Improving initial condition perturbations in a convection-permitting ensemble prediction system

Endi Keresturi1 | Yong Wang2 | Florian Meier2 | Florian Weidle2 | Christoph Wittmann2 | Aitor Atencia2

1NWP Section, Meteorological and Hydrological Service, Zagreb, Croatia
2Department of Forecasting Models, Central Institute for Meteorology and Geodynamics, Vienna, Austria

Correspondence
Endi Keresturi, Croatian Meteorological and Hydrological Service, Grič 3, 10000 Zagreb, Croatia.
Email: endi.keresturi@cirrus.dhz.hr

One of the main challenges presented by a limited-area model ensemble prediction system (LAMEPS) concerns the limited capacity for its initial condition perturbations to correctly represent large-scale flow uncertainties due to its limited-size domain and deficiencies in formulating lateral boundary conditions. In addition, a mismatch between LAMEPS (initial condition) and host EPS lateral boundary perturbations can form spurious gravity waves at the boundaries. In the present work, an ensemble $J_k$ blending method is proposed for improving representation of large-scale uncertainties and for addressing consistent initial conditions and lateral boundary perturbations. Our approach involves employing $J_k$ blending within a framework of three-dimensional variation (3D-Var) ensemble data assimilation (EDA). In such a system, small-scale perturbations are generated from 3D-Var EDA, while large-scale perturbations are generated from the host ensemble via $J_k$ blending. The ensemble $J_k$ method is implemented in the C-LAEF (Convection-permitting Limited-Area Ensemble Forecasting) system and is compared to the standard perturbed-observation EDA approach, i.e. perturbed-observation EDA without large-scale constraint. The comparison shows that the ensemble $J_k$ method gives a more skilful and reliable EPS, especially for upper-air variables. In addition, positive effects on the surface pressure and precipitation of large-scale perturbations are shown. Finally, the ensemble $J_k$ method’s capacity to alleviate perturbation mismatches is demonstrated.

KEYWORDS
blending, ensemble data assimilation, initial condition perturbations, $J_k$, limited-area ensemble

1 | INTRODUCTION

Since the very beginning of the development of numerical weather prediction (NWP), the research community has been aware of its shortcomings (Charney, 1951) and, most notably, of various sources of error (Lorenz, 1969; Kalnay, 2003; Buizza et al., 2005; Leutbecher and Palmer, 2008). These shortcomings became even more apparent after Lorenz demonstrated that numerical models exhibit chaotic behaviour (Lorenz, 1963). The chaotic nature and ever-present errors of the initial state necessarily lead to a complete loss of forecasting skill after a finite period of time, which Lorenz estimated to be 2 weeks (Lorenz, 1965). In the following decades, techniques of ensemble forecasting have been developed to address inherent uncertainties of the NWP system (Ehrendorfer, 1994; 1997; Lewis, 2005).

Different meteorological centres around the world have developed ways to construct operational ensemble prediction systems.
systems (EPSs). At first, only uncertainties in initial conditions (ICs) were considered. The first two methods used to generate IC perturbations operationally were the breeding method (Toth and Kalnay, 1993; 1997), developed by the National Centers for Environmental Prediction (NCEP), and singular vectors (Buizza and Palmer, 1995; Molteni et al., 1996), developed by the European Centre for Medium-range Weather Forecasts (ECMWF). These methods, although different, are based on the same main principle – perturbations should be added only in directions of phase space in which errors amplify the most. Later, additional methods were introduced (e.g. the Ensemble Transform (ET: Wei et al., 2008), Ensemble Transform Kalman Filter (ETKF: Wang and Bishop, 2003) and Ensemble Kalman Filter (EnKF: Evensen, 2003; Houtekamer and Mitchell, 2005). These methods are advantageous in that the perturbations that they produce are more consistent with the initial (analysis) probability distribution function (PDF) (Wei et al., 2008). All of the above methods are similar in that they only focus on IC perturbations. However, over the years it has been recognized that, to develop an optimal EPS, all sources of uncertainty must be perturbed (e.g. Houtekamer et al., 1996; Buizza et al., 2005; Vié et al., 2011; Wang et al., 2011; Bouttier et al., 2012; Nuissier et al., 2012; Romine et al., 2014), including IC, model and boundary condition errors.

In recent years, limited-area model ensemble prediction systems (LAMEPSs) have been developed to benefit from ensemble approach applied at the mesoscale and convective scale (Xue et al., 2007; Bowler et al., 2008; Clark et al., 2009; Vié et al., 2011; Wang et al., 2011; 2012; Peralta et al., 2012; Schellander-Gorgas et al., 2017; among others). Such ensembles are of particular importance due to a rapid loss of predictability observed at these scales (Hohenegger and Schär, 2007; Zhang et al., 2007; Judt, 2018). However, when using a limited-area model (LAM), one must address a new source of uncertainty – the specification of lateral boundary conditions (LBCs). This is typically done by nesting each LAMEPS member to a different host1 EPS member. Saito et al. (2011) discusses the importance of LBC perturbation in LAMEPSs. However, applying this procedure can lead to the following issue: when the method used to generate perturbations in a host EPS is different and independent from the method used in an LAMEPS, we observe a conflict between such perturbations at the lateral boundaries. This problem has been recognized by many authors (e.g. Bowler and Mylne, 2009; Brousseau et al., 2011; Wang et al., 2011; 2014; Caron, 2013; Kühnlein et al., 2014). Caron (2013) showed that correlations between the host EPS and LAMEPS tend to decrease with time and that the amplitude of LAMEPS perturbations can significantly differ from the amplitudes of the host model perturbations. As a consequence of this conflict, spurious waves are generated at the boundaries and quickly spread to the rest of the domain, resulting in excessive surface pressure spread in an LAMEPS. Kühnlein et al. (2014) demonstrated improved precipitation forecasts when IC perturbations were more consistent with LBC perturbations.

In addition to the above issue, as discussed in Guidard and Fischer (2008; GF08 hereafter) another challenge of limited-area modelling concerns the fact that LAMs are less effective at representing large-scale (e.g. synoptic) flow than global models due to their typically small domains, LBC formulation limitations, less effective LAM data assimilation (a fraction of the available observations are assimilated; see Gustafsson et al., 2018), etc. This problem can partially be addressed by applying a blending technique (Brožková et al., 2001; Yang, 2005; Wang et al., 2014; Hsiao et al., 2015). Different authors have developed various approaches to blending problem. Brožková et al. (2001) and Wang et al. (2014) proposed the use of a digital filter (DF) in the LAM and host model fields to obtain a blended analysis. This approach can be further extended to involve variational data assimilation where blending is performed before or after variational data assimilation.

Caron (2013) and Wang et al. (2014) showed that blending techniques (by DF) can also be used to address conflicting perturbations. Caron (2013) used the so-called selective ETKF method, where ETKF acts only at small scales, while large scales are kept from the host model, thus resulting in blending. Wang et al. (2014) used the DF-blending method under the framework of their ALADIN-LAEF (Aire Limitée Adaptation Dynamique développement International–Limited-Area Ensemble Forecasting) system (Wang et al., 2011) to combine IC perturbations generated through the breeding method with the global ECMWF Ensemble Prediction System (ECMWF-EPS) IC perturbations. DF-blending was used to blend small-scale ALADIN-LAEF and large-scale ECMWF-EPS IC perturbations to simultaneously introduce them into perturbed ICs. Wang et al. (2014) also achieved better upper-air scores by applying the blending method rather than the breeding method without blending.

However, as GF08 discusses, the DF-blending approach presents unique problems. Most notably, it has no relation to data assimilation theory, and there have been concerns regarding its optimality in the sense of minimum variance estimate. Thus, they proposed a method which includes global model information directly into a limited-area variational analysis (section 2.1) and implemented it in the ALADIN/Action de Recherche Petite Echelle et Grand Echelle (ARPEGE) framework within ALADIN three-dimensional variational assimilation (3D-Var). We refer to this as the Jk blending method. Dahlgren and Gustafsson (2012) explored this method further within a High-Resolution Limited-Area Model (HIRLAM)/ECMWF framework of the HIRLAM 3D-Var and 4D-Var. The method was found to

---

1In this article, the terms host model and host EPS denote models or EPSs that provide LBCs and that are also used when applying the Jk method (section 2).
improve on forecast scores in both the former and latter cases.

Large-scale information can also be included through the use of dual-resolution “hybrid” variational-ensemble data assimilation methods (e.g. Schwartz, 2016; Hu et al., 2017; Wu et al., 2017). Such methods incorporate ensemble-derived, flow-dependent background error covariances drawn from a coarser-resolution model into a high-resolution deterministic analysis. As opposed to the Jk method, which includes large-scale information as a new member in the 3D-Var (section 2.1), hybrid variational-ensemble data assimilation methods execute this through a background error covariance matrix. Schwartz (2016) demonstrated improved forecasts started from a hybrid analysis.

Issues of large-scale representation also reflect LAMEPSs in the sense that their IC perturbations, which are produced through techniques such as ensemble data assimilation (EDA), cannot correctly account for large-scale uncertainties. The simultaneous presence of small- and large-scale perturbations is beneficial for EPS (Wang et al., 2014; Johnson and Wang, 2016; Raynaud and Bouttier, 2016). To alleviate this problem, at Zentralanstalt für Meteorologie und Geodynamik (ZAMG), the LAMEPS ALADIN-LAEF has been developed within the framework of RC-LACE (Wang et al., 2018) and has been used in operations since 2009. It is based on the ALADIN model (ALADIN International Team, 1997; Termonia et al., 2018) and is currently run at 11 km horizontal grid spacing. However, with ever-increasing levels of computer power, it is now possible to run LAMEPSs at convection-permitting resolutions corresponding to a grid spacing of 1–4 km (e.g. Vié et al., 2011; Bouttier et al., 2012; Nuissier et al., 2012; Kühnlein et al., 2014; Romine et al., 2014; Schwartz et al., 2015; Johnson and Wang, 2016; Hagelin et al., 2017). For this reason, at ZAMG, a new convection-permitting LAMEPS named C-LAEF (Convection-permitting Limited-Area Ensemble Forecasting) based on the Applications of Research To Operations at MEsoscale model (AROME: Seity et al., 2011; Termonia et al., 2018) is being developed.

The goal of this article is to present and discuss an initial condition perturbation method that we plan to use in a first operational configuration of C-LAEF. We propose employing the Jk blending method within the C-LAEF system, for example, within the framework of its 3D-Var ensemble data assimilation system (Bouttier et al., 2012) to generate an ensemble of analyses. We refer to this as an ensemble Jk method. In such a system, small-scale perturbations are generated through 3D-Var EDA, while large-scale perturbations are generated from the host model via Jk blending. In turn, we hypothesize that final analyses are optimal, contain perturbed small and large scales which are, at the same time, consistent with perturbations coming from lateral boundaries. In this way, one can alleviate problems of large-scale forcing and inconsistencies between ICs and LBCs.

This article is organized as follows. In section 2, we present formalisms of the Jk blending method and introduce the ensemble Jk method. Section 3 describes the C-LAEF configuration, experiments and verification procedures performed. Sections 4–6 present our results, and a discussion and conclusions follow in section 7.

2 | INITIAL CONDITION PERTURBATIONS

2.1 | Jk blending method

We now present basic formalisms of the Jk blending method. For a more detailed description, the reader is referred to GF08. 3D-Var is a variational method where, in order to obtain optimal analysis $x_b$, a so-called cost function $J$ that measures the distance of the LAM state vector $x$ to observations $y$, and to the LAM background $x_0$, is minimized2:

$$J(x) = J_b(x) + J_o(x) = \frac{1}{2}(x - x_0)^T B^{-1}(x - x_0) + \frac{1}{2}(y - Hx)^T R^{-1}(y - Hx),$$

(1)

where $B$ and $R$ are error covariance matrices of the background and observations, respectively, and where $H$ is a nonlinear observation operator. In Equation 1, it is assumed that background and observation errors are uncorrelated. In addition, when we assume that host model errors are uncorrelated with observation and background errors, an extra term $J_k(x)$ measuring the distance from large-scale information in $x$ to the host model large-scale information $x_h$ can be added to Equation 1:

$$J(x) = J_b(x) + J_o(x) + J_k(x) = \frac{1}{2}(x - x_0)^T B^{-1}(x - x_0) + \frac{1}{2}(H_1 x - H_1 x_h)^T V^{-1}(H_1 x - H_1 x_h),$$

(2)

where $V$ is the host model large-scale error covariance matrix, $H_1$ is the operator which fetched $x$ from a nominal high-resolution to a low-resolution LAM space, and where $H_k$ is the operator which fetched $x_h$ from the host model low-resolution space to the same low-resolution LAM space as $H_1$. This is done to include only large scales from the host model (e.g. large-scale constraint) as discussed in the introduction. $H_k$ consists of two steps: (a) the interpolation of host model fields to a high-resolution LAM geometry, and (b) its truncation to a lower resolution (see section 3.3), such that only large scales are affected during analysis. We applied the Jk blending method to the AROME 3D-Var system (Fischer et al., 2005), in which $x$ is the state vector of five analysis variables (as is $x_h$): temperature, vorticity, divergence, specific humidity and the logarithm of surface pressure. A large-scale constraint can be applied to all of these or to only some. Our choices regarding this option are discussed in section 3.2.

As stated in GF08, correlations between host model and background errors are different from zero as a result of lateral boundary coupling. However, when one accepts inaccuracies

---

2In AROME, a cost function is formulated using an incremental approach (Courtier et al., 1994; GF08).
of roughly 10%, such correlations can be disregarded, and the total cost function can be represented as the sum of three parts as shown in Equation 2. In this work, it is assumed that this assumption also holds for our model configuration.

### 2.2 Ensemble Jk method

For the IC perturbation method of C-LAEF, we use the ensemble Jk method which combines the ensemble of AROME 3D-Var with the Jk blending method. As discussed in the introduction, one must obtain consistent IC and LBC, which is especially crucial if LAMEPS is integrated over a small domain. To combine large-scale perturbations drawn from the host EPS (the global EPS in this work), which provides the LBC for the LAMEPS, with the small-scale perturbations from LAMEPS, we propose an ensemble data assimilation blending technique. The approach is described as follows:

1. **3D-Var EDA** is used to generate perturbed analyses by:
   a. Random observation perturbations designed to simulate observation errors. For each observation, its error is taken from an observation database and is multiplied by a normally distributed random number with a zero mean and a unit standard deviation. This value is then added to the original observation value to obtain a perturbed observation.
   b. Background perturbations from short-range forecasts of the LAMEPS. For the time being, model errors are not considered. Perturbed-observation EDA perturbations are advantageous in that they are designed to estimate analysis uncertainty (e.g. Wei et al., 2008) which is a desirable property of initial condition perturbations. A theoretical justification of this approach is given in Žagar et al. (2005).

2. To generate correct large-scale information in perturbed analyses that are consistent with perturbed LBCs, we introduce a large-scale constraint by applying Jk blending to 3D-Var EDA. In other words, perturbed global analyses were incorporated as a new term of the 3D-Var cost function and were assimilated with observations described in the last two paragraphs. In brief, each EDA member assimilates the following three components:
   a. Perturbed observations.
   b. Perturbed background.
   c. Perturbed global analysis.

Mathematically, the cost function of one ensemble member of the ensemble Jk method can be written as

\[
J_i(x_i) = \frac{1}{2}(x_i - \bar{x}_{bi})^T B^{-1}(x_i - \bar{x}_{bi}) + \frac{1}{2}(y_i - Hx_i)^T R^{-1}(y_i - Hx_i) + \frac{1}{2}((H_1 x_i - H_b x_{bi}))^T V^{-1}(H_1 x_i - H_b x_{bi}), \tag{3}
\]

where subscript \(i\) denotes the ensemble member. Final perturbed analyses are then found by minimizing these cost functions as in the standard 3D-Var. The resulting blended analyses include small-scale perturbations (LAMEPS) and large-scale perturbations (global EPS) while being consistent with LBC perturbations. We believe that the ensemble Jk method can alleviate the problem of IC and LBC perturbation mismatch and improve the representation of large-scale uncertainties in ensemble systems.

### 3 EXPERIMENTAL DESIGN

#### 3.1 C-LAEF configuration

The C-LAEF system is based on the convection-permitting AROME model as it was operationally configured at ZAMG. AROME is a high-resolution limited-area spectral non-hydrostatic model. It takes most of the ALADIN code on the adiabatic part while its physics package is mainly an adaptation of that used in Mesoscale Non-Hydrostatic research model MESO-NH (Lafore et al., 1998), and it is well adapted to a small grid spacing of roughly 1 or 2 km (Termonia et al., 2018). In addition, the AROME 3D-Var system is almost identical to that developed for the ALADIN system (Fischer et al., 2005). A more detailed description of AROME dynamics, physics and assimilation properties can be found in Seity et al. (2011), Vié et al. (2011) and Brousseau et al. (2016).

In the present study, C-LAEF is configured by the AROME model with a horizontal grid spacing of 2.5 km and 90 vertical levels with 17 ensemble members (16 perturbed plus control). The integration domain used is shown in Figure 1. Due to high computational costs related to C-LAEF, the domain is relatively small at 1,080 km in the north/south direction and 1,500 km in the east/west direction. LBCs are provided by the ECMWF-EPS, and lateral boundary coupling is conducted using the Davies method (Davies, 1976). IC perturbations are described in the next subsection, and for LBC perturbations, 17 members of the ECMWF-EPS are used. Due to technical limitations, the same ECMWF-EPS members used for ALADIN-LAEF (the first 16 plus control) are also used for C-LAEF. As this work’s emphasis is on IC uncertainties,
model errors are not considered, and no model error representation scheme is used. 3D-Var is used for data assimilation. Conventional observations including Aircraft Meteorological Data Relay (AMDR), SYNOP, PILOT, TEMP, SHIP and EUROPROFILES plus two satellite products, GEOWIND atmospheric motion vectors and 25 km Advanced Scatterometer (ASCAT) ocean winds, are assimilated. In this testing phase, 6 h continuous assimilation cycles applied at 0000, 0600, 1200 and 1800 UTC are performed.

3.2 Experiments

The main goal of this work is to evaluate the ensemble \( J_k \) method and to assess its added value to standard perturbed-observation EDA (e.g. Houtekamer et al., 1996; Bouttier et al., 2012), that is, EDA without the \( J_k \) term. For this reason, two experiments are conducted: (a) a reference (REF) experiment in which C-LAERF is using perturbed-observation EDA to generate an ensemble of analyses, and (b) a JK experiment whereby C-LAERF uses the ensemble \( J_k \) method for the same purpose. With the exception of this difference, both experiments use the same configuration as that described in section 3.1. Assimilation is performed at a full model resolution for 17 different ensemble members. As the ensemble \( J_k \) method and perturbed-observation EDA act only on upper-air variables, to allow for a direct comparison, surface variables perturbation (e.g. land albedo, soil moisture, soil temperature, etc.) is not applied. An additional experiment (DOWN) is performed for the period studied in section 6. In DOWN, perturbed analyses were not generated by EDA or by the ensemble \( J_k \) but by the downscaling of ECMWF-ES analyses. This makes them fully consistent with LBC perturbations and DOWN serves as a reference experiment in section 6. Comparison of AROME EDA against the downscale experiment in terms of traditional scores is presented in Raynaud and Bouttier (2016) and will not be repeated here. They concluded that EDA “significantly improves the AROME EPS performance for surface weather variables at early ranges, namely up to 12 h depending on the variable, as measured by the spread/skill relationship and the CRPS over a one-month period.”

In the JK experiment, global ECMWF-ES analyses are used as a source of large-scale information in the ensemble \( J_k \) method – term \( x_i \) as described in section 2.1. Horizontal resolution of the ECMWF-ES in our experiments is spectral cubic T640 with 91 vertical levels with LBCs updated every 3 h. The ECMWF-ES uses singular vectors in combination with EDA for IC perturbation generation (Buzzi et al., 2008) and resulting perturbations are completely independent from ours.

In the \( J_k \) blending method, we test several means of tuning the impact of large-scale constraint. First, we evaluate the selection of variables included in \( x_i \). Multiple tests are performed and model behaviour studied from the use of different variables. We ultimately use all analysis variables with the exception of surface pressure, as this is the best performing option (not shown). Second, a scaling factor (for each analysis variable) controls the relative weight of the \( J_k \) term relative to \( J_b \) and \( J_o \) terms of the cost function, that is, it can be used to fine-tune the impact of all variables separately. This is a value that multiplies standard deviations within the \( V \) matrix. A similar scaling factor for background error standard deviations is standardly used in ALADIN/AROME variational assimilation. A probabilistic approach is used in this case. A scaling factor for each variable is randomly selected from between two reasonable empirically predetermined values, so that each member presents a slightly different combination. Each random number is determined from a seed, and our experiments can be reproduced exactly. Third, the inflation of ECMWF-ES analyses perturbations (calculated by subtracting control analysis from each ensemble member) by a factor of 2 (0) at the lowest (highest) model level with exponential change between is performed before their use in ensemble \( J_k \). After this procedure, newly calculated perturbations are again added to the ECMWF control to create inflated analyses. When this inflation procedure is not applied, the ensemble spread is reduced in JK relative to REF. Inflation restores the spread to similar values as in REF while not affecting other scores such as the RMSE.

3.3 V matrix and truncation selection

As illustrated by Equations 2 and 3, in addition to \( B \) and \( R \) covariance matrices, a \( V \) covariance matrix must also be provided. This is an important step because \( V \) controls the impact of large-scale constraint on the final analysed state. In calculating \( V \), we used the same procedure normally applied to \( B \), for example, the ensemble simulation method (ESM: Berre et al., 2006), meaning that \( V \) is climatological and does not contain any flow-dependent information. To obtain climatological host model large-scale error covariances using the ESM, forecast differences between 16 ECMWF-ES members at analysis time need to be calculated. This calculation is done for 0000 and 1200 UTC runs over two-week periods in January, April, July and October of 1 year to include seasonal variability of error covariances, as this can have a significant impact on forecast quality (Storto and Randriamampianina, 2010). As in GF08, \( V \) is formulated in a univariate manner.

The \( J_k \) term needs to be truncated such that only large scales are kept from the host model. To do so, we adopted the same procedure as that applied in GF08. Figure 2 shows the horizontal error variance spectrum (Berre, 2000) of \( V \) and \( B \) matrices for two levels (roughly 500 and 920 hPa) and variables (divergence and specific humidity). It can be observed that at scales of above the total wave number \( k^* \) (as defined in Berre, 2000) roughly 10, two spectra begin to diverge, and little energy is found in mesoscale spectra of the host model. We thus did not expect to derive any useful information from the host model at these scales, and \( J_k \) truncation was
set at $k^* = 8$ which corresponds to wavelengths of roughly 135 km.

### 3.4 Verification

JK and REF experiments described above are verified in 3 h intervals for surface variables, that is, 2 m temperature (T2M), 2 m relative humidity (RH2M), 10 m wind speed (W10M) and mean-sea-level pressure (MSLP) against observations drawn from 832 available surface stations within the verification domain (Figure 1). Model fields are interpolated via bilinear interpolation to observation locations. Verifying precipitation forecasts in convection-permitting models has its difficulties and traditional verification methods may not be appropriate (Ebert, 2008; Mittermaier, 2014). In this work, we use Fractions Skill Score (FSS: Roberts and Lean, 2008), which is a form of a neighbourhood approach (Gilleland et al., 2009), and it is computed following the ensemble formulation proposed by Duc et al. (2013). Integrated Nowcasting through Comprehensive Analyses (INCA: Haiden et al., 2011; Wang et al., 2017) analyses are used as a reference for precipitation verification. The INCA system is currently being developed at ZAMG, and it uses data drawn from automatic weather stations, remote-sensing data (radar, satellite), forecast fields of NWP models, and high-resolution topographic data (Haiden et al., 2011) to generate a high-resolution precipitation analysis field for a domain covering Austria and surroundings (Figure 1).

For upper-air variables, verification is done every 6 h against high-resolution deterministic ECMWF ($0.1^\circ \times 0.1^\circ$) and deterministic NCEP analyses ($0.5^\circ \times 0.5^\circ$). This approach is applied because the ensemble Jk method uses perturbed ECMWF analyses and, thus verification against deterministic ECMWF analyses can be biased toward the JK experiment. Verifying variables include temperature, relative humidity and wind speed at 500 and 850 hPa pressure levels (T500, T850, RH500, RH850, W500 and W850, respectively).

In general, probabilistic forecast verification is more difficult to perform than deterministic forecast verification because probabilistic forecast values are not always measured as 0 or 1. In addition, good ensemble prediction systems should generate superior ensemble mean forecasts over the control, high spread-skill relations and reliable probability forecasts (Du, 2007). Furthermore, Murphy (1993) described aspects or attributes that contribute to forecasting quality (e.g. accuracy, bias, reliability, resolution, discrimination, etc.). Given these sources of complexity, no single score can assess all attributes of probabilistic forecasts at once. For this reason, various scores are used to assess different aspects of forecasts.
and to draw conclusions about the EPSs under consideration. These include (a) the root-mean-square error (RMSE) of ensemble mean for assessing ensemble mean accuracy, (b) the continuous rank probability score (CRPS) for assessing overall EPS skill, (c) ensemble RMSE/spread relation and outlier statistics for assessing reliability, and (d) relative operational characteristics (ROC) for assessing discrimination. Detailed descriptions of these scores can be found in Talagrand et al. (1997) and Wilks (2006).

To determine if the difference in scores between the experiments is statistically significant, the moving-block bootstrap technique, following the procedure of Wilks (1997) and using 5,000 re-samples at a 90% confidence level, was applied. Block length (four in our case) was defined as a closest integer to a cube root of the sample length (Hall et al., 1995). Observation errors are not considered, as this extends beyond the scope of our verification package. Nevertheless, as observation errors can have a significant effect on final scores (e.g. Bowler, 2008; Candille and Talagrand, 2008; Duc and Saito, 2018), we discuss this limitation in section 7.

4 | RESULTS

In this section, we present the results of a two-month integration of REF and JK experiments to illustrate differences in average model behaviour. Experiments were performed for 62 days from 1 July until 31 August 2016 with 1200 UTC runs with a 24 h forecast range. These months involved several episodes of strong convection occurring over central Europe, rendering them appropriate periods for the verification of a convection-permitting EPS.

Due to the limited domain used, a forecast range of 24 h proved more than sufficient for our IC perturbation impact study, as LBC perturbations dominate even before this time is reached. For upper-air variables, only verification results done against independent NCEP analyses are presented, but conclusions remain the same if verification is done against ECMWF analyses. However, verifying against high-resolution ECMWF analyses yields even better JK results with an increased percentage of statistically significant results relative to those presented in the following subsections (not shown).

4.1 | RMSE of ensemble mean and spread

In a statistically consistent or reliable EPS, the RMSE of ensemble mean should match its spread (Fortin et al., 2014). This follows from the fact that in a perfect ensemble (i.e. a perfect sampling of true state uncertainty), truth should be statistically indistinguishable from any ensemble member...
Thus, the RMSE/spread relation is a good measure of reliability. Figure 3 shows RMSE and spread for T500, T850, RH500, RH850, W500 and W850. Positive effects of the ensemble Jk method on the accuracy of the ensemble mean can be clearly identified because RMSE is reduced for almost all variables and at both levels in JK. However, differences observed are mostly statistically significant within early forecast ranges (up to 12 h). Spread remains similar in JK and REF at 850 hPa but is reduced at 500 hPa in JK, which is a favourable result because JK would be over-dispersive otherwise. As REF is under-dispersive, reducing RMSE more than spread enhances the reliability of JK.

RMSE and spread for surface variables are shown in Figure 4, in which results are more neutral. There is a slight improvement in RMSE of JK for T2M, RH2M and W10M, mostly over the first 3–9 h and for MSLP from 9 to 18 h. However, none of the differences observed are significant. Spread remains almost the same for T2M and RH2M, while it is slightly higher in the case of MSLP and lower for W10M in JK.

Overall, both REF and JK are under-dispersive for both upper-air variables and especially for surface variables. These results are consistent with those of numerous other studies (Buizza et al., 2005). This is likely the case due to an absence of model and surface perturbations and because observation errors are not accounted for.

### 4.2 Continuous ranked probability score (CRPS)

The CRPS is related to the rank probability score (RPS) but for an infinite number of classes; thus, it compares the full predicted distribution with observations. The CRPS is a good measure of overall EPS skill. It is dependent on a predicted distribution’s first (ensemble mean) and second moment (spread) and is negatively oriented, meaning that lower values are better.

The CRPS for upper-air and surface variables are shown in Figures 5 and 6, respectively. Better CRPS is obtained for all upper-air variables for all levels for JK. Significance test results are similar to RMSE results – significant within early forecast ranges for upper-air (up to 12 h) and up to 18 h for RH850. For the surface, significant improvements in CRPS are obtained at 9 and 12 h for MSLP in JK. Insignificant improvements observed in the first forecast hours for T2M in JK are also visible here. Thus, the CRPS supports
conclusions drawn from Figures 3 and 4 that JK is more skilful for upper-air variables and MSLP.

4.3 Outlier statistics

Outliers are defined as the number of cases in which verifying observations lie outside of the whole ensemble and are typically expressed as percentages. The perfect outlier percentage is calculated as \(2/(N + 1) \times 100\%\) where \(N\) is the ensemble size (thus, a perfect score is 11\% in our experiments). The percentage of outliers serves as a good measure of reliability.

Figure 7 shows the percentage of outliers for upper-air variables. The number of outliers is greatly reduced in JK for T500 and T850, with weaker effect found for RH500 and RH850 and with almost neutral effect found for W850 and W500. This means that reliability is mostly improved for temperature and relative humidity. Significance test results support these conclusions.

Regarding the surface variables and results given in Figure 8, the greatest reduction in the number of outliers in JK is achieved for MSLP, which is also significant at 9 and 12 h forecast range, resulting in a more reliable ensemble. Slight and insignificant reductions are also observed for T2M.

4.4 Relative operating characteristics (ROC)

The ROC curve plots the hit rate against false alarm rate for a set of predefined probability thresholds (in our case, the difference between two consecutive thresholds is \(1/N\)). It compares two conditional distributions where the hit rate corresponds to occurrences and the false alarm to non-occurrences of an event measuring the attribute of discrimination.

The ROC curve of upper-air variables for the 6 h forecast range is shown in Figure 9. The figure shows a slight improvement in the discrimination of humidity and wind for all levels in JK, showing that JK can better distinguish between the occurrences and non-occurrences of an event. For the surface variables, the situation is more neutral with no clear improvement for JK (not shown).

More neutral results for surface variables and for all scores presented in this section support the conclusions of Dahlgren and Gustafsson (2012) and are as expected for a number of reasons. First, lower boundary conditions have a considerable impact on surface variables. Second, surface variables are sensitive to small-scale information while the ensemble JK method acts only on upper-air variables and at the largest scales.

4.5 Precipitation

Figure 10 shows FSS for 3 h precipitation of JK versus REF as a function of the lead-time for the whole verification period (62 days of July and August 2016). At a given lead-time, 62 skill scores (one for each day) are computed of the FSS of JK using the REF FSS as reference. This second normalization
of the score is computed to avoid the variability of the FSS score depending on the rainfall of both experiments. The left panel shows the median of the Skill Score of FSS of JK to FSS of REF illustrated as a matrix of colours (red means JK better than REF and blue the opposite, white is for no difference) for different thresholds, scales and forecast ranges. The right panel is the significance level for the comparison (meaning the percentage of times FSS of JK is higher than the FSS of REF). At 6 h forecast range, precipitation forecast is improved for all thresholds (scales above 195 km) and for thresholds up to 5 mm (scales above 45 km). For 12 and 18 h forecast ranges, neutral results are observed, while for 24 h forecast range, precipitation forecast is improved for all scales (threshold of 1 mm) and for higher thresholds (scales above 195 km). Results presented here indicate that, at early forecast ranges, mostly larger scales are affected (this is in agreement with the fact that only large scales are affected in the Jk method), while at later forecast ranges, nonlinear model integration spreads the influence on all scales.

5 | CASE-STUDY

We now illustrate benefits of including large-scale perturbations from the host model through a case-study. Figure 11 shows a synoptic situation observed on 11 July 2016 at 0000 and 1200 UTC. As is shown, a cold front related to a cyclone over the North Sea is moving over the C-LAEF domain. At the same time, on the ground strong advection involving warm and moist air from the Mediterranean area is observed. This synoptic set-up has a great potential to trigger convection in areas preceding the front.

We expect to observe benefits of including large-scale perturbations in cases such as the one described here (i.e. strong synoptic forcing located near the borders of the computational domain). Any information outside of the model domain is propagated inside through LBCs. However, due to well-known deficiencies related to lateral boundary formulation in LAMs (e.g. Warner et al., 1997; Nutter et al., 2004; Termonia et al., 2009) and due to LAM orientation toward the mesoscale, this information can be deficient. We believe that introducing large-scale information in an additional way – through Jk – can alleviate such problems.

Figure 12 (left column) illustrates ensemble mean MSLP for REF and JK experiments started at 0000 UTC on 11 July 2016 from 3 and 9 h forecasts. The overall distribution in pressure observed is similar across the whole domain. The biggest difference, considering all forecast hours, is visible over the upper half of the domain where lower pressure, related to an
incoming frontal system, propagates further to the south in JK (note how 1,011 and 1,014 hPa contours are displaced more to the south in the left column of Figure 12).

The right column of Figure 12 shows the difference between REF and JK ensemble mean absolute error (forecast – observation) plotted on surface station locations. As the front enters and moves through the domain, an interesting phenomenon is observed: improvements in MSLP forecast for JK follow the progression of the frontal system. Smaller errors of up to 1.5 hPa are observed in the northern part of the domain from which the cold front progresses towards inland Europe. Pressure distribution is clearly better represented in JK. Although surface pressure large-scale information was not used in the JK method directly, improvements in surface pressure forecast come from the other variables as surface pressure reflects the entire column.

In addition, Figure 13 (Supporting Information Figure S1) shows the probability of 3 h precipitation above 1 mm (10 mm) calculated from 3–6 to 6–9 h forecasts for REF and JK experiments started at 0000 UTC on 11 July 2016. Given that no precipitation is observed (Supporting Information Figure S2) for these hours, our REF experiment generates excessively high probability values (up to 60%) for precipitation over southern Germany and western Austria. For the JK, the area affected and probabilities are reduced with the latter remaining at below 20% for 3–6 h and at 30% for 6–9 h forecasts, thus minimizing the possibility of a false alarm. Later during that day, convection developed as the front moved through the domain. Figure 14 (Supporting Information Figure S3) shows the probability of 3 h precipitation above 1 mm (10 mm) calculated from 9 to 12 h forecasts for REF and JK experiments started at 1200 UTC on 11 July 2016 and INCA analysis. The JK experiment better captured overall precipitation distribution, especially over northern Austria and southern Czech Republic and a dry area over southeastern Germany which are both missed in REF. Intense precipitation in the north of Italy was better forecasted in JK with above 90% probability for rain above 10 mm compared to around 40% in REF (Supporting Information Figure S3). However, both experiments missed the intense precipitation over central Austria. The area over southern Switzerland is not considered in the analysis because, in INCA, observation data are scarce there and the quality of the radar data is suboptimal. Overall, the results presented here demonstrate the benefits of applying large-scale constraints on C-LAEF perturbations.

6 | PERTURBATION MISMATCHING

We now turn to the second main focus of this work – mismatched perturbations. An additional experiment (DOWN) was performed whereby IC perturbations are fully consistent with LBC perturbations. For this reason, the DOWN experiment serves as a new reference experiment.
In addition, all experiments are performed using 1 h LBC update frequency.

As stated by Warner et al. (1997), when solutions of the LAM and host model for areas close to lateral boundaries are different, such differences may generate spurious gradients and feedback between the two grids, which can influence the interior of the LAM domain. Given that EDA IC perturbations observed in the REF experiment are completely independent of LBC perturbations originating from ECMWF-EPS, such differences can arise.

One such event was observed for the REF run started on 17 July 2016 at 0000 UTC. Figure 15 shows the MSLP spread for the first hour of model integration. After only 5 min (Figure 15b), a large spread (with maximum value of roughly 3.5 hPa) can be observed close to the northeastern and along the southern domain boundary. As integration continues (Figure 15c,d), the anomaly in spread advances further into the domain at a velocity magnitude close to the speed of sound. After 1 h (Figure 15d), excessive spread (relative to the DOWN result shown in Figure 16a) has not left the domain and approaches the western and eastern lateral boundaries. Given that anomalies form at the boundaries and not in DOWN or the host model (not shown), it is clear that they are unrealistic and likely formed as a result of unrealistic waves generated by an IC and LBC perturbation mismatch. The fact that excessive spread is still present after 1 h of integration can create serious issues for 1 h assimilation cycles (e.g. Caron, 2013).

A potential solution to this problem could come from so-called space consistent coupling, whereby LBCs at initialization time do not come from the host model but from the LAM analysis (Brousseau et al., 2016). In turn, LBC perturbations are not imposed with full weight from the start and are instead slowly introduced over the first forecast hour to reduce the risk of perturbation conflict. Figure 16b shows the results of such integration observed after 1 h. Excessive spread is still observed close to the northwestern and eastern lateral boundaries and does not seem to be reduced at all (compared to Figure 15d), leading us to conclude that space consistent coupling is unable to solve this issue for such an event.

However, by using 3 h LBC update frequency, space consistent coupling can solve this problem (not shown) because in this case, LBC perturbations are introduced even more slowly.
Nevertheless, given the benefits of higher LBC update frequency (e.g. Nutter et al., 2004; Termonia et al., 2009), the availability of global EPS 1 h files (e.g. ECMWF) and continuing efforts made to improve LAMEPSs in general, we assume that results derived from a 1 h LBC update frequency are more significant.

An alternate approach involves using the ensemble \( J_k \) method, for which LAM IC perturbations are more consistent with LBC perturbations by design. The results of integration using the ensemble \( J_k \) method for the same period are shown in Figure 16c. Excessive spread is largely reduced after 1 h of integration, and overall spread distribution resembles the DOWN pattern most consistently (although the spread is still slightly larger than in DOWN). Increased spread to the east of the domain (also found in later forecast ranges – not shown) is also found in DOWN and can be explained by the presence of a cyclone over this part of the domain. From these results, we conclude that the ensemble \( J_k \) method can indeed be used to alleviate the problem of mismatching perturbations and that constraining only the largest scales (above 135 km as discussed in section 3.3) can achieve this goal.

**7 | SUMMARY AND CONCLUDING REMARKS**

Probabilistic approaches to weather prediction are widely recognized as superior to deterministic and ensemble prediction systems and are becoming more widely favoured forecast tools. This is especially true for LAMEPSs, which have been used more frequently over the last 10 years. Thus, motivations to resolve problems related to LAMEPSs are strong. Due to the typically small computational domains and deficiencies of LBC formulations, LAMEPSs are unable to correctly account for uncertainties of large-scale flow. Furthermore, mismatches between regional EPS (ICs) and host EPS lateral boundary perturbations can result when they are obtained independently.

In this article, we introduce an ensemble \( J_k \) method as a new idea for improving LAMEPS forecasts, whereby host model large-scale perturbations are introduced into LAMEPS IC perturbations using the \( J_k \) blending method. \( J_k \) blending is applied to the 3D-Var EDA system so that perturbed analyses are optimal as opposed to using the blending method developed by Wang et al. (2014). As a result, analyses contain perturbed small scales (LAM EDA) and large scales (host EPS). Furthermore, when including host model perturbations in LAM initial perturbations, final perturbations are more consistent with perturbations coming from lateral boundaries, thus minimizing the risk of potential mismatch.

The ensemble \( J_k \) method was implemented into convection-permitting ensemble system C-LAEF with ECMWF-EPS providing lateral boundary conditions and large-scale perturbations. To assess the effect of large-scale constraint on IC perturbations, two experiments were conducted,
FIGURE 10  The left panel is the median Skill Score of FSS of JK to FSS of REF (red means JK is better than REF and blue the opposite) as a function of scale and threshold. The right panel is the significance level for the comparison (percentage of times FSS of JK is higher than the FSS of REF). The plots are shown for (a) 6 h, (b) 12 h, (c) 18 h and (d) 24 h forecast ranges.
one using an ensemble \( J_k \) and the other using standard perturbed-observation EDA, that is, EDA without \( J_k \) blending. Impacts on average model performance were measured over 62 convectively active days of the summer of 2016.

Results show that using the ensemble \( J_k \) method provides a more accurate and reliable EPS. This is mostly the case due to a reduction of RMSE and thus due to increased accuracy. RMSE of ensemble mean is strongly reduced for...
upper-air variables, while spread is maintained on similar or slightly lower values. The percentage of outliers supports these conclusions as does the CRPS, which is reduced for all variables. The ROC score also benefits ensemble $Jk$ results, meaning that discrimination is improved.

Performance in terms of surface variables is more neutral, as was expected due to its high sensitivity to small scales and the forcing effects of lower boundary conditions. Nevertheless, slight but insignificant improvements resulting from using ensemble $Jk$ can be observed for early forecast ranges.
(up to 9 h) for temperature and humidity, and for surface pressure, improvements are visible up to 18 h. The strongest effect of ensemble $Jk$ is observed for surface pressure forecasts, which are more reliable and accurate. Discrimination remains similar for all variables. Precipitation forecasts were verified by a different methodology and by using the FSS. Improvements for thresholds up to 20 mm and for scales ranging from the highest (325 km) down to the smallest (5 km) are observed when using the ensemble $Jk$ method.

To further illustrate benefits of considering large-scale perturbations from the host model, a case-study of strong synoptic forcing observed near the borders of the computational domain is presented. It is demonstrated that using the ensemble $Jk$ method improves surface pressure distribution in areas positioned close to an incoming front, in turn minimizing the possibility of the false detection of precipitation and improving the precipitation forecast during a period of convective activity over Austria.

Finally, a case of strong perturbation mismatch was presented. Strong anomalies in surface pressure spread were observed close to the northeastern and southern domain boundaries after only 5 min of model integration. After 1 h, excessive spread was still present close to the edges of the domain. Introducing LBC perturbations gradually over the first hour of integration, for example, space consistent coupling, did not remedy this problem. However, when ensemble $Jk$ method was used and when large-scale IC perturbations were made more consistent with LBC perturbations, even only for the largest scales, excessive spread was mitigated after 1 h. From these findings, it is clear that future implementations of C-LAEF with hourly assimilation cycles must address the problem of perturbation mismatch, which was the leading cause of excessive spread observed here. The ensemble $Jk$ method can be used to alleviate such problems.

Overall, using the ensemble $Jk$ method benefits the ensemble performance of upper-air variables, precipitation and mean-sea-level pressure, while effects on the other surface variables are less significant and mostly reserved for special cases, such as those discussed in section 5. Observed impacts are strongest within the early forecast range, which is to be expected from an IC perturbation study (e.g. Vié et al., 2011). Furthermore, both of our experiments are under-dispersive, particularly for lower model levels. This is the case due to an absence of model and surface perturbations and to a lack of consideration of observation errors. In addition, due to the relatively large difference in grid spacing between C-LAEF and ECMWF-EPS, LBC “sweeping” effect (Nutter et al., 2004) is enhanced and plays a significant role in later forecast ranges, placing an additional constraint on ensemble spread growth.
As noted above, we did not conduct observation error simulations of any kind. Though impacts of observation error can change the outcome of forecast verification, we believe that this was not the case in our study. First, we were only interested in relative performance, that is, differences observed between experiments, and including the observation error would have had an effect of the same direction for both experiments. Second, RMSE/spread relation and outlier statistics are scores for which the absolute performance of each system is also important. As shown by other works (e.g. Bouttier et al., 2012; Romine et al., 2014), accounting for observation error would have increased ensemble spread given in Figures 3 and 4. This could mean that an EPS using the ensemble JK method could be declared over-dispersive in such a case. However, a slight modification of the ensemble JK method can be done to account for such over-dispersion. The inflation factor discussed in section 3.2 can be reduced to decrease the spread while keeping RMSE unchanged (not shown).

**ACKNOWLEDGEMENTS**

We gratefully acknowledge Jelena Bojarova for providing helpful comments during JK blending implementation, Kristian Horvath for his help in revising the manuscript and RC LACE for supporting this research.

**REFERENCES**

ALADIN International Team. (1997) The ALADIN project mesoscale modelling seen as basic tool for weather forecasting and atmospheric research. *WMO Bulletin*, 46, 317–324.

Berre, L. (2000) Estimation of synoptic and mesoscale forecast error covariances in a limited area model. *Monthly Weather Review*, 128, 644–667.

Berre, L., Ștefănescu, S.E. and Belo-Pereira, M. (2006) The representation of the analysis effect in three error simulation techniques. *Tellus*, 58A, 196–209.

Bouttier, F., Vié, B., Nuisierr, O. and Raynaud, L. (2012) Impact of stochastic physics in a convection-permitting ensemble. *Monthly Weather Review*, 140, 3706–3721.

Bouttier, F., Raynaud, L., Nuisier, O. and Ménétrier, B. (2015) Sensitivities of the AROME ensemble to initial and surface perturbations during HyMeX. *Quarterly Journal of the Royal Meteorological Society*, 142(S1), 390–403.

Bowler, N.E. (2008) Accounting for the effect of observation errors on verification of MOGREPS. *Meteorological Applications*, 15, 199–205.

Bowler, N.E. and Mylne, K.R. (2009) Ensemble transform Kalman filter perturbation for a regional ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 135, 757–766.

Brousseau, P., Berre, L., Bouttier, F. and Desroziers, G. (2011) Background-error covariances for a convective-scale data-assimilation system: AROME–France 3D-Var. *Quarterly Journal of the Royal Meteorological Society*, 137, 409–422.

Brousseau, P., Seity, Y., Ricard, D. and Léger, J. (2016) Improvement of the forecast of convective activity from the AROME-France system. *Quarterly Journal of the Royal Meteorological Society*, 142, 2231–2243.
Brožková, R., Klaric, D., Ivaťok-Săhălu, S., Geleyn, J.F., Cassé, V., Siroka, M., Radnòti, G., Janousek, M., Stadlbacher, K. and Seidl, H. (2001) DFI blending: an alternative tool for preparation of the initial conditions for LAM. In: Ritchie, H. (Ed.) Research Activities in Atmospheric and Oceanic Modelling. Geneva, Switzerland: WMO CAS/JSC WGNE Rep. 31, pp. 1.7–1.8.

Bowler, N.E., Arribas, A., Mylne, K.R., Robertson, K.B. and Beare, S.E. (2008) The MOGREPS short-range ensemble prediction system. Quarterly Journal of the Royal Meteorological Society, 134, 703–722.

Buizza, R. and Palmer, T.N. (1995) The singular vector structure of the atmospheric general circulation. Journal of the Atmospheric Sciences, 52, 1434–1456.

Buizza, R., Houtekamer, P.L., Toth, Z., Pellerin, G., Wei, M. and Zhu, Y. (2005) A comparison of the ECMWF, MSC and NCEP global ensemble prediction systems. Monthly Weather Review, 133, 1067–1097.

Buizza, R., Leutbecher, M. and Isaksen, L. (2008) Potential use of an ensemble of analyses in the ECMWF ensemble prediction system. Quarterly Journal of the Royal Meteorological Society, 134, 2051–2066.

Candille, G. and Talagrand, O. (2008) Impact of observational error on the validation of ensemble prediction systems. Quarterly Journal of the Royal Meteorological Society, 134, 959–971.

Caron, J.F. (2013) Mismatching perturbations at the lateral boundaries in limited area ensemble forecasting: a case study. Monthly Weather Review, 141, 356–374.

Charney, J.G. (1951) Dynamical forecasting by numerical process. In: Candille, G. and Talagrand, O. (2008) Impact of observational error on the validation of ensemble prediction systems. Quarterly Journal of the Royal Meteorological Society, 134, 703–722.

Ebert, E. (2008) Fuzzy verification of high-resolution gridded forecasts: a review. Quarterly Journal of the Royal Meteorological Society, 120, 1367–1387.

Dahlgren P, Gustaffson N. (2012) Assimilating host model information into a limited area model. Tellus, 64A, 15836. https://doi.org/10.3402/tellusa.v64i0.15836.

Davies, H.C. (1976) A lateral boundary formulation for multi-level prediction models. Quarterly Journal of the Royal Meteorological Society, 102, 405–418.

Du, J. (2007) Uncertainty and ensemble forecast. National Weather Service, Office of Science & Technology, Science & Technology Infusion Lecture Series, 42 pp. Retrieved from: https://www.nws.noaa.gov/ost/climate/STIP/STILecture1.pdf

Duc L, Saito K, Seko H. (2013) Spatial-temporal fractions verification for limited-area variational assimilation. Quarterly Journal of the Royal Meteorological Society, 134, 1416–1430.

Guidard, V. and Fischer, C. (2008) Introducing the coupling information in a limited-area variational assimilation. Quarterly Journal of the Royal Meteorological Society, 134, 723–735.

Gustaffson N, Janic J, Schraff C, Leuenberger D, Weissmann M, Reich H, Brousseau P, Montmerle T, Wattrelot E, Bučánek A and Milde M. (2018) Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. Quarterly Journal of the Royal Meteorological Society, 144, 1218–1256. https://doi.org/10.1002/qj.3179.

Hagelin, S., Svinhøfn J., Son, R., McCabe, A., Roberts, N. and Tennant, W. (2017) The Met Office convective-scale ensemble. MOGREPS-UK. Quarterly Journal of the Royal Meteorological Society, 143, 2846–2861.

Haiden, T., Kann, A., Wittmann, C., Piotniski, G., Bica, B. and Gruber, C. (2011) The Integrated Nowcasting through Comprehensive Analysis (INCA) system and its validation over the eastern Alpine region. Weather and Forecasting, 26, 166–183.

Hall, P., Horowitz, J.L. and Jing, B.-Y. (1995) On blocking rules for the bootstrap with dependent data. Biometrika, 82, 561–574.

Harnisch, F. and Keil, C. (2015) Initial conditions for convective-scale ensemble forecasting provided by ensemble data assimilation. Monthly Weather Review, 143, 1583–1600.

Hohenegger, C. and Schär, C. (2007) Atmospheric predictability at synoptic versus cloud-resolving scales. Bulletin of the American Meteorological Society, 88, 1783–1793.

Houtekamer, P.L., Lefaivre, L., Derome, J., Ritchie, H. and Mitchell, H.L. (1996) A system simulation approach to ensemble prediction. Monthly Weather Review, 124, 1225–1242.

Houtekamer, P.L. and Mitchell, H.L. (2005) Ensemble Kalman filtering. Quarterly Journal of the Royal Meteorological Society, 131, 3269–3289.

Hisso, L.F., Huang X.Y., Kuo, Y.H., Chen, D.S., Wang, H., Tsai, C.C., Yeh, T.C., Hong, J.S., Fong, C.T. and Lee, C.S. (2015) Blending of global and regional analyses with a spatial filter: application to typhoon prediction over the western North Pacific Ocean. Weather and Forecasting, 30, 754–770.

Hu, M., Benjamin, S.G., Ladwig, T.T., Dowell, D.C., Weygandt, S.S., Alexander, C.R. and Whitaker, J.S. (2017) GSI three-dimensional ensemble-variational hybrid data assimilation using a global ensemble for the regional Rapid Refresh model. Monthly Weather Review, 145, 4205–4225.

Johnson, A. and Wang, X. (2016) A study of multiscale initial condition perturbation methods for convective-permitting ensemble forecasts. Monthly Weather Review, 144, 2579–2604.

Judd, F. (2018) Insights into atmospheric predictability through global convection-permitting model simulations. Journal of the Atmospheric Sciences, 75, 1477–1497.

Kalnay, E. (2003) Atmospheric Modelling, Data Assimilation and Predictability. Cambridge: Cambridge University Press.

Kühnlein, C., Keil, C., Craig, G.C. and Gebhardt, C. (2014) The impact of down-scaled initial condition perturbations on convective-scale ensemble forecasts of precipitation. Quarterly Journal of the Royal Meteorological Society, 140, 1552–1562.

Lafore, J.P., Stein, J., Asencio, N., Bougeault, P., Ducrocq, V., Duron, V., Fischer, C., Hérail, P., Massart, P., Masson, V. and Pinty, I.P. (1998) The Meso-NH atmospheric simulation system. Part I: Adiabatic formulation and control simulations. Annales Geophysicae, 16, 90–109.

Leutbecher, M. and Palmer, T.N. (2008) Ensemble forecasting. Journal of Computational Physics, 227, 3515–3539.

Lewis, J.M. (2005) Roots of ensemble forecasting. Monthly Weather Review, 133, 1865–1885.

Lorenz, E.N. (1963) Deterministic nonperiodic flow. Journal of the Atmospheric Sciences, 20, 130–141.

Lorenz, E.N. (1965) A study of the predictability of a 28-variable atmospheric model. Tellus, 17, 321–333.

Lorenz, E.N. (1969) How much better can weather prediction become? Technology Rev., July/August: 39–49.

Mittermaier, M. (2014) A strategy for verifying near-convection-resolving model forecasts at observing sites. Weather and Forecasting, 29, 185–204.

Molteni, F., Buizza, R., Palmer, T.N. and Petroligas, T. (1996) The ECMWF ensemble prediction system: methodology and validation. Quarterly Journal of the Royal Meteorological Society, 122, 73–119.

Murphy, A.H. (1993) What is a good forecast? An essay on the nature of goodness in weather forecasting. Weather and Forecasting, 8, 281–293.

Nussier, O., Joly, B., Vié, B. and Ducrocq, V. (2012) Uncertainty on lateral boundary conditions in a convection-permitting ensemble: a strategy of selection for Mediterranean heavy precipitation events. Natural Hazards and Earth System Sciences, 12, 2993–3011.

Nutter, P., Stensrud, D. and Xue, M. (2004) Effects of coarsely resolved and temporally interpolated lateral boundary conditions on the dispersion of limited-area ensemble forecasts. Monthly Weather Review, 132, 2358–2377.
Perala, C., Ben Bouallègue, Z., Theis, S.E., Gebhardt, C. and Buchhold, M. (2012) Accounting for initial condition uncertainties in COSMO-DE-EPS. Journal of Geophysical Research, 117(D7), 1–13.

Raynaud, L. and Bourtier, F. (2016) Comparison of initial perturbation methods for ensemble prediction at convective scale. Quarterly Journal of the Royal Meteorological Society, 142, 854–866.

Roberts, N.M. and Lean, H.W. (2008) Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. Monthly Weather Review, 136, 78–97.

Romine, G.S., Schwartz, C.S., Berner, J., Fossell, K.R., Snyder, C., Anderson, J.L. and Weisman, M.L. (2014) Representing forecast error in a convection-permitting ensemble system. Monthly Weather Review, 142, 4519–4541.

Saito K, Seko H, Kunii M, Miyoshi T. (2011) Effect of lateral boundary condition temporal resolution problem and a proposed solution by means of ensemble transform (ET) technique in the NCEP global ensemble forecast system. Tellus, 63A, 11594. https://doi.org/10.1111/j.1600-0870.2011.00556.x.

Schellander-Gorgas, T., Wang, Y., Meier, F., Weidle, F., Wittmann, C. and Kann, A. (2017) On the forecast skills of a convection permitting ensemble. Geoscientific Model Development, 10, 35–56.

Seity, Y., Brousseau, P., Marelard, S., Hello, G., Bénard, P., Bourtier, F., Lac, C. and Masson, V. (2011) The AROME-France convective-scale operational model. Monthly Weather Review, 139, 976–991.

Storto A and Randriamampianina R. (2010) Ensemble variational assimilation for the representation of background error covariances in a high-latitude regional model. Journal of Geophysical Research, 115, 217204. https://doi.org/10.1029/2009JD012988.

Schwartz, C.S., Romine, G.S., Sobash, R.A., Fossell, K.R. and Weisman, M.L. (2015) NCAR’s experimental real-time convection-allowing ensemble prediction system. Weather and Forecasting, 30, 1645–1654.

Schwartz, C.S. (2016) Improving large-domain convection-allowing forecasts with high-resolution analyses and ensemble data assimilation. Monthly Weather Review, 144, 1777–1803.

Talagrand, O., Vautard, R. and Strauss, B. (1997) Evaluation of probabilistic prediction systems. In: Proc. ECMWF Workshop on Predictability, Shinfield Park, Reading, UK: ECMWF, pp. 1–26.

Termonia, P., Deckmyn, A. and Hamdi, R. (2009) Study of the lateral boundary condition problem and a proposed solution by means of boundary error restarts. Monthly Weather Review, 137, 3551–3566.

Termonia, P., Fischer, C., Bazile, E., Bousseyl, F., Brožková, R., Bénard, P., Bochenek, B., Degruawe, D., Derková, M., El Khatib, R. and Hamdi, R. (2018) The ALADIN system and its canonical model configurations AROME CY41T1 and ALARO CY40T1. Geoscientific Model Development, 11, 257–281.

Toth, Z. and Kalnay, E. (1993) Ensemble forecasting at NMC: the generation of perturbations. Bulletin of the American Meteorological Society, 74, 2317–2330.

Toth, Z. and Kalnay, E. (1997) Ensemble forecasting at NCEP: the breeding method. Monthly Weather Review, 125, 3297–3318.

Vié, B., Molinié, G., Nuissier, O., Vincendon, B., Ducrocq, V., Bourtier, F. and Richard, E. (2011) Hydro-meteorological evaluation of a convection-permitting ensemble prediction system for Mediterranean heavy precipitating events. Natural Hazards and Earth System Sciences, 12, 2631–2645.

Wang, X. and Bishop, C. (2003) A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. Journal of the Atmospheric Sciences, 60, 1140–1158.

Wang, Y., Belluš, M., Wittmann, C., Steinheimer, M., Weidle, F., Kann, A., Ivatek-Sahdan, S., Tian, W., Ma, X., Tascu, S. and Bazile, E. (2011) The Central European limited-area ensemble forecasting system: ALADIN-LAEF. Quarterly Journal of the Royal Meteorological Society, 137, 483–502.

Wang, Y., Tascu, S., Weidle, F. and Schmeisser, K. (2012) Evaluation of the added value of regional ensemble forecasts on global ensemble forecasts. Weather and Forecasting, 27, 972–987.

Wang, Y., Belluš, M., Geleyn, J.F., Ma, X., Tian, W. and Weidle, F. (2014) A new method for generating initial condition perturbations in a regional ensemble prediction system: blending. Monthly Weather Review, 142, 2043–2059.

Wang, Y., Meirold-Mautner, I., Kann, A., Slak, A.S., Simon, A., Vivoda, J., Bica, B., Böcksör, E., Brezóková, L., Dantinger, J. and Giszterowicz, M. (2017) Integrating nowcasting with crisis management and risk prevention in a transnational and interdisciplinary framework. Meteorologische Zeitschrift, 26, 459–473.

Wang, Y., Belluš, M., Ehrlich, A., Mile, M., Pristov, N., Smolikova, P., Španiel, O., Trojakova, A., Brožková, R., Cedlínk, J. and Klaríč, D. (2018) 27 years of Regional Cooperation for Limited Area modelling in Central Europe (RC LACE). Bulletin of the American Meteorological Society, 99, 1415–1432.

Warner, T.T., Peterson, R.A. and Treadon, R.E. (1997) A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction. Bulletin of the American Meteorological Society, 78, 2599–2617.

Wei, M., Toth, Z., Wobus, R. and Zhu, Y. (2008) Initial perturbations based on the ensemble transform (ET) technique in the NCEP global ensemble forecast systems. Tellus, 60A, 62–79.

Wilks, D.S. (1997) Resampling hypothesis tests for autocorrelated fields. Journal of Climate, 10, 65–82.

Wilks, D.S. (2006) Statistical Methods in the Atmospheric Sciences, 2nd edition. Cambridge, MA, USA: Academic Press.

Wu, W.-S., Parrish, D.F., Rogers, E. and Lin, Y. (2017) Regional ensemble–variational data assimilation using global ensemble forecasts. Weather and Forecasting, 32, 83–96.

Xue, M., Kong, F., Weber, D., Thomas, K.W., Wang, Y., Brewster, K., Droegemeier, K.K., Kain, J.S., Weiss, S.J., Bright, D.R., Wandishin, M.S., Coniglio, M.C. and Du, J. (2007) CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA Hazardous Weather Testbed 2007 spring experiment. Preprints. In: 22nd Conf. on Weather Analysis and Forecasting/18th Conf. on Numerical Weather Prediction, Vol. 3B.1. Salt Lake City, UT: Amer. Meteorol. Soc. http://ams.confex.com/ams/pdfpapers/124387.pdf. [Accessed 19 February 2019]

Yang X. 2005. Analysis blending using a spatial filter in gridpoint model coupling. HIRLAM Newsletter, no. 48, Article 10, HIRLAM Programme, de Bilt, Netherlands, 49–55.

Žagar, N., Andersson, E. and Fisher, M. (2005) Balanced tropical data assimilation based on a study of equatorial waves in ECMWF short-range forecast errors. Quarterly Journal of the Royal Meteorological Society, 131, 987–1011.

Zhang, F.Q., Bei, N.F., Rotunno, R., Snyder, C. and Epifanio, C.C. (2007) Mesoscale predictability of moist baroclinic waves: convection-permitting experiments and multistage error growth dynamics. Journal of the Atmospheric Sciences, 64, 3579–3594.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Keresturi E, Wang Y, Meier F, Weidle F, Wittmann C, Atencia A. Improving initial condition perturbations in a convection-permitting ensemble prediction system. Q J R Meteorol Soc., 2019;145:993–1012. https://doi.org/10.1002/qj.3473