be seen but the advantage of E_SAN, E_SANP, and E_SANB over E_BASE is much more significant at the beginning but also declines gradually with increasing lead time. The verification based on the FAR is also consistent with the FSS.

To conclude, the SAN produces perturbations of large amplitude and contributes greatly to the ensemble spread but it does not necessarily improve the RMSE within data assimilation cycles. It produces analyses associated with much more horizontal and vertical kinetic energy at convective and mesoscales. The large part of the added horizontal kinetic energy still remains after 1-h forecasts while the added part of vertical kinetic energy at the convective scale may overdissipate. The use of the SAN can significantly improve precipitation forecasts up to 6 h. Moreover, combining the SAN with either the PSP scheme or the warm bubble technique results in a smaller RMSE in cycles, larger

←

but for threshold value of 5.0 mm h^{-1} . Each FSS (FAR) value is computed as an average over all 63 forecast runs (the study period contains 7 days and each day has 9 forecast runs). The lines are marked as filled dots at the forecast lead times where the differences compared to E_BASE are statistically significant at 95% confidence intervals after 10 000 bootstrap resampling based on 63 difference samples. Same as Zeng et al. (2018), the values of MSE and MSE_{ref} required for calculation of FSS (false alarms and hits for calculation of FAR) are collected over all ensemble members when calculating the FSS (FAR).

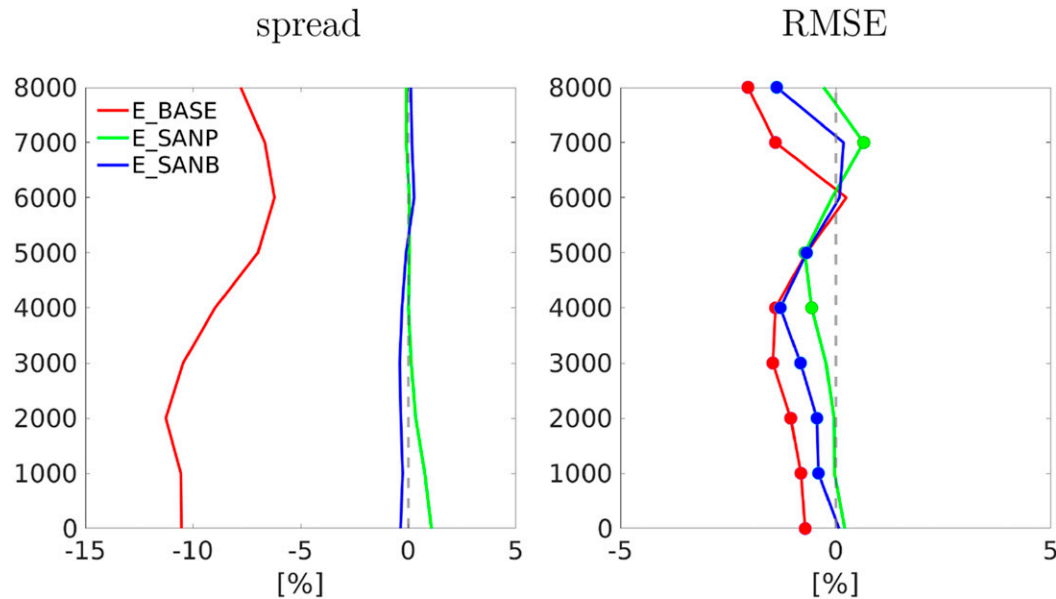


FIG. 11. As in Fig. 6, but for comparison of E_BASE, E_SAN, E_SANP, and E_SANB in Study 3.

horizontal and vertical kinetic energy over all scales for both analyses and 1-h forecasts, and further improvements in precipitation forecasts. The latter combination is overall more beneficial.

5. Summary

In this work, different approaches of representing model error at small/unresolved scale are presented, including the PSP scheme for turbulence (Kober and Craig 2016; Rasp et al. 2018), an advanced warm bubble approach, and small-scale additive noise (SAN; Zeng et al. 2019). Their performance is investigated and compared in the context of convective-scale data assimilation and subsequent forecasts for a convective period in Germany with weak synoptic forcing, using the operational KENDA system at the DWD.

It is found that the application of the PSP scheme results in more background ensemble spread in assimilation cycles but its effects on the RMSE of background ensemble are neutral. Moreover, it generates analysis states with increased horizontal and vertical kinetic energy at convective and mesoscales. The added horizontal energy declines much more slowly than the added vertical energy in 1-h forecasts, and better precipitation forecasts up to 3 h can be achieved. For the warm bubble technique, it is shown that it does not effectively create more ensemble spread but it is useful to reduce the RMSE within assimilation cycles. It results in analysis states with mild increase in horizontal kinetic energy at convective and mesoscales and with almost equal

increase in vertical kinetic energy over different scales. The added vertical energy is less than that added by the PSP scheme but it dissipates more slowly. Also here improved precipitation forecasts can be expected up to 3 h. In comparison, the use of the SAN contributes greatly to ensemble spread but it results in a larger RMSE in assimilation cycles. It significantly increases horizontal and vertical kinetic energy at convective and mesoscales in analyses. The large part of added horizontal energy still remains after 1-h forecasts but an overdissipation of vertical energy may occur at the convective scale. Skills of precipitation forecasts can be considerably improved up to 6 h. Combining the SAN with the PSP scheme or the warm bubble technique results in smaller RMSE due to their advantage in flow dependency. Both combinations also further amplify the horizontal and vertical kinetic energy and generate better precipitation forecasts than the SAN alone. However, since the SAN and the warm bubble complement each other more (the former one is useful to provide spread and the latter one to reduce the RMSE), this combination appears to be a better choice.

The improvements of precipitation forecasts are positively correlated to the added horizontal kinetic energy in analyses and forecasts. Comparatively, the added vertical kinetic energy dissipates rapidly in the forecasts and is therefore unlikely the dominant factor for improvement at longer forecast lead time. Furthermore, the vertical kinetic energy is also considered as an indicator for model imbalance as surface pressure tendency (Lange et al. 2017). The SAN is the most effective

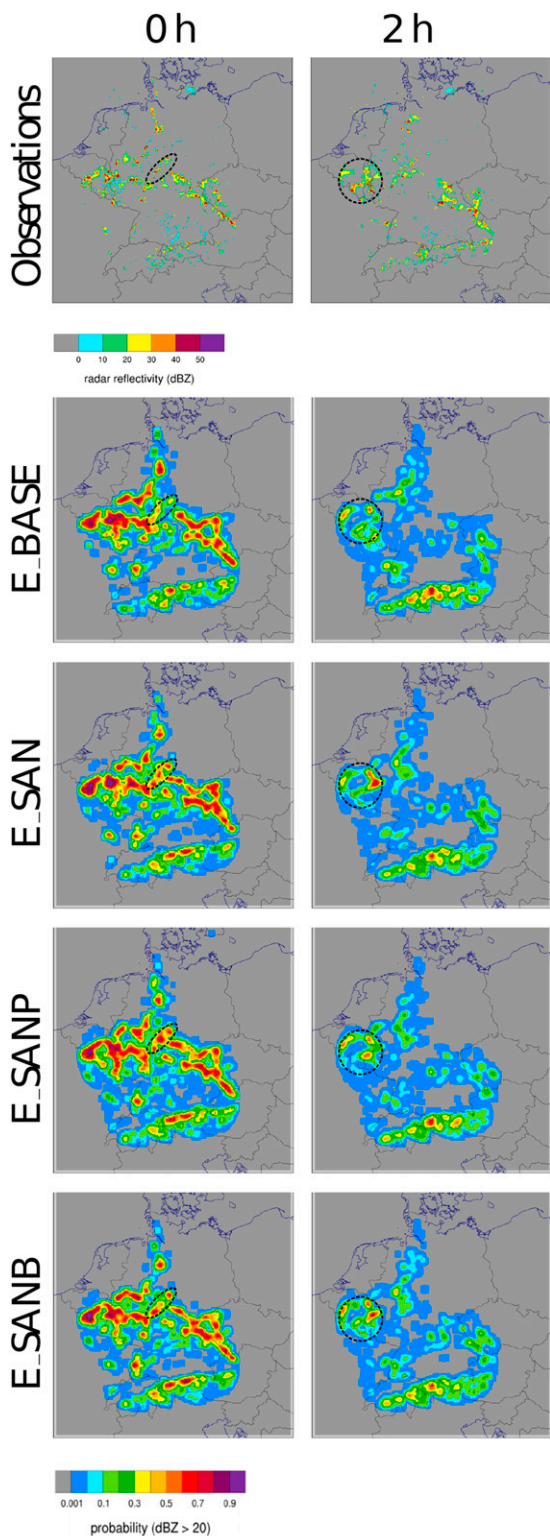


FIG. 12. Radar reflectivity composite at the elevation 0.5° at 1400 UTC (initial time) and 1600 UTC 6 Jun 2016 (2-h forecast). (top row) Observations; (row two) ensemble probability of E_BASE for threshold 20 dBZ; (rows three–five) as in row 2, but for E_SAN, E_SANP, and E_SANB.

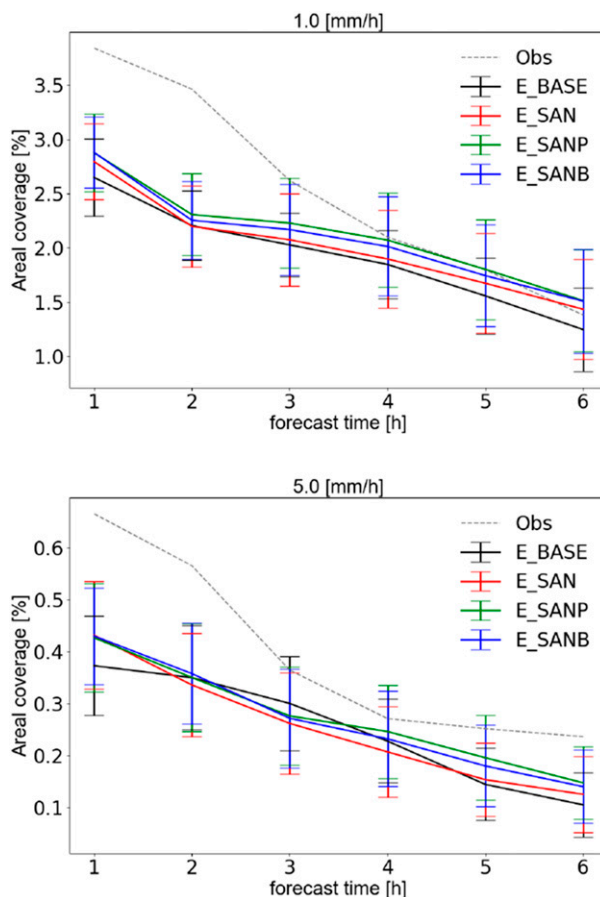


FIG. 13. As in Fig. 8, but for comparison of E_BASE, E_SAN, E_SANP, and E_SANB with precipitation observations.

method to improve the quality of precipitation forecasts although it is associated with fairly imbalanced states as shown in Zeng et al. (2019). However, imbalanced states do not necessarily lead to better precipitation forecasts, for instance, the RTPS as shown in Zeng et al. (2018). It may be worth exploring possibilities that can improve the balance of analyses of existing methods (e.g., the SAN and PSP) while maintaining the skills of precipitation forecasts. Furthermore, all these three methods mainly introduce small-scale perturbations that usually will be dumped within 6 h. The potential influences on forecasts of longer lead time also require further investigation. At last, it should be noted that Sobash and Wicker (2015) proposed an additive noise method to add the storm-scale random noise where the absolute innovation (observation–first guess) is greater than a certain value (i.e., 10 dBZ). This method also can automatically increase the storm through the comparison between observation and first guess. Moreover, this method can additionally suppress the spurious storms through assimilation of no reflectivity data, since it also

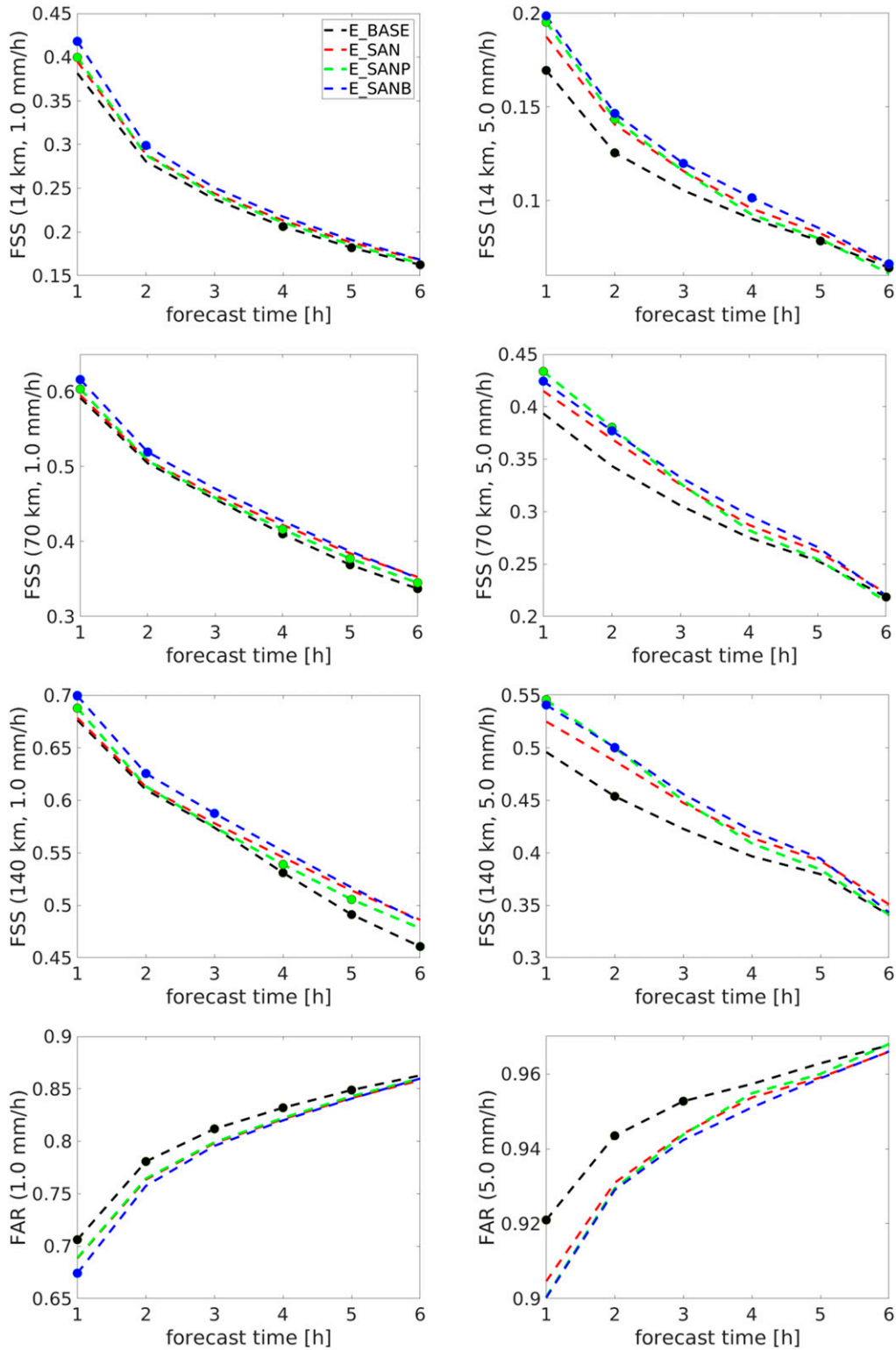


FIG. 14. As in Fig. 9, but for comparison of E_BASE, E_SAN, E_SANP, and E_SANB in Study 3. Recall that E_SAN is the reference run instead of E_BASE.

Downloaded from <http://journals.ametsoc.org/mwr/article-pdf/148/6/2457/4947966/mwr148064.pdf> by guest on 29 August 2020

enhances the spread at these areas as shown in [Hu et al. \(2019\)](#). We have tested this method with the random noise using model truncation error samples in [Zeng et al. \(2019\)](#), but no satisfactory results were obtained, but it may be worth combining it with the SAN or PSP.

Acknowledgments. The study was partially supported by the Transregional Collaborative Research Center SFB/TRR 165 Waves to Weather funded by the German Science Foundation (DFG) and partially supported by the DFG Priority Program 2115: PROM through the project JA 1077/5-1. The work of T. Janjic is supported through DFG Heisenberg Programm JA 1077/4-1. The work of A. de Lozar has been supported by the Innovation Programm for applied Research and Development (IAFE) of Deutscher Wetterdienst in the framework of the SINFONY project. Thanks are given to Hans Ertel Centre for Weather Research ([Weissmann et al. 2014](#); [Simmer et al. 2016](#)). Thanks are also given to Kevin Bachmann and Tobias Selz from LMU for the code of spectrum computation. The assimilated data used were obtained from the DWD.

REFERENCES

- Aksoy, A., D. C. Dowell, and C. Snyder, 2009: A multiscale comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analyses. *Mon. Wea. Rev.*, **137**, 1805–1824, <https://doi.org/10.1175/2008MWR2691.1>.
- Anderson, J., 2009: Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus*, **61A**, 72–83, <https://doi.org/10.1111/j.1600-0870.2008.00361.x>.
- , T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The Data Assimilation Research Testbed: A community facility. *Bull. Amer. Meteor. Soc.*, **90**, 1283–1296, <https://doi.org/10.1175/2009BAMS2618.1>.
- Baldauf, M., A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, 2011: Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities. *Mon. Wea. Rev.*, **139**, 3887–3905, <https://doi.org/10.1175/MWR-D-10-05013.1>.
- Berner, J., K. R. Fossell, S.-Y. Ha, J. P. Hacker, and C. Snyder, 2015: Increasing the skill of probabilistic forecasts: Understanding performance improvements from model-error representations. *Mon. Wea. Rev.*, **143**, 1295–1320, <https://doi.org/10.1175/MWR-D-14-00091.1>.
- Bick, T., and Coauthors, 2016: Assimilation of 3D radar reflectivities with an ensemble Kalman filter on the convective scale. *Quart. J. Roy. Meteor. Soc.*, **142**, 1490–1504, <https://doi.org/10.1002/qj.2751>.
- Bierdel, J., P. Friederichs, and S. Bentzien, 2012: Spatial kinetic energy spectra in the convection-permitting limited-area NWP model COSMO-DE. *Meteor. Z.*, **21**, 245–258, <https://doi.org/10.1127/0941-2948/2012/0319>.
- Bouttier, F., B. Vié, O. Nuissier, and L. Raynaud, 2012: Impact of stochastic physics in a convection-permitting ensemble. *Mon. Wea. Rev.*, **140**, 3706–3721, <https://doi.org/10.1175/MWR-D-12-00031.1>.
- Bowler, N. E., and Coauthors, 2017: Inflation and localization tests in the development of an ensemble of 4D-ensemble variational assimilations. *Quart. J. Roy. Meteor. Soc.*, **143**, 1280–1302, <https://doi.org/10.1002/qj.3004>.
- Buizza, R., M. Milleer, and T. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908, <https://doi.org/10.1002/qj.49712556006>.
- Denis, B., J. Cote, and R. Laprise, 2002: Spectral decomposition of two-dimensional atmospheric fields on limited-area domains using the discrete cosine transform (DCT). *Mon. Wea. Rev.*, **130**, 1812–1829, [https://doi.org/10.1175/1520-0493\(2002\)130<1812:SDOTDA>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<1812:SDOTDA>2.0.CO;2).
- Doms, G., and M. Baldauf, 2018: A description of the non-hydrostatic regional COSMO-Model. Part I: Dynamics and numerics. Consortium for Small-Scale Modelling Rep. COSMO-Model 5.5, 167 pp., <http://www.cosmo-model.org/content/model/documentation/core/cosmoDyncsNumcs.pdf>.
- , and Coauthors, 2011: A description of the nonhydrostatic regional COSMO model. Part II: Physical parameterization. Consortium for Small-Scale Modelling Rep. LM_F90 4.20, 161 pp., <http://www.cosmo-model.org/content/model/documentation/core/cosmoPhysParamtr.pdf>.
- Dowell, D. C., and L. J. Wicker, 2009: Additive noise for storm-scale ensemble data assimilation. *J. Atmos. Oceanic Technol.*, **26**, 911–927, <https://doi.org/10.1175/2008JTECHA1156.1>.
- , —, and D. J. Stensrud, 2004: High resolution analyses of the 8 May 2003 Oklahoma City Storm. Part II: EnKF data assimilation and forecast experiments. Preprints, *22nd Conf. on Severe Local Storms*, Hyannis, MA, Amer. Meteor. Soc., 12.5, <http://ams.confex.com/ams/pdfpapers/81393.pdf>.
- Duda, J. D., X. Wang, F. Kong, M. Xue, and J. Berner, 2016: Impact of a stochastic kinetic energy backscatter scheme on warm season convection-allowing ensemble forecasts. *Mon. Wea. Rev.*, **144**, 1887–1908, <https://doi.org/10.1175/MWR-D-15-0092.1>.
- Efron, B., 1987: Better bootstrap confidence intervals. *J. Amer. Stat. Assoc.*, **82**, 171–185, <https://doi.org/10.1080/01621459.1987.10478410>.
- Errico, R. M., 1985: Spectra computed from a limited area grid. *Mon. Wea. Rev.*, **113**, 1554–1562, [https://doi.org/10.1175/1520-0493\(1985\)113<1554:SCFALA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1985)113<1554:SCFALA>2.0.CO;2).
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, <https://doi.org/10.1029/94JC00572>.
- Fujita, T., D. J. Stensrud, and D. C. Dowell, 2007: Surface data assimilation using an ensemble Kalman filter approach with initial condition and model physics uncertainties. *Mon. Wea. Rev.*, **135**, 1846–1868, <https://doi.org/10.1175/MWR3391.1>.
- Ha, S., J. Berner, and C. Snyder, 2015: A comparison of model error representation in mesoscale ensemble data assimilation. *Mon. Wea. Rev.*, **143**, 3893–3911, <https://doi.org/10.1175/MWR-D-14-00395.1>.
- Hamill, T. M., and J. S. Whitaker, 2005: Accounting for the error due to unresolved scales in ensemble data assimilation. *Mon. Wea. Rev.*, **133**, 3132–3147, <https://doi.org/10.1175/MWR3020.1>.
- Hamrud, M., M. Bonavita, and L. Isaksen, 2015: EnKF and hybrid gain ensemble data assimilation. Part I: EnKF implementation. *Mon. Wea. Rev.*, **143**, 4847–4864, <https://doi.org/10.1175/MWR-D-14-00333.1>.
- Hirt, M., S. Rasp, U. Blahak, and G. C. Craig, 2019: Stochastic parameterization of processes leading to convective initiation

- in kilometer-scale models. *Mon. Wea. Rev.*, **147**, 3917–3934, <https://doi.org/10.1175/MWR-D-19-0060.1>.
- Houtekamer, P. L., H. L. Mitchell, and X. Deng, 2009: Model error representation in an operational ensemble Kalman filter. *Mon. Wea. Rev.*, **137**, 2126–2143, <https://doi.org/10.1175/2008MWR2737.1>.
- Hu, J., N. Yussouf, D. D. Turner, T. A. Jones, and X. Wang, 2019: Impact of ground-based remote sensing boundary layer observations on short-term probabilistic forecasts of a tornadic supercell event. *Wea. Forecasting*, **34**, 1453–1476, <https://doi.org/10.1175/WAF-D-18-0200.1>.
- Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126, <https://doi.org/10.1016/j.physd.2006.11.008>.
- Keil, C., F. Bauer, K. Bachmann, S. Rasp, L. Schneider, and C. Barthlott, 2019: Relative contribution of soil moisture, boundary layer and microphysical perturbations on convective predictability in different weather regimes. *Quart. J. Roy. Meteor. Soc.*, **145**, 3102–3115, <https://doi.org/10.1002/qj.3607>.
- Kober, K., and G. C. Craig, 2016: Physically-based stochastic perturbations (PSP) in the boundary layer to represent uncertainty in convective initiation. *J. Atmos. Sci.*, **73**, 2893–2911, <https://doi.org/10.1175/JAS-D-15-0144.1>.
- Lange, H., and T. Janjić, 2016: Assimilation of Mode-S EHS aircraft observations in COSMO-KENDA. *Mon. Wea. Rev.*, **144**, 1697–1711, <https://doi.org/10.1175/MWR-D-15-0112.1>.
- , G. C. Craig, and T. Janjić, 2017: Characterizing noise and spurious convection in convective data assimilation. *Quart. J. Roy. Meteor. Soc.*, **143**, 3060–3069, <https://doi.org/10.1002/qj.3162>.
- Leutbecher, M., R. Buizza, and L. Isaksen, 2007: Ensemble forecasting and flow-dependent estimates of initial uncertainty. *Proc. ECMWF Workshop on Flow-Dependent Aspects of Data Assimilation*, Reading, United Kingdom, ECMWF, 185–201.
- Lin, Y., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.*, **22**, 1065–1092, [https://doi.org/10.1175/1520-0450\(1983\)022<1065:BPOTSF>2.0.CO;2](https://doi.org/10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2).
- Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys.*, **20**, 851–887, <https://doi.org/10.1029/RG020i004p00851>.
- Meng, Z., and F. Zhang, 2008: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part III: Comparison with 3DVAR in a real-data case study. *Mon. Wea. Rev.*, **136**, 522–540, <https://doi.org/10.1175/2007MWR2106.1>.
- Murphy, J. M., D. M. Sexton, G. S. Barnett, D. N. Jones, M. J. Webb, M. Collins, and D. Sainforth, 2004: Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, **430**, 768–772, <https://doi.org/10.1038/nature02771>.
- Piccolo, C., and M. Cullen, 2016: Ensemble data assimilation using a unified representation of model error. *Mon. Wea. Rev.*, **144**, 213–224, <https://doi.org/10.1175/MWR-D-15-0270.1>.
- , —, W. Tennant, and A. Semple, 2018: Comparison of different representations of model error in ensemble forecasts. *Quart. J. Roy. Meteor. Soc.*, **145**, 15–27, <https://doi.org/10.1002/qj.3348>.
- Raschendorfer, M., 2001: The new turbulence parameterization of LM. COSMO Newsletter, Vol. 1, Consortium for Small-Scale Modeling, Offenbach, Germany, 89–97, http://www.cosmo-model.org/content/model/documentation/newsLetters/newsLetter01/newsLetter_01.pdf.
- Rasp, S., T. Selz, and G. C. Craig, 2018: Variability and clustering of midlatitude summertime convection: Testing the Craig and Cohen theory in a convection-permitting ensemble with stochastic boundary layer perturbations. *J. Atmos. Sci.*, **75**, 691–706, <https://doi.org/10.1175/JAS-D-17-0258.1>.
- Reinhardt, T., and A. Seifert, 2006: A three-category ice scheme for LMK. COSMO Newsletter, Vol. 6, Consortium for Small-Scale Modeling, Offenbach, Germany, 115–120, http://www.cosmo-model.org/content/model/documentation/newsLetters/newsLetter06/cnl6_reinhardt.pdf.
- Reynolds, C. A., J. Teixeira, and J. G. Mclay, 2008: Impact of stochastic convection on the ensemble transform. *Mon. Wea. Rev.*, **136**, 4517–4526, <https://doi.org/10.1175/2008MWR2453.1>.
- Rhodin, A., H. Lange, R. Potthast, and T. Janjić, 2013: Documentation of the DWD data assimilation system. Deutscher Wetterdienst (DWD), 448 pp.
- Roberts, N. M., and H. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, <https://doi.org/10.1175/2007MWR2123.1>.
- Schättler, U., 2014: A description of the nonhydrostatic regional COSMO-model. Part V: Preprocessing: Initial and boundary data for the COSMO-Model. Tech. Rep. INT2LM 2.1, Consortium for Smallscale Modeling (COSMO), 78 pp., https://www.hzg.de/imperia/md/assets/clm/2015int2lm_2.01.pdf.
- Schraff, C., H. Reich, A. Rhodin, A. Schomburg, K. Stephan, A. Periañez, and R. Potthast, 2016: Kilometre-scale Ensemble Data Assimilation for the COSMO model (KENDA). *Quart. J. Roy. Meteor. Soc.*, **142**, 1453–1472, <https://doi.org/10.1002/qj.2748>.
- Selz, T., L. Bierdel, and G. C. Craig, 2019: Estimation of the variability of mesoscale energy spectra with three years of COSMO-DE analyses. *J. Atmos. Sci.*, **76**, 627–637, <https://doi.org/10.1175/JAS-D-18-0155.1>.
- Shutts, G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quart. J. Roy. Meteor. Soc.*, **131**, 3079–3102, <https://doi.org/10.1256/qj.04.106>.
- Simmer, C., and Coauthors, 2016: HErZ: The German Hans-Ertel Centre for weather research. *Bull. Amer. Meteor. Soc.*, **97**, 1057–1068, <https://doi.org/10.1175/BAMS-D-13-00227.1>.
- Sobash, R., and L. Wicker, 2015: On the impact of additive noise in storm-scale EnKF experiments. *Mon. Wea. Rev.*, **143**, 3067–3086, <https://doi.org/10.1175/MWR-D-14-00323.1>.
- Weisman, M. L., and J. B. Klemp, 1982: The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Mon. Wea. Rev.*, **110**, 504–520, [https://doi.org/10.1175/1520-0493\(1982\)110<0504:TDONSC>2.0.CO;2](https://doi.org/10.1175/1520-0493(1982)110<0504:TDONSC>2.0.CO;2).
- Weissmann, M., and Coauthors, 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. *Meteor. Z.*, **23**, 193–208, <https://doi.org/10.1127/0941-2948/2014/0558>.
- Whitaker, J. S., and T. M. Hamill, 2012: Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Wea. Rev.*, **140**, 3078–3089, <https://doi.org/10.1175/MWR-D-11-00276.1>.
- , —, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP global forecasting system. *Mon. Wea. Rev.*, **136**, 463–482, <https://doi.org/10.1175/2007MWR2018.1>.
- Zängl, G., D. Reinert, P. Ripodas, and M. Baldauf, 2015: The ICON (ICOSahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Quart. J. Roy. Meteor. Soc.*, **141**, 563–579, <https://doi.org/10.1002/qj.2378>.

- Zeng, Y., U. Blahak, M. Neuper, and D. Jerger, 2014: Radar beam tracing methods based on atmospheric refractive index. *J. Atmos. Oceanic Technol.*, **31**, 2650–2670, <https://doi.org/10.1175/JTECH-D-13-00152.1>.
- , —, 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, <https://doi.org/10.1002/qj.2904>.
- , T. Janjić, 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, <https://doi.org/10.1029/2018MS001375>.
- , —, M. Sommer, A. de Lozar, U. Blahak, and A. Seifert, 2019: Representation of model error in convective-scale data assimilation: Additive noise based on model truncation error. *J. Adv. Model. Earth Syst.*, **11**, 752–770, <https://doi.org/10.1029/2018MS001546>.
- 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**, 1238–1253, [https://doi.org/10.1175/1520-0493\(2004\)132<1238:IOIEAO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2).

Copyright of Monthly Weather Review is the property of American Meteorological Society and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.