A Localized Adaptive Particle Filter within an Operational NWP Framework

Roland Potthast, Anne Walter, and Andreas Rhodin

Data Assimilation Unit, Deutscher Wetterdienst, Offenbach am Main, Germany

ABSTRACT

Particle filters are well known in statistics. They have a long tradition in the framework of ensemble data assimilation (EDA) as well as Markov chain Monte Carlo (MCMC) methods. A key challenge today is to employ such methods in a high-dimensional environment, since the naïve application of the classical particle filter usually leads to filter divergence or filter collapse when applied within the very high dimension of many practical assimilation problems (known as the curse of dimensionality). The goal of this work is to develop a localized adaptive particle filter (LAPF), which follows closely the idea of the classical MCMC or bootstrap-type particle filter, but overcomes the problems of collapse and divergence based on localization in the spirit of the local ensemble transform Kalman filter (LETKF) and adaptivity with an adaptive Gaussian resampling or rejuvenation scheme in ensemble space. The particle filter has been implemented in the data assimilation system for the global forecast model ICON at Deutscher Wetterdienst (DWD). We carry out simulations over a period of 1 month with a global horizontal resolution of 52 km and 90 layers. With four variables analyzed per grid point, this leads to 6.6×10^6 degrees of freedom. The LAPF can be run stably and shows a reasonable performance. We compare its scores to the operational setup of the ICON LETKF.

1. Introduction

Data assimilation is concerned with the use of observations in combination with a numerical model to determine the state of a dynamical system. In the framework of weather forecasting, data assimilation has a long history, ranging from the early works of Bjerknes (1904) and Richardson (1922) to modern ensemble data assimilation systems (cf. Bauer et al. 2015). Data assimilation links the model world with reality, usually by using a wide range of observational data to correct the development of the dynamical system step by step through an assimilation cycle and, thus, providing initial conditions for forecasting.

For an introduction into data assimilation methods, we refer to Kalnay (2003), Evensen (2009), Anderson and Moore (2012), van Leeuwen et al. (2015), Reich and Cotter (2015), and Nakamura and Potthast (2015). The history of data assimilation methods used within an operational framework started with optimal interpolation from the 1960s to the 1990s. Variational methods such as three-dimensional variational assimilation (3D-VAR)

have been employed operationally since about 1990, with four-dimensional variational assimilation (4D-VAR) since approximately 2000. Variational methods have the strength to calculate a best estimate for either one time slice (3D-VAR) or over some temporal window (4D-VAR). Ensemble data assimilation development started in the mid-1990s, with operational use since about 2010.

The ensemble Kalman filter (EnKF) was developed by Evensen (1994; see also Evensen and van Leeuwen 2000; Evensen 2009). The idea was applied to global numerical weather prediction by Houtekamer and Mitchell (1998, 2001, 2005) and Houtekamer et al. (2005). In the area of geophysical data assimilation, Burgers et al. (1998) developed the theoretical basis of the EnKF methods based on perturbations of the observations. Whitaker and Hamill (2002) proposed the alternative approach, called the ensemble square root filter (EnSRF). It does not use randomly perturbed observations, but formulates a deterministic calculation of the posterior ensemble.

Further variants of ensemble filters include, for example, the singular evolutive extended Kalman filter (SEEKF; Pham et al. 1998), the ensemble adjustment Kalman filter (EAKF; Anderson 2001), and the ensemble transform Kalman filter (ETKF; Bishop et al. 2001). Localization is a key ingredient of the ensemble Kalman filter (Houtekamer and Mitchell 1998; denoted as data

---

Corresponding author: Anne Walter, anne.walter@dwd.de

selection with a cutoff radius), the work by Brusdal et al. (2003), the local ensemble Kalman filter (LEKF; Ott et al. 2004), and the local ensemble transform Kalman filter (LETKF; Hunt et al. 2007), where all locally available observations are assimilated in one step. Various other forms of explicit filters are developed (e.g., the GIGG filter; Bishop 2016). For an overview of ensemble-based data assimilation methods, we refer to Vetra-Carvalho et al. (2018).

Important current research topics on ensemble Kalman filters are covariance localization and inflation (see van Leeuwen 2003b; Miyoshi et al. 2007; Miyoshi and Sato 2007; Campbell et al. 2010; Greybush et al. 2011; Janjić et al. 2011; Periáñez et al. 2014). Localization for particle filters has been described by, for example, Bengtsson et al. (2003) and van Leeuwen (2003a). Flow-adaptive localization has been described by Bishop and Hodyss (2007, 2009a,b) and Anderson (2007a); multiscale localization by Miyoshi and Kondo (2013); and flow-adaptive inflation by Anderson (2007b, 2009), Li et al. (2009), and Miyoshi (2011). The investigation of large ensembles has been carried out by, for example, Miyoshi et al. (2014).

Leaving the Gaussian regime, for which ensemble Kalman filters are best, particle filters take into account the full nonlinearity of both the model dynamics and observation operators in applications, leading to strongly non-Gaussian distributions on all temporal and spatial scales. Particle filters have a long history in stochastic modeling, where they have been used since the 1960s under the name of iterative Markov chain Monte Carlo (MCMC) methods (Bain and Crisan 2009; Crisan and Rozovskii 2011). The idea is to sample some probability distribution, where the number of samples reflects the local strength of the probability density, their weights are adapted using observations, and then resampling is carried out. Several particle filter methods have been formulated and tested for small-dimensional problems, ranging from early work by Gordon et al. (1993) to the review of van Leeuwen (2009). More recently, Ades and van Leeuwen (2013, 2015) have adapted their equivalent-weights particle filter (EWPF) to high-dimensional systems using a simple relaxation technique and a proposal to ensure equal weights for the particles. Zhu et al. (2016) have further improved the EWPF to their implicit equal weights particle filter (IEWPF). Further, Gaussian mixture models have been developed, based on the estimation of the model error probability distribution by a set of Gaussians (see Hoteit et al. 2008; Stordal et al. 2011 for a hybrid method of a Gaussian mixture and the particle filter). Frei and Künsch (2013) develop a hybrid method for an EnKF and a particle filter. For a recent review, we refer to van Leeuwen (2009) and Reich and Cotter (2015; cf. Nakamura and Potthast 2015).

It is well known that in a high-dimensional framework, particle filters suffer from so-called filter collapse or filter divergence under the curse of dimensionality (cf. van Leeuwen 2010; Snyder et al. 2008, 2015; Bickel et al. 2008). This means that usually, only very few or just one of the ensemble members carry all the weight in the assimilation step. This immediately destroys the diversity of the ensemble and leads to useless behavior when applied in an iterative way. Different ideas have been developed to overcome filter collapse: for example, guiding the particles to the right places in the high-dimensional space (van Leeuwen 2010). Also, using localization for particle filters has become popular (see, e.g., Reich and Cotter 2015; Poterjoy and Anderson 2016). Only recently, Robert et al. (2018) and Poterjoy (2016) started to apply particle filters in an operational environment for a large-scale operational global weather model. Robert et al. (2018) employ the data assimilation coding environment of Deutscher Wetterdienst (DWD) working with the convection-permitting COSMO Model. To avoid filter divergence, they combine the particle filter step with an ensemble Kalman filter step. Poterjoy (2016) works in a convective-scale framework in comparison to the global setup; he also employs localization as a core ingredient of his localized particle filter (see also Poterjoy et al. 2017).

Here, our goal is to develop a localized adaptive particle filter (LAPF) for global atmospheric data assimilation to fit into the framework of a global operational weather prediction model. We want to show that with appropriate adaptation, the classical particle filter can be a stable and useful approach in a large-scale operational framework.

Our reference is the implementation of the LETKF (Hunt et al. 2007) for the Icosahedral Nonhydrostatic (ICON) model. This has been operational at DWD since 20 January 2015. The ICON-LETKF provides initial conditions for the global ensemble prediction system (ICON-EPS) and feeds the dynamic covariance matrix of DWD's ensemble–variational data assimilation system (EnVAR). The basic ingredients of the LAPF will be described in detail in sections 3b–f. Section 3b describes the classical particle filter; section 3c describes the projection onto ensemble space to calculate the particle weights; section 3d describes a classical resampling step; section 3e describes spread control by optimal spread estimation; and section 3f describes a Gaussian rejuvenation or resampling step, carried out by modulated global draws around each particle after the first classical resampling step.

Further, in section 3a, we describe 1) the set of observation operators that are included in our assimilation tests, 2) the LETKF reference implementation, 3) the localization in observation space, 4) multiplicative inflation and relaxation to prior perturbation (RTPP), 5)

the assimilation grid and interpolation, 6) additive co-variance inflation, and 7) the incremental analysis up-date; all these components are important parts of the system.

Adaptive Gaussian rejuvenation in combination with optimal spread estimation acts as spread control for the ensemble, such that filter collapse or filter divergence can be avoided. It ensures the stability of the particle filter, in the sense that the ensemble spread stays in a feasible range, the filter does not crash, and its core task to syn-chronize the system to real observations works. The en-semble space projection helps to keep the ensemble weights in a reasonable range. The careful treatment of the rejuvenation by modulated global draws in ensemble space around each ensemble member after resampling keeps atmospheric fields intact for further integration.

By running the implementation for a test period of 1 month with a global horizontal resolution of 52 km, 90 vertical layers, and four variables analyzed per grid point (i.e., with state space dimension of approximately $n = 6.6 \times 10^6$), we demonstrate the feasibility of the method. A comparison of scores with the operational setup shows that the localized adaptive particle filter is able to provide a reasonable atmospheric analysis in a large-scale environment.

In section 2, we give an introduction into the ICON model setup and the ensemble prediction system ICON-EPS. Section 3 describes the localized adaptive particle filter in ensemble space with Gaussian resampling and spread control. The ensemble data assimilation (EDA) suite of DWD includes different tools, in particular multiplicative and additive covariance inflation, and various stochastic schemes for generating spread in surface fields. We survey our system and experimental setup in section 4 and evaluate the assimilation cycle and forecasts in detail in sections 4a and 4b, studying the development of spread, bias, scores, and the stability of the system. Finally, conclusions and further develop-ment steps are discussed in section 5.

2. The ICON model and ICON-EPS

In this part, we briefly introduce the ICON model in section 2a, the ICON-EPS in section 2b, and the pre-operational experimental setup in section 2c.

a. The ICON model and deterministic forecast system

ICON is the operational global numerical weather prediction (NWP) model of DWD. It is a joint project of DWD and the Max Planck Institute for Meteorology (MPI-M; Zängl et al. 2015). ICON is based on the prog-nostic variables suggested by Gassmann and Herzog [2008; for further information, see also Reinert et al.

(2018)], but instead of using the three-dimensional Lamb transformation, it uses the two-dimensional version to convert the nonlinear momentum advection into a vector-invariant form (Zängl et al. 2015).

The model grid is based on an unstructured triangular grid that is generated by successive refinement of a spherical icosahedron, which consists of 20 equilateral triangles with an edge length of about 7054 km. These triangles are subdivided (e.g., by bisection, trisection) into smaller triangles, leading to a model grid with the desired spatial resolution. The use of this icosahedral grid provides a nearly homogeneous coverage of the globe. After dividing the triangles, the operational grid of the ICON model consists of 2 949 120 triangles on each horizontal level. Each triangle has an average area of 173 km². This corresponds to the global horizontal resolution of 13 km for the deterministic run. For the operational setup, a two-way nested area over Europe with 6.5-km horizontal resolution is included. The en-sembles run with an operational horizontal resolution of 40 km, with a 20-km-resolution nest over Europe.

The scalar prognostic variables (e.g., temperature, humidity) are located in the center of the triangles, whereas the wind components are located at the edge midpoints of the triangles. The most important prog-nostic variables (e.g., wind, humidity, cloud water, cloud ice, temperature, snow, precipitation) are calculated for all grid cells on 90 terrain-following vertical model levels, which range from the surface up to a height of 75 km, leading to 265 million grid points in the opera-tional setup. Additional prognostic equations are solved over land on seven soil levels for soil temperature and soil water content. If snow is present, several snow var-iables are also determined. Once per day (at 0000 UTC), the sea surface temperature (just over ice-free ocean) is analyzed from observations and is kept constant during the forecasts. For the sea ice fraction of the ice-covered oceans, we proceed in the same way. However, ice thickness and ice surface temperature are determined by a simple sea ice model.

Beyond the adiabatic processes in the atmosphere (horizontal and vertical transport processes), diabatic processes (e.g., radiation, turbulence) play a major role in NWP. Describing these small-scale processes is part of the physics parameterization of ICON.

Within the operational workflow, we distinguish the data assimilation cycle and the forecast mode. During the data assimilation cycle, a 3-h forecast starting from the previous analyses (the first guess) is blended with all observations valid for a 3-h time window centered at the analysis date. It is important to note that traditionally, the atmospheric analysis calculates increments to four core variables (tem-perature, humidity, and two wind components) at each grid

point in its three-dimensional grid, with pressure adapted according to the hydrostatic equation. The model adapts further prognostic variables itself. Additional two-dimensional fields that are also adapted have been neglected in our variable counts. To obtain an optimal initial state for the subsequent forecasts, we calculate a variational analysis with a dynamic covariance matrix, known as ensemble variational data assimilation (EnVAR). In total, 70% of the covariance matrix is calculated from the ensemble runs based on a LETKF. Further, 30% of the covariance matrix is given by its climatological part based on the National Meteorological Center (NMC) method (Parrish and Derber 1992). The EnVAR and the LETKF were made operational on 20 January 2016.

Based on the analyses for 0000 and 1200 UTC, ICON provides a 180-h forecast in just 1-h wall-clock time. Forecasts over 120 h are based on the analyses of the 0600 and 1800 UTC run, and the 30-h forecasts are based on the 0300, 0900, 1500, and 2100 UTC analyses.

b. The ICON-EPS

ICON-EPS has run preoperationally since January 2016, providing background error correlations for the operational global EnVAR system of DWD, with operational forecasts including all ensemble products since 17 January 2018. It also has provided boundary conditions for the operational local-area kilometer-scale ensemble data assimilation (KENDA; Schraff et al. 2016) and ensemble prediction system COSMO-DE-EPS (Bouallègue et al. 2013) since March 2017.

The ICON-EPS initial conditions are provided by the LETKF, which is part of the hybrid data assimilation suite LETKF+EnVAR for the ICON model. The ICON-EPS is run up to 180 h at 0000 and 1200 UTC, up to 120 h at 0600 and 1800 UTC, and up to 30 h at the 3-hourly intervals in between, with the main purpose of generating forecasts and ensemble boundary conditions in the short range of up to 30-h lead time every 3 h.

c. Development environment, experimental setup, and period

The main goal of our work is to investigate the feasibility and performance of a stable particle filter for global numerical weather prediction with the ICON model in the above operational setup. For our experimental test, we chose the period of 1–31 May 2016.

Since quality control is carried out based on the deterministic run, it is always part of experiments. Here, for the development of the LAPF, we choose the established 52-km experimental resolution for the ensemble and 26 km for the deterministic run. In standard DWD experiments, EDA tests are usually run with the

operational ensemble size $L = 40$ members, which is also our choice for the LAPF development. In this setup, we study the performance and stability of the particle filter in direct comparison to the LETKF-based operational setup.

The development of the LAPF takes place in the data assimilation coding environment of DWD. The suite includes modules for snow analysis (SNOW) every 3 h, sea surface temperature (SST) analysis, and soil moisture analysis (SMA) once per day. The surface analysis consists of separate modules in which, among others, random perturbations are added to the ensemble members. For our experiments, this part has been kept identical to the operational setup (cf. Reinert et al. 2018).

3. An LAPF with Gaussian resampling

In this section, we introduce the localized adaptive particle filter with adaptive Gaussian resampling and spread control. First, section 3a describes the operational LETKF implementation, which serves as a reference for comparison and whose core algorithm is replaced by the LAPF algorithm. Section 3b presents the classical particle filter basis of the method, as an important first step, we describe the ensemble transform version of this particle filter in section 3c. Then, we go into details of classical resampling in section 3d and describe the indicator for spread control in section 3e and the adaptive Gaussian resampling or rejuvenation in section 3f. We note that the adaptive Gaussian rejuvenation is carried out on top of classical resampling (i.e., this is an additional tool for spread control added to the classical particle filter).

a. The operational LETKF implementation

In this section, we first briefly describe the operational ensemble data assimilation method, its components, and setup. For the LAPF implementation, the core LETKF algorithm II is replaced by a particle filter step, keeping the observation handling, quality control, and part of the inflation and localization facilities unchanged.

1) OBSERVATION OPERATORS AND ANALYSIS FREQUENCY

Observation operators are evaluated at 3-hourly intervals in the cycled data assimilation code. Observation types currently include TEMP, PILOT,[1] SHIP, SYNOP,[2]

---

[1] TEMPs and PILOTs are particular weather balloons measuring profiles of, for example, temperature and humidity.

[2] SYNOPs are the classical land-based weather stations measuring, for example, temperature, humidity, and pressure.

BUOY, wind profiler, aircraft, atmospheric motion vector (AMV),[3] radio occultations, scatterometer, and satellite radiances. For further information about the observation types, we refer to ECMWF (2014, p. 32). Observational quality control (observation minus first-guess check) and bias correction (for radiances and individual aircrafts) are performed in the deterministic data assimilation system. Bias-corrected observations and quality control flags are then further passed to the ensemble data assimilation system.

2) LETKF

The formulation of the LETKF at DWD is based on the proposal of Hunt et al. (2007). The implementation is shared with the regional KENDA system (Schraff et al. 2016) for the COSMO EPS. The ensemble Kalman filter equations are solved in ensemble space (39 dimensions in case of 40 members). In principle, the Kalman gain matrix uses the background error covariance matrix $\mathbf{P}^b$ in order to determine the analysis increment of the ensemble mean and the symmetric square root of the analysis ensemble covariance matrix $\mathbf{P}^a$. Practically, a weight matrix $\mathbf{W}$ is derived, which is used for the construction of the analysis ensemble as a linear combination of the forecast ensemble members. Since the LAPF implementation directly imitates the LETKF transform, we need to look into more detail here. The operational LETKF system implements Eqs. (20) and (21) of Hunt et al. (2007), that is,

$$\overline{\mathbf{w}}^a = \tilde{\mathbf{P}}^a (\mathbf{Y}^b)^{\mathrm{T}} \mathbf{R}^{-1} (\mathbf{y}^0 - \overline{\mathbf{y}}^b), \tag{1}$$

for calculating the mean of the analysis ensemble and $\tilde{\mathbf{P}}^a$ given by

$$\tilde{\mathbf{P}}^a = [(L-1)\mathbf{I} + (\mathbf{Y}^b)^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{Y}^b]^{-1}, \tag{2}$$

where we use the letter L for the number of ensemble members and the notation $\overline{\mathbf{w}}^a$ for the linear coefficients of the analysis mean, $\tilde{\mathbf{P}}^a$ denotes the $L \times L$ analysis covariance in the space of ensemble coefficients, $\mathbf{R}$ is the observation error covariance matrix, $\mathbf{y}^0$ are the observations, $\overline{\mathbf{y}}^b$ is the mean of the model equivalents $H\mathbf{x}^b$ of the observations, H is the observation operator, and $\mathbf{Y}^b$ is the matrix of ensembles minus mean in observation space. Equation (2) in model space leads to Eqs. (22) and (23) of Hunt et al. (2007):

$$\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^b + \mathbf{X}^b \overline{w}^a, \tag{3}$$

$$\mathbf{P}^a = \mathbf{X}^b \tilde{\mathbf{P}}^a (\mathbf{X}^b)^{\mathrm{T}}, \tag{4}$$

where $\overline{\mathbf{x}}^a$ is the analysis mean, and $\mathbf{X}^b$ is the matrix of ensemble minus its mean. The analysis ensemble is calculated as in (24) of Hunt et al. (2007). We obtain

$$\mathbf{X}^a = \mathbf{X}^b \mathbf{W} \tag{5}$$

using

$$\mathbf{W} = [(L-1)\tilde{\mathbf{P}}^a]^{1/2}, \tag{6}$$

with the symmetric square root denoted by the 1/2 power of the symmetric matrix $\tilde{\mathbf{P}}^a$. We note that a derivation of this algorithm with its links to classical inverse problems theory can also be found in Nakamura and Potthast (2015, chapter 5).

3) LOCALIZATION ON $\mathbf{R}$

Localization is performed by calculating independent analyses (weight matrices $\mathbf{W}$) at each analysis grid point using only the observations in the vicinity of that location. The observations are weighted smoothly in dependence on their distance to that point, according to a localization function chosen as the fifth-order polynomial described by Gaspari and Cohn (1999), which is similar to a Gaussian but has compact support. We use a horizontal localization length scale of 300 km and a vertical length scale varying from 0.3 [given in $\ln(p)$] at the surface to 0.8 at the model top (75 km). The length scales are defined following Daley (1993), using the second derivative of the localization function c at its origin: $l = \sqrt{-1/[\nabla^2 c(0)]}$. For a Gaussian, this coincides with the standard deviation of the distribution. Formally, the inverse of the observation error covariance matrix $\mathbf{R}$ [in (2), respectively (6)] is weighted by a pointwise multiplication with the function defined by Gaspari and Cohn, such that observations that are located at a larger distance from the current analysis grid point receive less weight when calculating the analysis. Therefore, this procedure is often denoted as localization on $\mathbf{R}$.

4) MULTIPLICATIVE INFLATION AND RTPP

The analysis ensemble spread is adjusted by multiplicative inflation with a factor ranging from 0.9 to 1.5, based on an online estimate of spread and ensemble-mean root-mean-square error (RMSE) in observation space, following Houtekamer et al. (2005). The inflation factor is estimated locally based on the statistics on observation minus first-guess differences, as described in section 3e, and the $\mathbf{W}$ matrices are adjusted respectively. In addition, an RTPP is applied following Whitaker and Hamill (2012), with a rate of 0.75. The latter preserves a

---

[3] Calculated from subsequent satellite images.

reasonable situation-dependent spread–skill relationship in the analysis cycle.

5) ASSIMILATION GRID AND INTERPOLATION

Tapering the observations with a smooth function and taking the symmetric square root in the LETKF algorithm ensures that the weight matrices only change on scales on the order of (or larger than) the localization length scale. For this reason, it is sufficient to derive the weight matrices **W** on a coarser analysis grid $\mathscr{G}$ (e.g., Yang et al. 2009), with spacing on the order of this prescribed length scale. Afterward, the **W** are interpolated to the model grid, and the final analyses are derived by taking linear combinations of the forecast ensemble members according to the interpolated weight matrices.

6) ADDITIVE COVARIANCE INFLATION

To account for model errors, additive random perturbations consistent with 25% of the amplitude of the climatological **B** matrix used in the deterministic EnVAR assimilation system are added to the analysis ensemble members. In addition, the SST is perturbed by random perturbations of 1 K, which are a linear combination of perturbations with spatial correlation length scales of 100 and 1000 km and a time scale of 1 day.

7) INCREMENTAL ANALYSIS UPDATE

The analysis increments applied by the cycled data assimilation system, as well as the stochastic perturbations, introduce imbalances and spinup effects, which are diminished by using an incremental analysis update (IAU) scheme (see Bloom et al. 1996). The combined analysis increments from the LETKF, inflation schemes, and additive perturbations are added to the model trajectory, dispensing them over a time interval of 3 h, symmetrically adjusted around the analysis date.

b. *The classical particle filter*

The concept of the LAPF is chosen to be as parallel as possible to the ensemble transform Kalman filter described in sections 3a(1)–(7) above. Here, we basically replace the analysis step 2 by the steps described in sections 3b–f.

Let us recall the stochastic notation and background of the particle filter (following Nakamura and Potthast 2015), where $\mathbf{x}_k \in \mathbb{R}^n$ denotes model states of dimension n, and $\mathbf{y}_k \in \mathbb{R}^m$ is the vector of m observations at time t_k. Bayesian data assimilation starts with some prior distribution. The analysis step employs observations to derive an analysis distribution. This analysis distribution

is propagated to the next time step, where it acts as a prior distribution for the subsequent assimilation step. For our implementation, we consider the analysis distribution $p^{(a)}(\mathbf{x})$ for time t_0 as initial state. Then, the analysis distribution is propagated in time by a short-range ensemble forecast, from t_{k-1} to t_k, to get the first-guess distribution $p_k^{(b)}(\mathbf{x})$. Afterward, the Bayes formula

$$p(\mathbf{x}|\mathbf{y}_k) = \frac{p_k^{(b)}(\mathbf{x})p(\mathbf{y}_k|\mathbf{x})}{p(\mathbf{y}_k)}, \quad \mathbf{x} \in \mathbb{R}^n, \quad \mathbf{y}_k \in \mathbb{R}^m,$$

is employed to calculate the new analysis distribution

$$p_k^{(a)}(\mathbf{x}) := p(\mathbf{x}|\mathbf{y}_k) = cp(\mathbf{y}_k|\mathbf{x})p_k^{(b)}(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n, \quad (7)$$

where c is a normalization constant so that

$$\int_{\mathbb{R}^n} p_k^{(a)}(\mathbf{x})\, d\mathbf{x} = 1. \quad (8)$$

The classical particle filter uses an ensemble $\mathbf{x}^{(l)}$ of states, which represents the prior probability distribution $p_k^{(b)}$ at time t_k in the form of δ functions. Alternatively, particles are considered as draws from this prior distribution, and in the case of L particles, they each carry a weight of $1/L$. To carry out the analysis step at time t_k, weights are calculated by

$$w_{k,\ell} := cp(\mathbf{y}_k|\mathbf{x}^{(l)}), \quad \ell = 1, \dots, L, \quad (9)$$

for the particles $\mathbf{x}^{(l)}$ corresponding to (7), where c is a normalization constant. We note that sometimes we use the normalization to L for easier discussion (i.e., $\sum_{\ell=1,\dots,L} w_{k,\ell} = L$). Then, for the prior each particle carries the weight 1.

c. *Projection onto ensemble space*

In the following, we drop the explicit declaration of the time index k and write $\mathbf{y}^o$ for the observations $\mathbf{y}_k$ at time t_k.

As seen from (1), the LETKF is based on the projection of the observation onto ensemble space. For brevity, we use **Y** for $\mathbf{Y}^b$. We note that the orthogonal projection of the observation difference $\mathbf{y}^o - \bar{\mathbf{y}}^b$ onto the ensemble space $\{\mathbf{Yw} : \mathbf{w} \in \mathbb{R}^L\}$ with respect to the scalar product weighted by $\mathbf{R}^{-1}$ in $\mathbb{R}^m$ is given by

$$P(\mathbf{y}^o - \bar{\mathbf{y}}^b) = \mathbf{Y}(\mathbf{Y}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Y})^{-1}\mathbf{Y}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}^o - \bar{\mathbf{y}}^b). \quad (10)$$

Compare, for example, lemma 3.2.3 of Nakamura and Potthast (2015) using the adjoint $\mathbf{Y}^* = \mathbf{Y}^{\mathrm{T}}\mathbf{R}^{-1}$

based on the weighted scalar product in $\mathbb{R}^m$ and the Euclidean scalar product in the space $\mathbb{R}^L$ of ensemble coefficients.[4] The corresponding particle filter weights (9) at time t_k based on this projection are given by its ensemble transform projection:

$$w_{k,\ell} := ce^{-1/2[P(\mathbf{y}^o - H\mathbf{x}^{(\ell)})]^T \mathbf{R}^{-1}[P(\mathbf{y}^o - H\mathbf{x}^{(\ell)})]}, \quad \ell = 1, \ldots, L,$$
(11)

with normalization constant c. Abbreviating $A := \mathbf{Y}^T\mathbf{R}^{-1}\mathbf{Y}$ and $C := A^{-1}\mathbf{Y}^T\mathbf{R}^{-1}(\mathbf{y}^o - \overline{\mathbf{y}}^b)$, we first note

$$\mathbf{y}^o - H\mathbf{x}^{(\ell)} = \mathbf{y}^o - (\overline{\mathbf{y}}^b + \mathbf{Y}\mathbf{e}_\ell) = (\mathbf{y}^o - \overline{\mathbf{y}}^b) - \mathbf{Y}\mathbf{e}_\ell,$$
$$\ell = 1, \ldots, L,$$
(12)

and

$$P(\mathbf{y}^o - H\mathbf{x}^{(\ell)}) = \mathbf{Y}A^{-1}\mathbf{Y}^T\mathbf{R}^{-1}[(\mathbf{y}^o - \overline{\mathbf{y}}^b) - \mathbf{Y}\mathbf{e}_\ell]$$
$$= \mathbf{Y}(C - \mathbf{e}_\ell), \quad \ell = 1, \ldots, L,$$
(13)

where $\mathbf{e}_\ell$ is the standard unit vector, which is 1 in its ℓth component and zero otherwise. Now, the exponent of (11) is transformed into

$$[P(\mathbf{y}^o - H\mathbf{x}^{(\ell)})]^T\mathbf{R}^{-1}P(\mathbf{y}^o - H\mathbf{x}^{(\ell)}) = [C - \mathbf{e}_\ell]^T A[C - \mathbf{e}_\ell],$$
$$\ell = 1, \ldots, L,$$
(14)

leading to

$$w_{k,\ell} = ce^{-1/2[C - \mathbf{e}_\ell]^T A[C - \mathbf{e}_\ell]}, \quad \ell = 1, \ldots, L.$$
(15)

The classical weight (9) is known to lead to filter divergence in high-dimensional spaces. Here, the ensemble transform and the projection P onto ensemble space lead to a significant reduction of the dimensionality. The observation vector $\mathbf{y} \in \mathbb{R}^m$ is mapped onto the vector $A^{-1}\mathbf{Y}^T\mathbf{R}^{-1}(\mathbf{y}^o - \overline{\mathbf{y}}^b)$ in the space $\mathbb{R}^L$ with ensemble size L. The weights (11) now penalize the distance of the ensemble members $\mathbf{e}_\ell$, $\ell = 1, \ldots, L$ in $\mathbb{R}^L$ to the projection C of the observations onto ensemble space. The histograms in Fig. 1 show the result of this projection step, which, in combination with adaptive Gaussian resampling and localization, leads to a feasible behavior of the particle filter weights.

To evaluate the relationship between the classical particle filter weights and the ensemble space particle filter weights, we note

$$w_{k,\ell}^{\text{classical}} = ce^{-1/2[(\mathbf{y}^o - H\mathbf{x}^{(\ell)})]^T\mathbf{R}^{-1}[(\mathbf{y}^o - H\mathbf{x}^{(\ell)})]}$$
$$= ce^{-1/2\{[P + (I-P)](\mathbf{y}^o - H\mathbf{x}^{(\ell)})^T\mathbf{R}^{-1}\{[P + (I-P)](\mathbf{y}^o - H\mathbf{x}^{(\ell)})\}}$$
$$= ce^{-1/2[P(\mathbf{y}^o - H\mathbf{x}^{(\ell)})]^T\mathbf{R}^{-1}[P(\mathbf{y}^o - H\mathbf{x}^{(\ell)})]} \underbrace{e^{-1/2[(I-P)(\mathbf{y}^o - H\mathbf{x}^{(\ell)})]^T\mathbf{R}^{-1}[(I-P)(\mathbf{y}^o - H\mathbf{x}^{(\ell)})]}}_{=\tilde{c}},$$
(16)

where we use the orthogonality of the projection P with respect to the scalar product with weight $\mathbf{R}^{-1}$ such that the mixed terms of P with $I - P$ vanish. The second exponential factor in the last line of (16) is equal to a constant $\tilde{c}$ for all $\ell = 1, \ldots, L$ since we have

$$(I - P)(\mathbf{y}^o - H x^{(\ell)}) = (I - P)(\mathbf{y}^o - \overline{\mathbf{y}}^b + \mathbf{Y}\mathbf{e}_\ell)$$
$$= (I - P)(\mathbf{y}^o - \overline{\mathbf{y}}^b) - \underbrace{(I - P)\mathbf{Y}\mathbf{e}_\ell}_{=0}.$$
(17)

If the ensemble space spans only a small part of the full state space, the constant $\tilde{c}$ can be very small, so the ensemble transformation effectively removes

---

[4] This is readily obtained from $\langle \mathbf{z}, \mathbf{Y}\mathbf{w} \rangle_{\mathbf{R}^{-1}} = \langle \mathbf{z}, \mathbf{R}^{-1}\mathbf{Y}\mathbf{w} \rangle = \langle \mathbf{Y}^T\mathbf{R}^{-1}\mathbf{z}, \mathbf{w} \rangle = \langle \mathbf{Y}^*\mathbf{z}, \mathbf{w} \rangle$, where $\langle \cdot, \cdot \rangle$ denotes an inner product and where $\mathbf{z} \in \mathbb{R}^m$ and $\mathbf{w} \in \mathbb{R}^L$. See Nakamura and Potthast (2015) for a detailed introduction.

a very small but uniform factor from the ensemble weights.

After the determination of the weights [(15)], the classical resampling (section 3d) is carried out. Then, the spread control (section 3e) will be prepared. Subsequently, the adaptive Gaussian resampling step (section 3f) will be executed.

d. Classical resampling

The LAPF carries out a classical resampling step based on (15), suggested already by Gordon et al. (1993). For resampling, accumulated weights w_{ac_ℓ}, $\ell = 1, \ldots, L$, are defined by

$$w_{\text{ac}_0} = 0, \qquad w_{\text{ac}_\ell} = w_{\text{ac}_{\ell-1}} + w_{k,\ell}, \qquad \ell = 1, \ldots, L, \quad (18)$$

where now we employ normalization to the total weight of L. Then, similar to Doucet et al. (2001; see also Elsheikh et al. 2014; Crisan and Rozovskii 2011), we
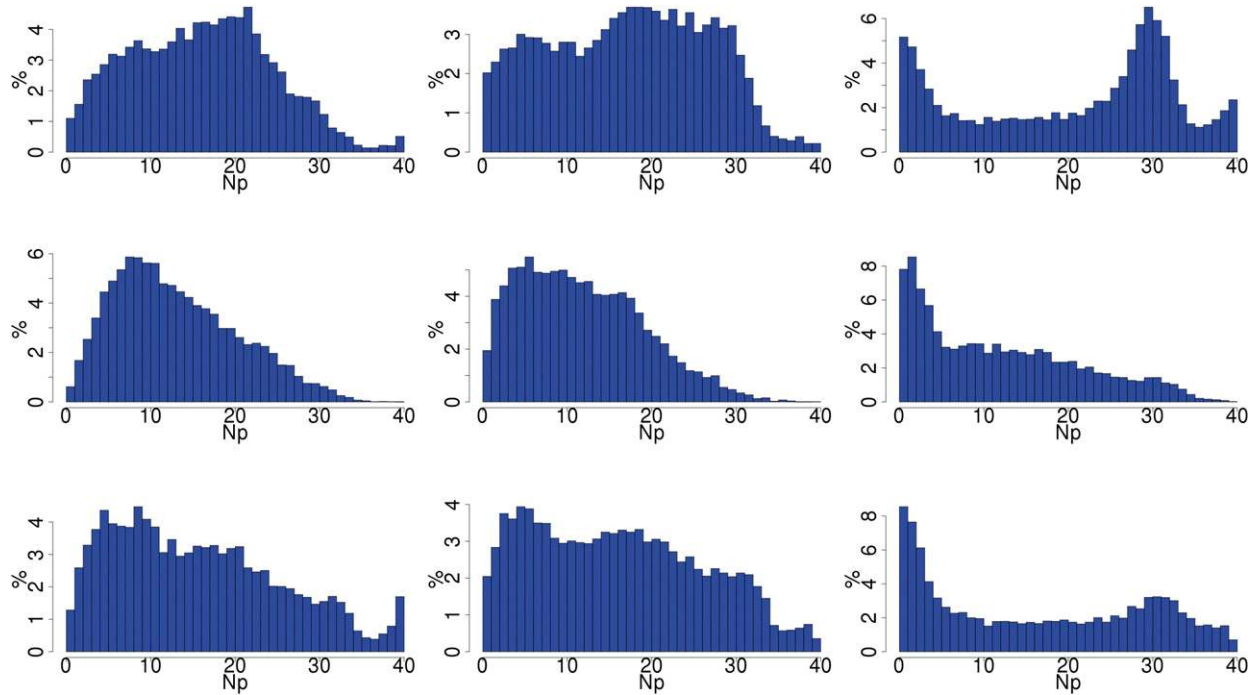
FIG. 1. Global histograms of the number of particles with weights above 1 (when the total weight is given by the number of particles L) for three dates and three levels: (top) 0300 UTC 20 May, (middle) 1200 UTC 25 May, and (bottom) 2100 UTC 31 May 2016 at levels (left) 100, (center) 500, and (right) 1000 hPa. The x axis shows the number of particles (Np) with weight larger than 1, and the y axis shows the percentage of analysis grid points with these numbers.

draw $r_\ell \sim U([0, 1])$, $\ell = 1, \ldots, L$, set $R_\ell = \ell - 1 + r_\ell$, and define the transform matrix for the particles by

$$\breve{\mathbf{W}}_{i,\ell} = \begin{cases} 1, & \text{if } R_\ell \in \left(w_{\mathrm{ac}_{\ell-1}}, w_{\mathrm{ac}_\ell} \right], \\ 0, & \text{otherwise}, \end{cases} \qquad (19)$$

$i, \ell = 1, \ldots, L$ with $\breve{\mathbf{W}} \in \mathbb{R}^{L \times L}$, where $(s, t]$ denotes the interval of values $s < \eta \le t$. This is carried out for each analysis grid point $p \in \mathscr{G}$; for brevity, however, we use $\breve{\mathbf{W}}$ instead of $\breve{\mathbf{W}}(p)$.

e. Spread control

In ensemble data assimilation systems, the spread of the ensemble evolves as a result of model dynamics, model errors (represented by additive perturbations or multiplicative inflation), and active observations[5] and thus relies on a correct specification of model and observational errors. As it is very difficult to properly estimate and model these errors, the spread of the ensemble is adjusted. In the operational LETKF, an adaptive inflation factor ρ is estimated, based on statistics of observations minus first guess (cf. Desroziers et al.

2005; Li et al. 2009). For this purpose, we use adaptive Gaussian resampling with parameters based on the estimate of ρ in the LAPF. It is derived from the observation minus background $(o - b)$ statistics, the current ensemble spread, and the assumed observation error. Its determination is based on

$$\mathbf{d}_{o-b} = \mathbf{y}^o - H(\mathbf{x}^b) = \mathbf{y}^o - H(\mathbf{x}^t) + H(\mathbf{x}^t) - H(\mathbf{x}^b)$$
$$\approx \boldsymbol{\varepsilon}^o - \mathbf{H}\boldsymbol{\varepsilon}^b, \qquad (20)$$

with the true background state $\mathbf{x}^t$, the background state $\mathbf{x}^b$, the linearization $\mathbf{H}$ of H, the vector of observation errors $\boldsymbol{\varepsilon}^o$, and the vector of background errors $\boldsymbol{\varepsilon}^b$. Then, if the observation errors and background errors are uncorrelated, we obtain

$$\mathbb{E}[\mathbf{d}_{o-b}\mathbf{d}_{o-b}^{\mathrm{T}}] = \mathbb{E}[\boldsymbol{\varepsilon}^o(\boldsymbol{\varepsilon}^o)^{\mathrm{T}}] + \mathbf{H}\mathbb{E}[\boldsymbol{\varepsilon}^b(\boldsymbol{\varepsilon}^b)^{\mathrm{T}}]\mathbf{H}^{\mathrm{T}} \qquad (21)$$

(see Desroziers et al. 2005). To estimate the inflation factor, we substitute the expectation values of the background and observation errors with the actual ensemble covariance matrix $\mathbf{P}^b$ multiplied by the inflation factor ρ and the nominal covariance of observation error $\mathbf{R}$, respectively: $\mathbb{E}[\boldsymbol{\varepsilon}^b(\boldsymbol{\varepsilon}^b)^{\mathrm{T}}] \approx \rho\mathbf{P}^b$ and $\mathbb{E}[\boldsymbol{\varepsilon}^o(\boldsymbol{\varepsilon}^o)^{\mathrm{T}}] \approx \mathbf{R}$, resulting in

---

[5] Active observations are those that passed the quality control.

$$\mathbb{E}[\mathbf{d}_{o-b}\mathbf{d}_{o-b}^{\mathrm{T}}] \approx \mathbf{R} + \rho\mathbf{H}\mathbf{P}^b\mathbf{H}^{\mathrm{T}}. \quad (22)$$

Now, by taking the trace $\mathrm{Tr}(A) = \sum_{j=1}^{m} a_{jj}$ of the matrices on both sides, using $\mathrm{Tr}(A + \tilde{A}) = \mathrm{Tr}(A) + \mathrm{Tr}(\tilde{A})$, $\mathrm{Tr}(\rho\tilde{A}) = \rho\mathrm{Tr}(\tilde{A})$, and $\mathrm{Tr}(vv^{\mathrm{T}}) = \mathrm{Tr}(v^{\mathrm{T}}v)$, the inflation factor ρ is estimated by

$$\rho = \frac{\mathbb{E}[\mathbf{d}_{o-b}^{\mathrm{T}}\mathbf{d}_{o-b}] - \mathrm{Tr}(\mathbf{R})}{\mathrm{Tr}(\mathbf{H}\mathbf{P}^b\mathbf{H}^{\mathrm{T}})}, \quad (23)$$

following Desroziers et al. (2005) and Li et al. (2009).

Equation (23) computes a scalar inflation factor ρ based on a set of observations and the corresponding ensemble spread in observation space. It is carried out locally as a localized ensemble data assimilation method is employed; that is, we calculate

$$\rho(p) = \frac{\mathbf{d}_{o-b}^{\mathrm{T}}\mathbf{d}_{o-b} - r^2}{q^2} \quad (24)$$

at each point $p \in \mathscr{G}$ with the local innovation vector $\mathbf{d}_{o-b}$, the observation error $r^2 = \mathrm{Tr}(\mathbf{R})$, and the local estimate $q^2 := \mathrm{Tr}(\mathbf{H}\mathbf{P}^b\mathbf{H}^{\mathrm{T}})$ of the background error covariance in observation space. The factor $\rho(p)$ is the estimate for the local variance inflation at the analysis point p in the LETKF.

Because of the localization procedure, the number of observations used in this method may be small, and the estimated value of ρ may be based on limited statistics. To make the estimate more robust, we first limit ρ by lower and upper bounds of 0.9 and 1.5 and afterward perform a temporal smoothing. A weighting factor $\alpha = 0.05$ is chosen to combine the $\tilde{\rho}_k$ estimated by (24) in the current cycle k and the ρ_{k-1} used in the previous analysis cycle (3 h in the past) to get the ρ_k to be applied in the current cycle k:

$$\rho_k = \alpha\tilde{\rho}_k + (1 - \alpha)\rho_{k-1}, \qquad k = 1, 2, 3, \ldots. \quad (25)$$

In the LETKF, ρ is used at each analysis grid point to calculate the filter transformation matrix $\mathbf{W}_{\mathrm{infl}}$ by

$$\mathbf{W}_{\mathrm{infl}}(p) = \sqrt{\rho(p)}\mathbf{W}(p), \qquad p \in \mathscr{G}, \quad (26)$$

where $\mathbf{W}$ is the transform matrix defined in (6). For the localized adaptive particle filter, pure multiplicative inflation is not appropriate, since it would just inflate the distribution of the remaining duplicate ensemble members. Instead, we apply the Gaussian resampling based on (23) and (24), as described below.

f. Gaussian resampling or rejuvenation

The LAPF first calculates the ensemble weights according to the ensemble space projection (15) and the resampling (19) at each of the analysis points $p \in \mathscr{G}$. Usually, this resampling leads to part of the total number of particles only getting the majority of the weights. Often, a rejuvenation step [see Doucet et al. 2001; van Leeuwen et al. 2015, Eq. (2.39)] is carried out around each of the remaining particles (i.e., new particles are generated based on a pseudo-random draw in ensemble space). We note that this rejuvenation can be considered as classical resampling from some posterior represented by a superposition of Gaussian functions in the spirit of classical MCMC methods (Nakamura and Potthast 2015, chapters 4 and 5).

In our implementation of the LAPF, we draw from a Gaussian distribution around each remaining particle in ensemble space, used with appropriate multiplicity as constructed by $\breve{\mathbf{W}}$ in (19). Using a Gaussian with mean given by the column vector $\breve{\mathbf{W}}_\ell$ of $\breve{\mathbf{W}}$ and with a covariance matrix $\sigma^2\mathbf{I} \in \mathbb{R}^{L \times L}$, this leads to a draw from a distribution in physical space with mean given by the ensemble $\{x^{(\ell)}, \ell = 1, \ldots, L\}$ and the rescaled ensemble covariance matrix $\sigma^2\mathbf{P}^b$.

1) GLOBAL REJUVENATION

A pseudorandom matrix $\mathbf{N} \in \mathbb{R}^{L \times L}$ with each element draw from a Gaussian distribution with mean zero and covariance 1 is chosen globally to ensure the best possible continuity of the meteorological variables in physical space. Since resampling is global, without modulating the random matrix $\mathbf{N}$, the resulting perturbations would be scaled superpositions of the original ensemble members, keeping linear balances completely and nonlinear balances to some degree. However, adaptivity of the pseudorandom draws turns out to be crucial to avoid filter divergence and filter collapse. Only after adapting the spread of the rejuvenation or resampling carefully as follows did the filter start to be stable.

2) ADAPTIVITY

For the adaptive resampling step, the size of the draw (given by $\mathbf{N}$) is modulated by applying a scalar perturbation factor σ for each analysis grid point. Scaling of the draw around each member at time t_k is carried out by

$$\mathbf{W}(p) = \breve{\mathbf{W}}(p) + \mathbf{N} \times \sigma[\rho_k(p)]. \quad (27)$$

The specification of the factor σ is based on the inflation parameter ρ_k estimated in (25):
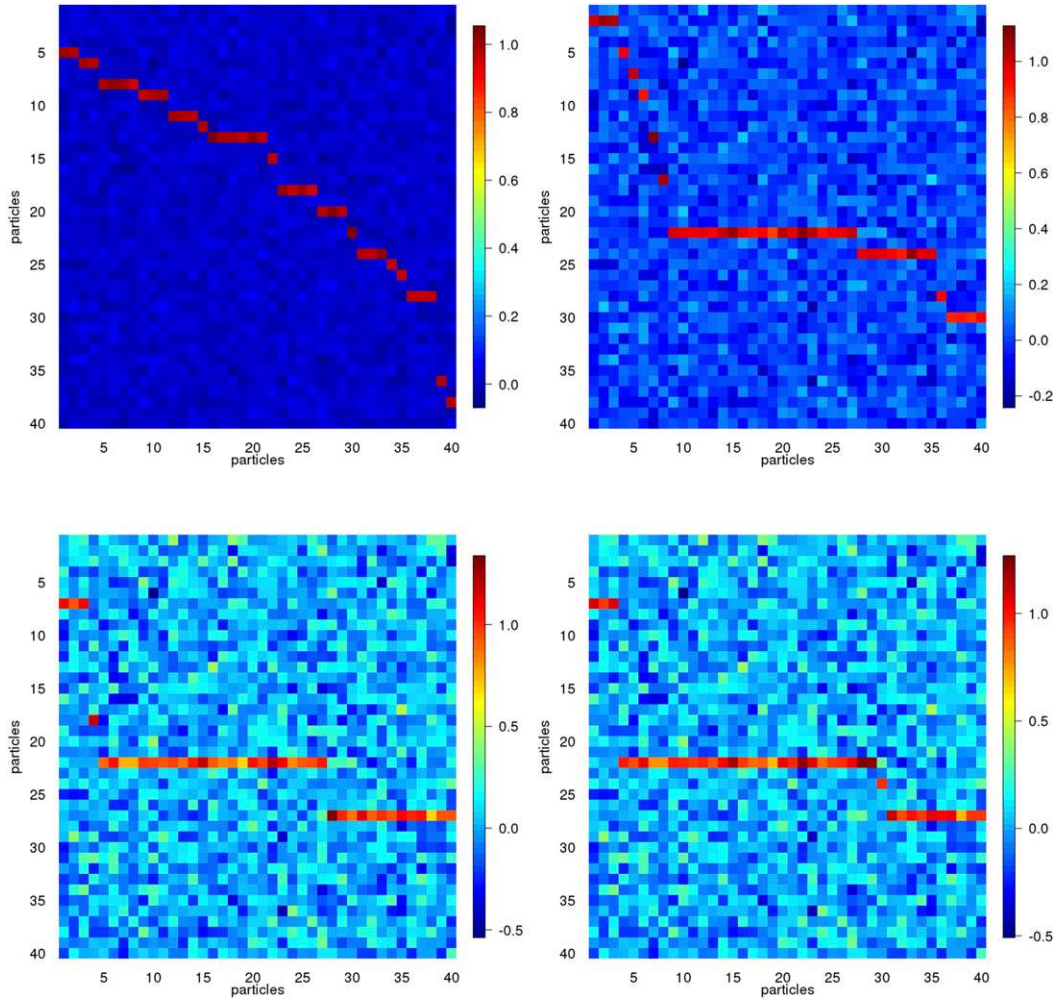
FIG. 2. Examples for transformation matrices **W** of the LAPF after Gaussian resampling. We show **W** at one analysis grid point [60°N, 90°E (Siberia)], at levels of (from left to right and top to bottom) 100, 500, 750, and 1000 hPa for 0000 UTC 26 May 2016. The x axis shows the analysis ensemble index, and the y axis shows the first-guess ensemble index.

$$\sigma(\rho) := \begin{cases} c_0, & \rho < \rho^{(0)}, \\[2ex] c_0 + (c_1 - c_0) \times \dfrac{\rho - \rho^{(0)}}{\rho^{(1)} - \rho^{(0)}}, & \rho^{(0)} \leq \rho \leq \rho^{(1)} \\[2ex] c_1, & \rho > \rho^{(1)}, \end{cases}$$

$$\text{(28)}$$

where elementary tuning tests lead to the values $c_0 = 0.02$, $c_1 = 0.2$, $\rho^{(0)} = 1.0$, and $\rho^{(1)} = 1.4$ (which might not be an optimal choice). The function continuously depends on ρ with $\sigma = c_0$ if $\rho \leqq \rho_0$ and $\sigma = c_1$ if $\rho \geqq \rho_1$.

We summarize that we always perturb each of the remaining members of the filter after the classical resampling step by a Gaussian with standard deviation

of at least c_0 and at most c_1. The scaling is a continuous function of the input parameters of our estimator of ρ (i.e., it depends continuously on space, on the observations, and on the ensemble members). The only discontinuities can occur when the matrix **W** in dependence of the spatial point p itself is discontinuous according to a change in the weights resulting from the classical resampling procedure.

3) NUMBER OF SURVIVING PARTICLES

We remark that the adaptive Gaussian rejuvenation ensures that with probability one, the analysis ensemble consists of L distinct particles. However, in the first classical resampling step, the number of distinct particles can be significantly smaller with the limiting case where only one particle gains all the weight, such that
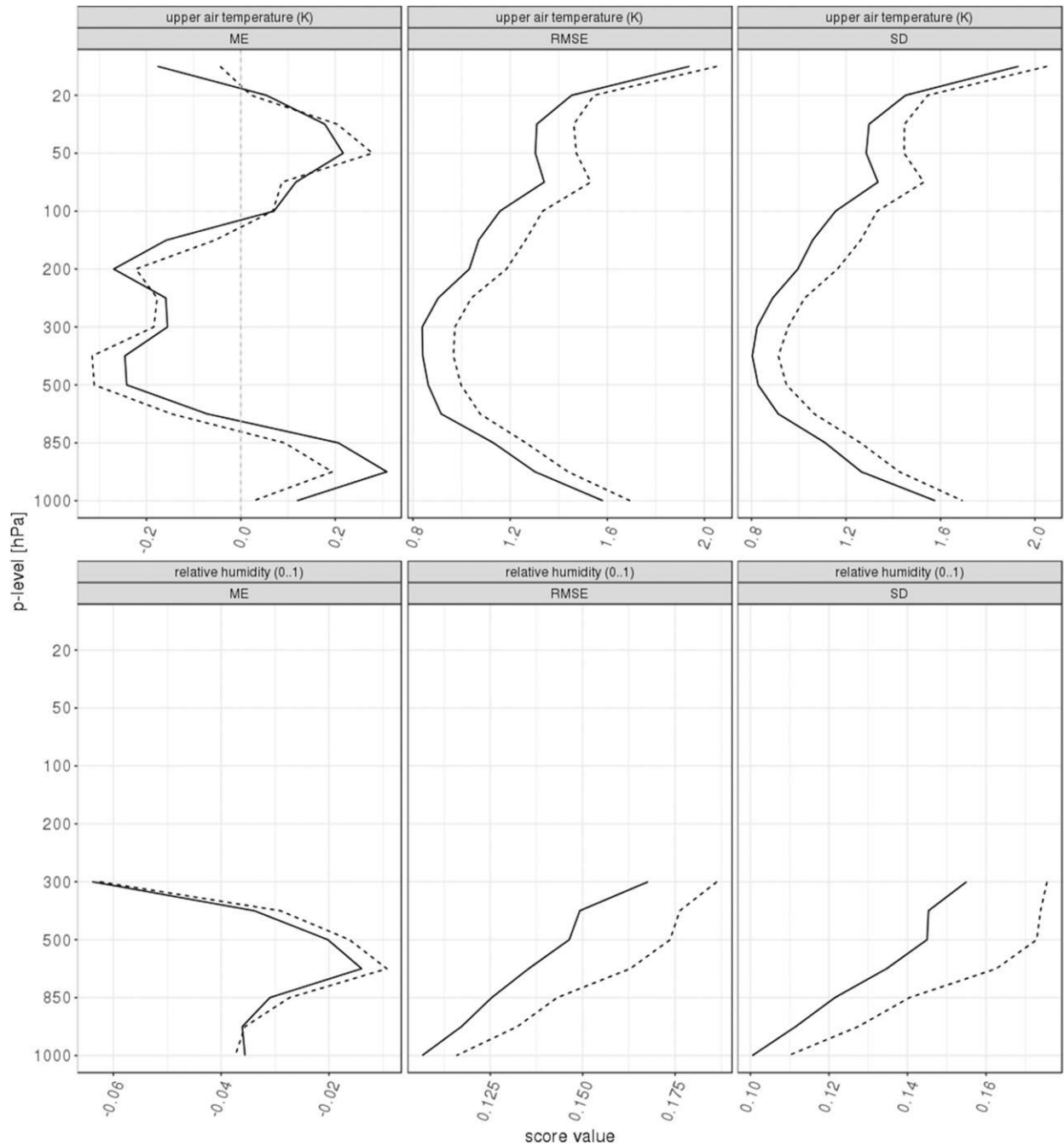
FIG. 3. Observation minus ensemble mean first-guess statistics in the 3-hourly assimilation cycle for the period 8–31 May 2016 in the global domain: (left) ME, (middle) RMSE, and (right) SD. (top) Upper-air temperature (K); (bottom) relative humidity (0, . . . , 1). Solid lines indicate the reference (LETKF), and dashed lines indicate the experiment (LAPF).

after resampling and before rejuvenation, we have L identical copies of this particle in a given localization point. It is, of course, highly interesting to study the statistics of how many particles remain at each analysis grid point and in what way the ensemble projection (11) helps the filter to stay away from collapse.

In Fig. 1, for three dates close to the end of the experimental period, we show some histograms of the number of particles at each analysis grid point surviving the first resampling step (i.e., those particles with weights $w_{k,l} \geq 1$ before rejuvenation when normalizing the total weight to L). The results show that at 100 hPa
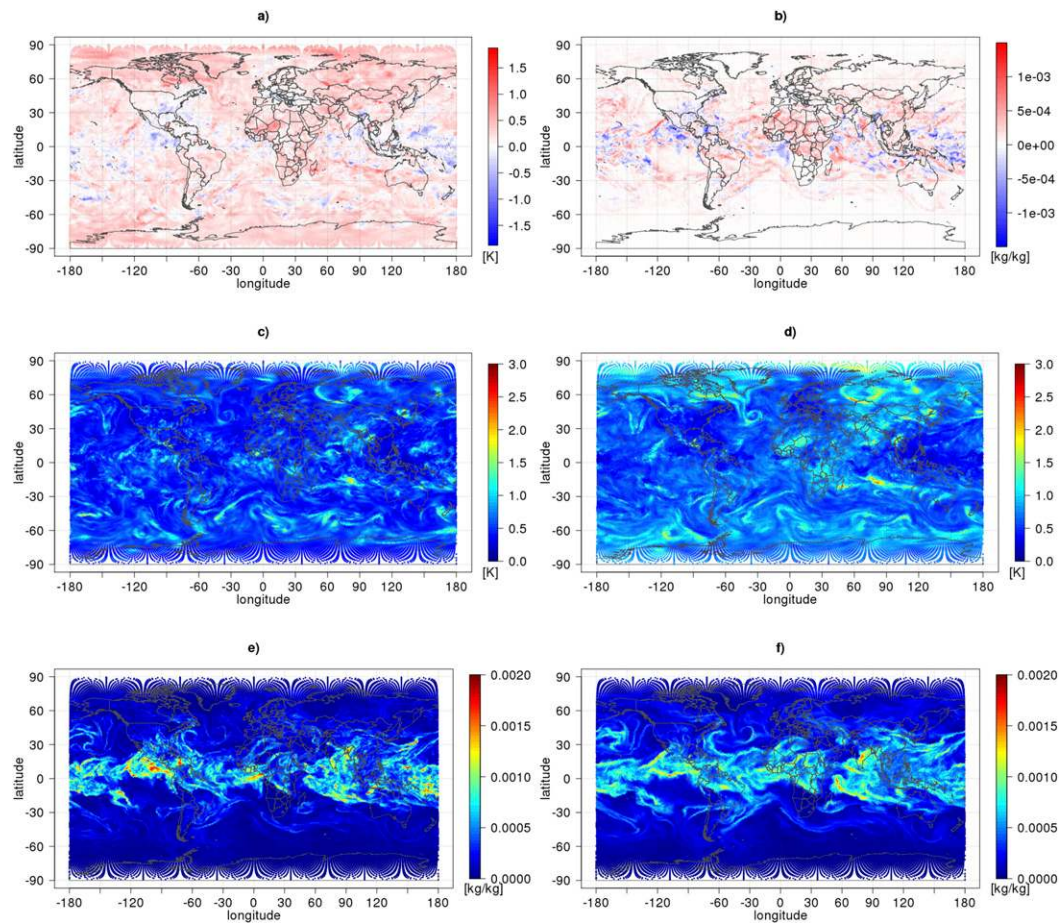
FIG. 4. Spread at model level 64 (~500 hPa) for 0000 UTC 31 May 2016. Differences for (a) upper-air temperature (K) and (b) specific humidity (kg kg^{-1}) between LETKF and LAPF. Fields for upper-air temperature in (c) LAPF and (d) LETKF. Fields for specific humidity in (e) LAPF and (f) LETKF.

(high level with few data; left column), mainly five to 20 particles obtain most of the weight. There are considerably fewer cases where 20 to 40 particles survive the classical resampling step.

In the midtroposphere (500 hPa; middle column) it can clearly be seen that the first date (0300 UTC 20 May 2016) and the third date (2100 UTC 31 May 2016) show a quite similar distribution. The number of cases with one up to 30 particles with weight larger than one is very similar. For the second date (second row; 1200 UTC 25 May 2016), there are only a few cases with more than 20 surviving particles. This is probably due to the larger amount of synoptic data at 1200 UTC, compared to 0300 and 2100 UTC.

The result for the bottom level (1000 hPa; right column) differs from the other two levels. In all these subfigures, there exists one peak at one to five particles. For the first and the third rows, there is also a peak at 30 particles. The first peak might be due to model biases in

the boundary layer in combination with a high number of observations.

A high number of observations leads to a small number of particles surviving the classical resampling, which is the well-known filter divergence phenomenon. We remark, however, that due to the ensemble transform and projection step of section 3b, this divergence only occurs in a part of the localization boxes. Further, adaptive Gaussian rejuvenation in ensemble space guarantees the calculation of L distinct analysis ensemble members with a controlled spread and distribution.

4) EXAMPLES

Examples for the matrix **W** from (27) are displayed in Fig. 2. The **W** matrices show how the analysis ensemble is constructed from the first-guess ensemble. Entries close to one indicate that an analysis ensemble member is sampled from the respective first-guess member. In general, some first-guess members lead to multiple
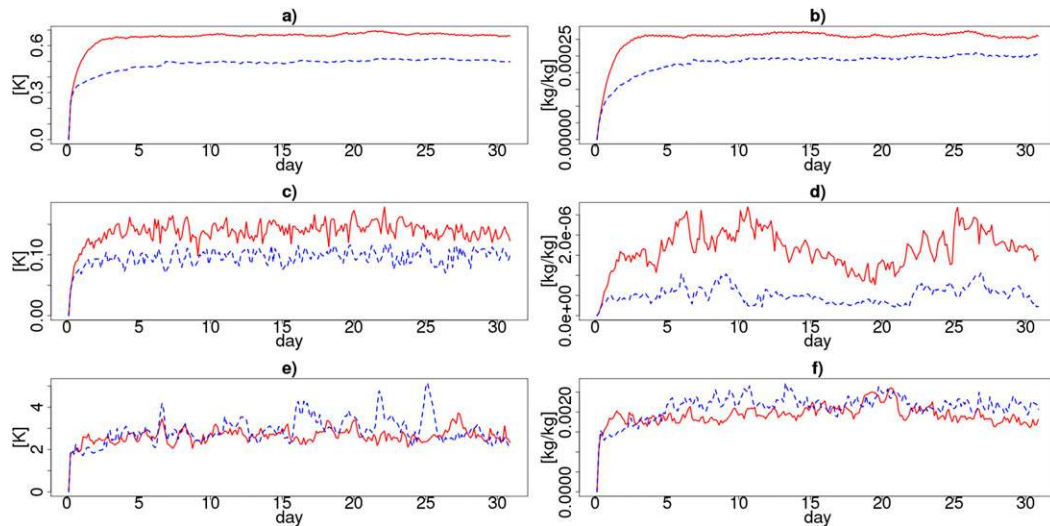
FIG. 5. Time series of the first-guess ensemble spread at ICON level 64 (~500 hPa) averaged globally. (left) Spread for upper-air temperature (K); (right) specific humidity (kg kg^{-1}). (top) Mean, (middle) minimum, and (bottom) maximum of the spread (red solid line = LETKF, blue dashed line = LAPF).

analysis members, whereas others are dismissed totally (only entries close to zero in that row). The deviations from zero or one are consequences of the Gaussian resampling step.

In the first case (top left; in the lower stratosphere with few observations), the analysis members are sampled from a large number of first-guess particles. Going farther to the ground, fewer first-guess particles are selected, but are replicated multiple times due to their large weight. This is caused by the larger number of observations farther down in the atmosphere putting more constraints on the prior distribution.

It may also be noted that in Fig. 2a, the strength of the Gaussian resampling is weak so that entries remain close to zero or one. In the subsequent figures, the estimated inflation factor is larger so that final weights almost fill the range in between −0.5 and 1.

4. Numerical tests in the operational framework and results

Here, we present tests of the localized adaptive particle filter for the global ICON model for a typical experimental setup and for a time period of 1 month. In particular, we investigate the assimilation cycle and forecast scores in some detail, diagnosing the development of spread, bias, scores, and the stability of the system.

We start with evaluating the assimilation cycle of the LAPF experiment in section 3a. As reference, we use the LETKF implementation that is operational at DWD. With the help of spread control, the LAPF provides reasonable results and is stable over the full experimental

period. In section 3b, we investigate the quality of forecast runs and compare it to the performance of the system in the operational setup.

a. Assimilation cycle

For the experimental diagnostics, a spinup period of 1 week is excluded from the observation minus first-guess statistics, as bias-correction algorithms and ensemble spread have to adapt to the new method.

We first investigate upper-air observation minus first-guess (obs − fg) statistics based on radiosonde observations (TEMP). Some selection of results is displayed in Fig. 3, in particular bias (ME), RMSE, and standard deviation (SD) for 3-h forecasts (first guesses), generated by the reference (LETKF) and the particle filter (LAPF). These statistics are based on observations that passed the quality control (i.e., that were actually used in the assimilation). Figure 3 visualizes statistics for the global domain and for the time period 8–31 May 2016.

The results show that the root-mean-square error and standard deviation of the current LAPF are about 10%–15% worse than those of the LETKF. They also show that the system is functioning and shows comparable features to the LETKF-based statistics. The LAPF shows better results for upper-air temperature than for relative humidity. For the bias, the results for relative humidity determined by the LAPF are slightly better than the results of the LETKF (in the sense that they are closer to zero). The values of RMSE and SD in the vertical column show similar shapes, but with higher values for LAPF. Overall, this first implementation of the LAPF scheme, where we were not yet able to carry
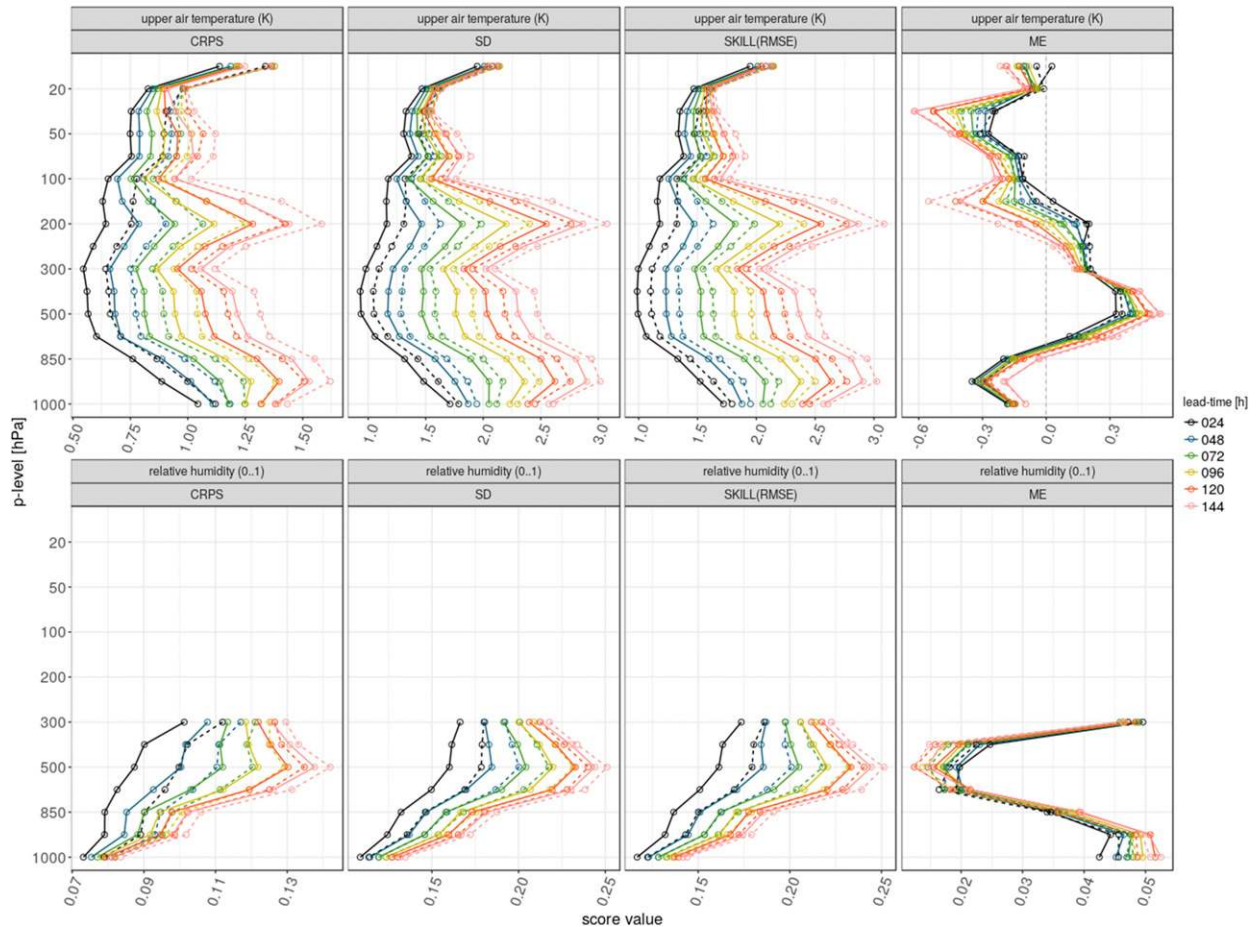
FIG. 6. The global statistics of forecasts against observations for the period of 2–24 May 2016. (left to right) CRPS, SD, RMSE, and ME. (top) Statistics for upper-air temperature (K); (bottom) relative humidity (0, . . . , 1). The solid line indicates the reference (LETKF), and the dashed line indicates the experiment (LAPF). Colors indicate different lead times (from 24 to 144 h).

out the time-consuming full tuning, which is usually done for an operational system, shows a very reasonable behavior in comparison to the LETKF.

Since filter collapse and filter divergence are of high interest, we next investigate the behavior of corresponding diagnostics, investigating spread behavior and the number of surviving particles in each resampling step. Figure 4 shows the spread averaged over all ensemble members for 0000 UTC 31 May 2016 (i.e., at the end of the first month of cycling). Displayed is the spread at ICON level 64 (approximately 500 hPa) for upper-air temperature and specific humidity. In this context, spread is the pointwise variance of the ICON-EPS.

- The left panels in the second and third rows show the fields for the LAPF; the right panels show those for the LETKF. In the first row, the differences of the spread between LETKF and LAPF are displayed for upper-air temperature in Fig. 4a and specific

humidity spread in Fig. 4b. It can clearly be seen that the structures are similar, but the spread for both variables is higher for the LETKF ($\overline{\sigma_T} = 0.6653$ K and $\overline{\sigma_q} = 0.0003$ kg kg^{-1}) than for the LAPF ($\overline{\sigma_T} = 0.5037$ K and $\overline{\sigma_q} = 0.0002$ kg kg^{-1}).

- The main structures of the spread of LAPF and LETKF show similar physical features. The differences between the two filters are more random and linked to the stochastic parts of the methods.

- The regions with the maximum spread are also the places where the LAPF and the LETKF differ most. One example of this is the temperature over Madagascar. Here, the spread of the LETKF is much larger than that of the LAPF. Also, the vortex over western Siberia shows a big difference between the two filters.

- For the specific humidity, the biggest differences are situated in the tropics, where humidity values are large. The spread difference plot clearly resembles the patterns of the spread in specific humidity itself.
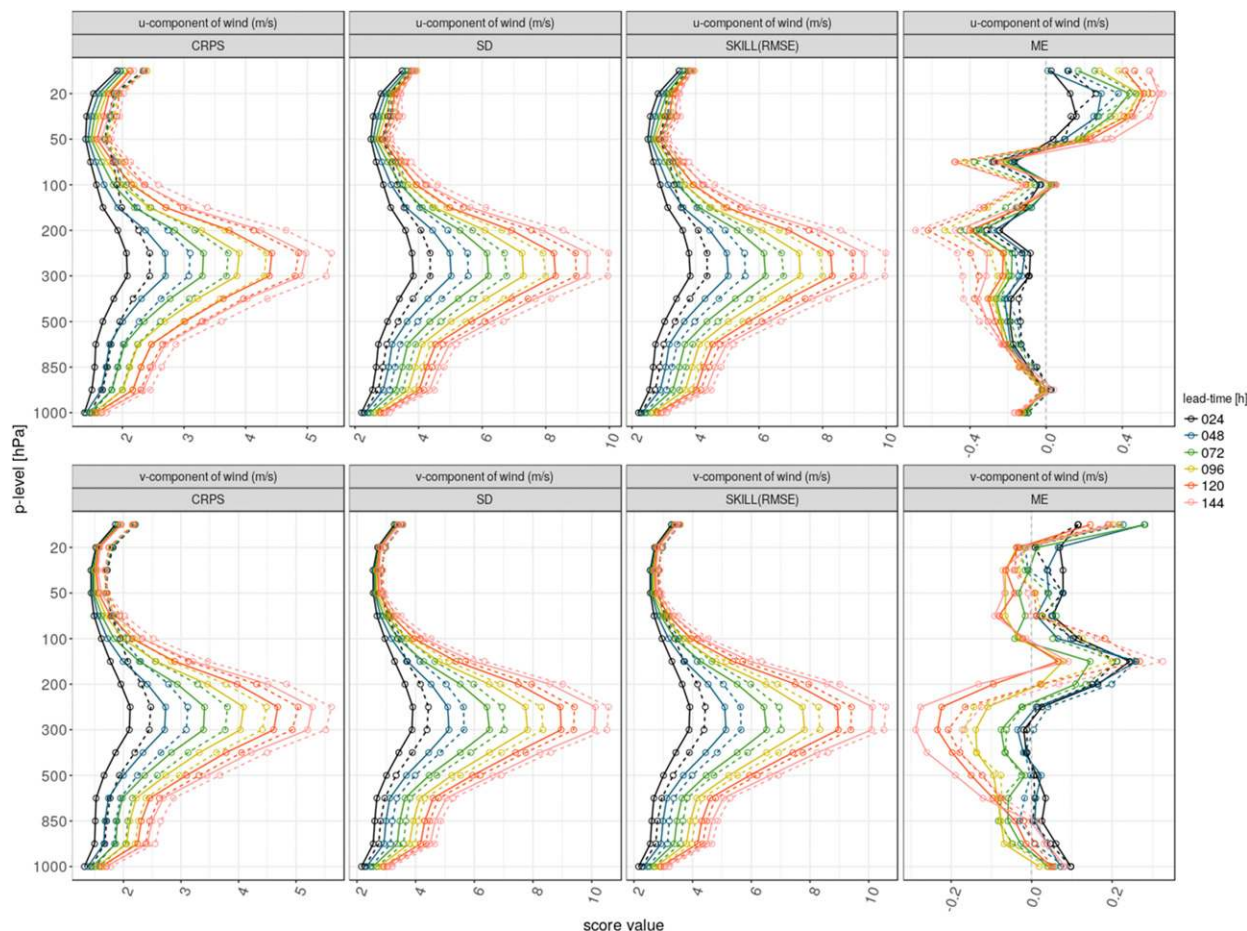
FIG. 7. As in Fig. 6, but for (top) u component and (bottom) v component of wind m s^{-1}.

It indicates that the differences between LAPF and LETKF are often situated at the borders of the big vortices. The largest differences are situated, for example, in the region around the west coast of Mexico, where the LAPF has a maximum but the LETKF does not.

For studying the development of the spread and the stability of the filter, time series from 1 to 31 May 2016 for both filters are plotted in Fig. 5. The spread starts from zero and needs a couple of days to settle, because we started all members from the same state duplicated $L = 40$ times. The first row shows the mean of the spread for upper-air temperature and specific humidity, calculated at each point in time and for one horizontal level of the atmospheric grid. For both quantities, the LAPF shows lower values for mean spread than the LETKF. The same holds for the minimum value of the spread. However, the maximum values of the spread of the LETKF and the LAPF show quite similar values.

b. Forecast verification

Forecasts were run twice a day at 0000 and 1200 UTC. In Figs. 6 and 7, a verification of temperature, relative humidity, and wind components compared to radiosonde observations is shown: continuous ranked probability score (CRPS), SD, RMSE, and ME for lead times of 24, 48, 72, 96, 120, and 144 h for the forecasts based on the LETKF and the LAPF analyses. The determination of SD, RMSE, and ME is based on the ensemble mean. At this stage, the results show that in the current development stage, the LAPF does not outperform the LETKF. However, the shapes of CRPS, SD, and RMSE are comparable, indicating that the LAPF is stable and has a reasonable behavior. The humidity bias of the LAPF is reduced in comparison to the LETKF at all heights above 850 hPa. It is consistent with theory and implementation that the particle filter does not draw the model fields to the observations as strongly as the LETKF, a distance that is maintained throughout all forecast lead times and explains the behavior of the scores.

5. Conclusions

Standard algorithms for data assimilation used for large-scale atmospheric analysis in operational centers include the ensemble Kalman filter and 4D-VAR. These data assimilation methods are either inherently or practically based on the assumption that the underlying distribution is Gaussian. If the ensemble distribution is not Gaussian, these methods are not optimal. In the case of non-Gaussianity, more general Bayesian methods such as the particle filter have been proposed. The core idea of the particle filter is to realize the Bayesian approach giving a weight to each particle, depending on its distance to the observations. The adaptation of particles is carried out in different ways: by resampling, nudging particles toward some proposal distribution, or optimal transport processes.

Classical particle filters in high-dimensional dynamical systems suffer from filter divergence or filter collapse due to the curse of dimensionality. In this work, we have developed and implemented a localized adaptive particle filter (LAPF) in ensemble space with spread control and Gaussian resampling or rejuvenation. With the help of modulated rejuvenation, we prevent the filter divergence as well as filter collapse. It has been implemented for global atmospheric data assimilation to fit into the framework of the global operational weather prediction model ICON of DWD.

The LAPF was tested over a period of 1 month with 40 ensemble members, a global horizontal resolution of 52 km, and 90 vertical layers in an operational setup with slightly reduced resolution. A comparison of the scores with those of the operational system of DWD (with some modest reduction of resolution) shown in section 4 demonstrates that the localized adaptive particle filter is able to provide reasonable atmospheric analysis in a large-scale environment. We have shown that for this first attempt, the RMSE quality of forecasts based on the LAPF is 10%–15% behind the forecast quality of the LETKF for forecasts up to several days (cf. Figs. 6 and 7), and BIAS is partly improved, in particular for humidity. Altogether, for the assimilation cycle and forecasts, the LAPF shows promising results (Fig. 3). Furthermore, we are able to demonstrate the stability of the LAPF over a period of 1 month (cf. Figs. 4 and 5) and show that atmospheric data assimilation within an operational modeling environment is possible based on a localized adaptive particle filter approach.

REFERENCES

Ades, M., and P. J. van Leeuwen, 2013: An exploration of the equivalent weights particle filter. *Quart. J. Roy. Meteor. Soc.*, **139**, 820–840, https://doi.org/10.1002/qj.1995.

——, and ——, 2015: The equivalent-weights particle filter in a high-dimensional system. *Quart. J. Roy. Meteor. Soc.*, **141**, 484–503, https://doi.org/10.1002/qj.2370.

Anderson, B. D. O., and J. B. Moore, 2012: *Optimal Filtering.* Courier Corporation, 368 pp.

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903, https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

——, 2007a: An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus*, **59A**, 210–224, https://doi.org/10.1111/j.1600-0870.2006.00216.x.

——, 2007b: Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D*, **230**, 99–111, https://doi.org/10.1016/j.physd.2006.02.011.

——, 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.

Bain, A., and D. Crisan, 2009: *Fundamentals of Stochastic Filtering.* Stochastic Modelling and Applied Probability Series, Vol. 60, Springer, 390 pp., https://doi.org/10.1007/978-0-387-76896-0.

Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55, https://doi.org/10.1038/nature14956.

Bengtsson, T., C. Snyder, and D. Nychka, 2003: Toward a nonlinear ensemble filter for high-dimensional systems. *J. Geophys. Res.*, **108**, 8775, https://doi.org/10.1029/2002JD002900.

Bickel, P., B. Li, and T. Bengtsson, 2008: Sharp failure rates for the bootstrap particle filter in high dimensions. *Pushing the Limits of Contemporary Statistics: Contributions in Honor of Jayanta K. Ghosh*, Vol. 3, Institute of Mathematical Statistics, 318–329, https://doi.org/10.1214/074921708000000228.

Bishop, C. H., 2016: The GIGG–EnKF: Ensemble Kalman filtering for highly skewed non-negative uncertainty distributions. *Quart. J. Roy. Meteor. Soc.*, **142**, 1395–1412, https://doi.org/10.1002/qj.2742.

——, and D. Hodyss, 2007: Flow-adaptive moderation of spurious ensemble correlations and its use in ensemble-based data assimilation. *Quart. J. Roy. Meteor. Soc.*, **133**, 2029–2044, https://doi.org/10.1002/qj.169.

——, and ——, 2009a: Ensemble covariances adaptively localized with ECO-RAP. Part 1: Tests on simple error models. *Tellus*, **61A**, 84–96, https://doi.org/10.1111/j.1600-0870.2008.00371.x.

——, and ——, 2009b: Ensemble covariances adaptively localized with ECO-RAP. Part 2: A strategy for the atmosphere. *Tellus*, **61A**, 97–111, https://doi.org/10.1111/j.1600-0870.2008.00372.x.

——, B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129**, 420–436, https://doi.org/10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2.

Bjerknes, V., 1904: Das problem der wettervorhersage, betrachtet von standpunkt der mechanik und der physik. *Meteor. Z.*, **21**, 1–7.

Bloom, S. C., L. L. Takacs, A. M. da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.*, **124**, 1256–1271, https://doi.org/10.1175/1520-0493(1996)124<1256:DAUIAU>2.0.CO;2.

Bouallègue, Z. B., S. E. Theis, and C. Gebhardt, 2013: Enhancing COSMO-DE ensemble forecasts by inexpensive techniques. *Meteor. Z.*, **22**, 49–59, https://doi.org/10.1127/0941-2948/2013/0374.

Brusdal, K., J. M. Brankart, G. Halberstadt, G. Evensen, P. Brasseur, P. J. van Leeuwen, E. Dombrowsky, and J. Verron, 2003: A demonstration of ensemble-based assimilation methods with a layered OGCM from the perspective of operational ocean forecasting systems. *J. Mar. Syst.*, **40–41**, 253–289, https://doi.org/10.1016/S0924-7963(03)00021-6.

Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126**, 1719–1724, https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.

Campbell, W. F., C. H. Bishop, and D. Hodyss, 2010: Vertical co-variance localization for satellite radiances in ensemble Kalman filters. *Mon. Wea. Rev.*, **138**, 282–290, https://doi.org/10.1175/2009MWR3017.1.

Crisan, D., and B. Rozovskii, 2011: *The Oxford Handbook of Nonlinear Filtering*. Oxford University Press, 1080 pp.

Daley, R., 1993: *Atmospheric Data Analysis*. Cambridge University Press, 472 pp.

Desroziers, G., L. Berre, B. Chapnik, and P. Poli, 2005: Diagnosis of observation, background and analysis-error statistics in observation space. *Quart. J. Roy. Meteor. Soc.*, **131**, 3385–3396, https://doi.org/10.1256/qj.05.108.

Doucet, A., N. de Freitas, and N. Gordon, 2001: *Sequential Monte Carlo Methods in Practice*. Springer, 582 pp., https://doi.org/10.1007/978-1-4757-3437-9.

ECMWF, 2014: IFS documentation—Cy40r1. Part I: Observations. ECMWF Tech. Doc., 76 pp., https://www.ecmwf.int/sites/default/files/IFS_CY40R1_Part1.pdf.

Elsheikh, A. H., I. Hoteit, and M. F. Wheeler, 2014: A nested sampling particle filter for nonlinear data assimilation. *Quart. J. Roy. Meteor. Soc.*, **140**, 1640–1653, https://doi.org/10.1002/qj.2245.

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.

——, 2009: *Data Assimilation: The Ensemble Kalman Filter*. Springer, 280 pp.

——, and P. J. van Leeuwen, 2000: An ensemble Kalman smoother for nonlinear dynamics. *Mon. Wea. Rev.*, **128**, 1852–1867, https://doi.org/10.1175/1520-0493(2000)128<1852:AEKSFN>2.0.CO;2.

Frei, M., and H. Künsch, 2013: Bridging the ensemble Kalman and particle filters. *Biometrika*, **100**, 781–800, https://doi.org/10.1093/biomet/ast020.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757, https://doi.org/10.1002/qj.49712555417.

Gassmann, A., and H.-J. Herzog, 2008: Towards a consistent numerical compressible non-hydrostatic model using generalized Hamiltonian tools. *Quart. J. Roy. Meteor. Soc.*, **134**, 1597–1613, https://doi.org/10.1002/qj.297.

Gordon, N. J., D. J. Salmond, and A. F. M. Smith, 1993: Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proc.*, **140F**, 107–113, https://doi.org/10.1049/ip-f-2.1993.0015.

Greybush, S. J., E. Kalnay, T. Miyoshi, K. Ide, and B. R. Hunt, 2011: Balance and ensemble Kalman filter localization techniques. *Mon. Wea. Rev.*, **139**, 511–522, https://doi.org/10.1175/2010MWR3328.1.

Hoteit, I., D.-T. Pham, G. Triantafyllou, and G. Korres, 2008: A new approximate solution of the optimal nonlinear filter for data assimilation in meteorology and oceanography. *Mon. Wea. Rev.*, **136**, 317–334, https://doi.org/10.1175/2007MWR1927.1.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811, https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.

——, and ——, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137, https://doi.org/10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2.

——, and ——, 2005: Ensemble Kalman filtering. *Quart. J. Roy. Meteor. Soc.*, **131**, 3269–3289, https://doi.org/10.1256/qj.05.135.

——, ——, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.*, **133**, 604–620, https://doi.org/10.1175/MWR-2864.1.

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.

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

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 341 pp.

Li, H., E. Kalnay, and T. Miyoshi, 2009: Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter. *Quart. J. Roy. Meteor. Soc.*, **135**, 523–533, https://doi.org/10.1002/qj.371.

Miyoshi, T., 2011: The Gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform Kalman filter. *Mon. Wea. Rev.*, **139**, 1519–1535, https://doi.org/10.1175/2010MWR3570.1.

——, and Y. Sato, 2007: Assimilating satellite radiances with a local ensemble transform Kalman filter (LETKF) applied to the JMA global model (GSM). *SOLA*, **3**, 37–40, https://doi.org/10.2151/sola.2007-010.

——, and K. Kondo, 2013: A multi-scale localization approach to an ensemble Kalman filter. *SOLA*, **9**, 170–173, https://doi.org/10.2151/sola.2013-038.

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

——, K. Kondo, and T. Imamura, 2014: The 10,240-member ensemble Kalman filtering with an intermediate AGCM. *Geophys. Res. Lett.*, **41**, 5264–5271, https://doi.org/10.1002/2014GL060863.

Nakamura, G., and R. Potthast, 2015: *Inverse Modeling*. IOP Publishing, 509 pp., https://doi.org/10.1088/978-0-7503-1218-9.

Ott, E., and Coauthors, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, **56A**, 415–428, https://doi.org/10.3402/tellusa.v56i5.14462.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763, https://doi.org/10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2.

Periáñez, A., H. Reich, and R. Potthast, 2014: Optimal localization for ensemble Kalman filter systems. *J. Meteor. Soc. Japan*, **92**, 585–597, https://doi.org/10.2151/jmsj.2014-605.

Pham, D. T., J. Verron, and M. C. Roubaud, 1998: A singular evolutive extended Kalman filter for data assimilation in oceanography. *J. Mar. Syst.*, **16**, 323–340, https://doi.org/10.1016/S0924-7963(97)00109-7.

Poterjoy, J., 2016: A localized particle filter for high-dimensional nonlinear systems. *Mon. Wea. Rev.*, **144**, 59–76, https://doi.org/10.1175/MWR-D-15-0163.1.

——, and J. L. Anderson, 2016: Efficient assimilation of simulated observations in a high-dimensional geophysical system using a localized particle filter. *Mon. Wea. Rev.*, **144**, 2007–2020, https://doi.org/10.1175/MWR-D-15-0322.1.

——, R. A. Sobash, and J. L. Anderson, 2017: Convective-scale data assimilation for the Weather Research and Forecasting

Model using the local particle filter. *Mon. Wea. Rev.*, **145**, 1897–1918, https://doi.org/10.1175/MWR-D-16-0298.1.

Reich, S., and C. Cotter, 2015: *Probabilistic Forecasting and Bayesian Data Assimilation.* Cambridge University Press, 297 pp., https://doi.org/10.1017/CBO9781107706804.

Reinert, D., F. Prill, H. Frank, M. Denhard, and G. Zängl, 2018: Database reference manual for ICON and ICON-EPS. Deutscher Wetterdienst Tech. Rep., 90 pp., https://doi.org/10.5676/DWD_pub/nwv/icon_1.2.5.

Richardson, L. F., 1922: *Weather Prediction by Numerical Process.* Cambridge University Press, 258 pp.

Robert, S., D. Leuenberger, and H. R. Künsch, 2018: A local ensemble transform Kalman particle filter for convective-scale data assimilation. *Quart. J. Roy. Meteor. Soc.*, **144**, 1279–1296, https://doi.org/10.1002/qj.3116.

Schraff, C., H. Reich, A. Rhodin, A. Schomburg, K. Stephan, A. Periáñez, 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.

Snyder, C., T. Bengtsson, P. Bickel, and J. Anderson, 2008: Obstacles to high-dimensional particle filtering. *Mon. Wea. Rev.*, **136**, 4629–4640, https://doi.org/10.1175/2008MWR2529.1.

——, ——, and M. Morzfeld, 2015: Performance bounds for particle filters using the optimal proposal. *Mon. Wea. Rev.*, **143**, 4750–4761, https://doi.org/10.1175/MWR-D-15-0144.1.

Stordal, A. S., H. A. Karlsen, G. Naevdal, H. J. Skaug, and B. Vallès, 2011: Bridging the ensemble Kalman filter and particle filters: The adaptive Gaussian mixture filter. *Comput. Geosci.*, **15**, 293–305, https://doi.org/10.1007/s10596-010-9207-1.

van Leeuwen, P. J., 2003a: Nonlinear ensemble data assimilation for the ocean. *Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean*, Shinfield Park, Reading, United Kingdom, ECMWF, 265–286, https://www.ecmwf.int/en/elibrary/13249-nonlinear-ensemble-data-assimilation-ocean.

——, 2003b: A variance-minimizing filter for large-scale applications. *Mon. Wea. Rev.*, **131**, 2071–2084, https://doi.org/10.1175/1520-0493(2003)131<2071:AVFFLA>2.0.CO;2.

——, 2009: Particle filtering in geophysical systems. *Mon. Wea. Rev.*, **137**, 4089–4114, https://doi.org/10.1175/2009MWR2835.1.

——, 2010: Nonlinear data assimilation in geosciences: An extremely efficient particle filter. *Quart. J. Roy. Meteor. Soc.*, **136**, 1991–1999, https://doi.org/10.1002/qj.699.

——, Y. Cheng, and S. Reich, 2015: *Nonlinear Data Assimilation.* Frontiers in Applied Dynamical Systems: Reviews and Tutorials Series, Vol. 2, Springer, 118 pp., https://doi.org/10.1007/978-3-319-18347-3.

Vetra-Carvalho, S., P. J. van Leeuwen, L. Nerger, A. Barth, M. U. Altaf, P. Brasseur, P. Kirchgessner, and J.-M. Beckers, 2018: State-of-the-art stochastic data assimilation methods for high-dimensional non-Gaussian problems. *Tellus*, **70A**, 1445364, https://doi.org/10.1080/16000870.2018.1445364.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.

——, and ——, 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.

Yang, S.-C., E. Kalnay, B. Hunt, and N. E. Bowler, 2009: Weight interpolation for efficient data assimilation with the local ensemble transform Kalman filter. *Quart. J. Roy. Meteor. Soc.*, **135**, 251–262, https://doi.org/10.1002/qj.353.

Zängl, G., D. Reinert, P. Rípodas, 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.

Zhu, M., P. J. van Leeuwen, and J. Amezcua, 2016: Implicit equal-weights particle filter. *Quart. J. Roy. Meteor. Soc.*, **142**, 1904–1919, https://doi.org/10.1002/qj.2784.