A New Method for Generating Initial Condition Perturbations in a Regional Ensemble Prediction System: Blending

YONG WANG,* MARTIN BELLUS,† JEAN-FRANCOIS GELEYN,# XULIN MA,@ WEIHONG TIAN,& AND FLORIAN WEIDLE*

* Department of Forecasting Models, Central Institute for Meteorology and Geodynamics, Vienna, Austria
† NWP Division, Slovak Hydro-meteorological Institute, Bratislava, Slovakia
# CNRM, Météo-France, Toulouse, France
@ College of Atmospheric Science, Nanjing University of Information Science and Technology, Nanjing, China
& NMC, China Meteorological Administration, Beijing, China

(Manuscript received 18 December 2012, in final form 9 January 2014)

ABSTRACT

A blending method for generating initial condition (IC) perturbations in a regional ensemble prediction system is proposed. The blending is to combine the large-scale IC perturbations from a global ensemble prediction system (EPS) with the small-scale IC perturbations from a regional EPS by using a digital filter and the spectral analysis technique. The IC perturbations generated by blending can well represent both large-scale and small-scale uncertainties in the analysis, and are more consistent with the lateral boundary condition (LBC) perturbations provided by global EPS. The blending method is implemented in the regional ensemble system Aire Limitée Adaptation Dynamique Développement International-Limited Area Ensemble Forecasting (ALADIN-LAEF), in which the large-scale IC perturbations are provided by the European Centre for Medium-Range Weather Forecasts (ECMWF-EPS), and the small-scale IC perturbations are generated by breeding in ALADIN-LAEF. Blending is compared with dynamical downscaling and breeding over a 2-month period in summer 2007. The comparison clearly shows impact on the growth of forecast spread if the regional IC perturbations are not consistent with the perturbations coming through LBC provided by the global EPS. Blending can cure the problem largely, and it performs better than dynamical downscaling and breeding.

1. Introduction

Ensemble prediction techniques have been applied in most numerical weather prediction (NWP) centers as a dynamical way of accounting for the forecast uncertainty. The optimal design of an ensemble prediction system (EPS) strongly depends on the quantification of uncertainties due to errors in initial conditions (ICs), model formulation, and physical parameterizations. Additional challenges posed for a skillful regional EPS include, for example, the problem of quantifying the uncertainties due to errors in lateral boundary conditions (LBC).

The construction of the IC perturbation is crucial for a skillful EPS. Previous work has shown significant short-range forecast uncertainties on mesoscale and in local detail (Zhang et al. 2002; McMurdie and Mass 2004; Bowler et al. 2008). The representation of those mesoscale uncertainties in a regional EPS is particularly important for forecasting high-impact weather, quantitative precipitation prediction, low cloud and visibility, wind gusts, etc. For perturbing IC in a regional EPS, there are at least three key requirements:

- The IC perturbations should be effective immediately from the initial time; this means that analysis errors should be quantified in the IC perturbations.
- The spatial scale of the IC perturbations should be in accordance with the scales of variability resolved by the mesoscale regional model.
- The IC perturbations should be consistent with the perturbation coming through the lateral boundary.

Various approaches are designed for dealing with IC uncertainties related with the forecast in regional EPS. The dynamical downscaling (Grimit and Mass 2002; Frogner et al. 2006; Bowler et al. 2008; Marsigli et al. 2011)
in which the IC perturbations of regional EPS are obtained by interpolating the IC perturbations of the global EPS providing the LBC perturbations, is the most popular way for simulating the IC uncertainties. This is because of its simplicity and good performance. The dynamical downscaling is a priori incapable of meeting the second key requirement, because the regional small-scale uncertainties cannot be explicitly simulated with dynamical downscaling, which is usually just following the governance of the global ensemble that is driving it.

An alternative is the breeding technique (Toth and Kalnay 1993). When applied to a regional EPS, it creates the perturbations including, in principle, all scales resolved by the mesoscale regional model, and takes at least partly the analysis uncertainties into account. There are two other improved versions of breeding: the ensemble transform (ET; Bishop and Toth 1999; Wei et al. 2008) and the ensemble transform Kalman filter (ETKF; Bishop et al. 2001; Wang and Bishop 2003). ET and ETKF are more optimal at sampling the analysis uncertainties than the original breeding approach because they take account of the distribution of assimilated observations.

In most regional EPSs the LBC perturbations are obtained by the forecast from a global EPS. There is a concern that because of different perturbation treatment in regional EPS and its global driving EPS, the regional IC perturbations might conflict with perturbations coming from the LBC (Warner et al. 1997; Bowler et al. 2008). It is unclear if an ill-posed setup like that would be superior to the dynamical downscaling.

For example it is very likely, that a regional ensemble using regional breeding IC perturbations coupled to a global EPS using singular vector (SV) perturbations (Buizza and Palmer 1995), would lead to introducing a global EPS using singular vector (SV) perturbations using regional breeding IC perturbations coupled to global EPS. That would be superior to the dynamical downscaling and was introduced by Mylne 2009). It is unclear if an ill-posed setup like that would be superior to the dynamical downscaling.

At Zentral Anstalt für Meteorologie und Geodynamik (ZAMG), the regional EPS Aire Limitée Adaptation Dynamique Développement International-Limited Area Ensemble Forecasting (ALADIN-LAEF; Wang et al. 2010, 2011) has been developed for the purpose of providing a reliable short-range probabilistic forecast to the national weather service of ALADIN-Limited Area modeling for Central Europe (LACE) partners; and to allow the probabilistic information be propagated into downstream models, for example, of hydrology and energy industry.

For ALADIN-LAEF the natural choice for the LBC perturbations are those from European Centre for Medium-Range Weather Forecasts (ECMWF-EPS) forecasts. This is not only because of the similarity in model physics in the ECMWF Integrated Forecast System (IFS) and ALADIN, but also because of the quality of ECMWF-EPS forecasts and their operational availability at ZAMG.

The IC perturbations for ECMWF-EPS are generated using SV technique. This is an appropriate method for medium-range forecasting, but it is still unclear whether it is appropriate for use in regional EPS. Research on regional SVs is in a very early stage (Stappers and Barkmeijer 2011); the design of SVs is surely not optimal for a short-range ensemble, which has to quantify the uncertainties in the analysis. Furthermore, the SV technique is computationally expensive.

In ALADIN-LAEF we employ the breeding technique, as it is simple, inexpensive, and has been successfully applied at NCEP (Du et al. 2003). To deal with the aforementioned inconsistency problem of coupling with ECMWF-EPS, a blending idea for generating regional IC perturbations has been implemented in ALADIN-LAEF. The idea is to use the ALADIN blending technique (Giard 2001; Brůžková et al. 2001, 2006; Derkova and Bellus 2007) to combine the large-scale uncertainties provided by ECMWF-EPS with the small-scale uncertainties generated by breeding with ALADIN model. The ALADIN blending technique is a spectrally based
digital filter. The blended perturbations should have the large-scale perturbations from ECMWF-EPS, and the small-scale part from ALADIN breeding.

We believe that the new perturbations meet the three key requirements for regional IC perturbations. Through ALADIN breeding the perturbations provided by blending attempt to estimate the errors in the initial analysis based on the past information about the flow. On the large scale, the IC perturbations are now consistent with the LBC perturbations, with both of them being based on the ECMWF-EPS perturbations. The small-scale uncertainty in the analysis is more detailed and accurate due to the higher resolution of ALADIN, and it is more in balance with the orographic and surface forcing used in the regional model. This should be a better representation of reality than interpolated large-scale perturbations from the global EPS.

A similar blending scheme was proposed by Caron (2013). He uses a scale-selective ETKF in the context of a convective-scale ensemble, including different possible truncation scales and a demonstration of the impact of inconsistent LBCs.

In this study, we give a detailed description of the blending method for generating IC perturbations in ALADIN-LAEF, and document the performance of blending compared with breeding and dynamical downscaling. Here, we also investigate the impact of the inconsistent coupling of ALADIN breeding IC perturbations with ECMWF-EPS LBC perturbations.

Section 2 introduces the blending method. Section 3 describes the ALADIN-LAEF configuration, breeding, and experimental setup. Section 4 evaluates the results of a 2-month comparison between blending, dynamical downscaling and breeding. A case study is also investigated in section 4. A summary and conclusions follow in section 5.

2. The blending method

The idea of blending is to obtain a new perturbed initial state by combining large-scale features of global ensemble ICs with small-scale features provided by regional ensemble ICs. The underlying hypothesis is that the information on small-scale uncertainties is more reliably represented by the regional ensemble than by the global ensemble, since the small scales are not resolved in the global ensemble.

The blending method is a spectral technique using standard nonrecursive low-pass Dolph–Chebyshev digital filter (Giard 2001; Brožková et al. 2001). The blending procedure is schematically and graphically illustrated in Fig. 1. It consists of several consecutive steps: 1) interpolating model fields of global ensemble ICs on the spectral resolution of the regional ensemble; 2) mapping fields from the spectral resolution of the regional ensemble to a lower spectral resolution, blending truncation, which is predefined by the blending ratio; 3) applying a digital filter to both perturbed ICs of global ensemble and regional ensemble on the original grid of the regional model but at the blending truncation; 4) remapping the fields from blending truncation to the original spectral resolution of the regional ensemble after digital filtering; and 5) computing the difference between those filtered fields, which represents a large-scale increment. This increment contains almost pure low-frequency perturbation information, and is then added to the original high-frequency signal of the perturbed high-resolution regional ensemble ICs. The combination (blending) of both spectra is performed in a spectral transition zone, which is implicitly defined by the use of an incremental digital filter initialization technique (Lynch et al. 1997). The detailed description of blending is given mathematically in below.

Following the idea by Machenhauer and Haugen (1987) and considering perturbed variable $G$ from the global ensemble, and $R$ from the regional ensemble, both $G$ and $R$ being valid at the resolution of the regional model, their full harmonic Fourier expansions are given by

$$G(x) = \sum_{m=-M}^{M} G_m e^{2i\pi(mx/L_x)} \quad \text{and}$$

$$R(x) = \sum_{m=-M}^{M} R_m e^{2i\pi(mx/L_x)}, \quad (1)$$

FIG. 1. Schematic description of spectral blending.
where \( M \) is the maximum wavenumber and \( L_x \) is the horizontal domain length in the \( x \) direction. Let \( J \) be the number of grid points of the whole regional domain along the \( x \) direction, the inverse truncated Fourier transform; that is, the spectral coefficients \( G_m \) and \( R_m \) are obtained by

\[
G_m = \frac{1}{J} \sum_{j=0}^{J-1} G(j) e^{-2\pi i (jm/J)} \quad \text{and} \quad R_m = \frac{1}{J} \sum_{j=0}^{J-1} R(j) e^{-2\pi i (jm/J)}.
\]

Similar equations are applied in the \( y \) direction. As we mentioned above, blending is applied on full gridpoint resolution of the regional ensemble but with a lower spectral resolution (we call it the blending truncation), which represents the scale resolved by IC perturbations of the global ensemble. When changing the spectral truncation from the original resolution to the blending spectral resolution, for both perturbed variables from the global ensemble and the small scales to be kept from the regional ensemble, but being valid at the blending spectral resolution, denoted as \( G_{LB} \) and \( R_{LB} \), their full harmonic Fourier expansions are given by

\[
G_{LB}(x) = \sum_{m=-M_L}^{M_L} G_{LOW}(m)e^{2\pi i (mx/L_x)} \quad \text{and} \quad R_{LB}(x) = \sum_{m=-M_L}^{M_L} R_{LOW}(m)e^{2\pi i (mx/L_x)},
\]

where \( M_L \) is the maximum wavenumber of the blending spectral truncation. The inverse truncated Fourier transform (i.e., the spectral coefficients \( G_{m,LOW}^{LB} \) and \( R_{m,LOW}^{LB} \)) are obtained by

\[
G_{m,LOW}^{LB} = \frac{1}{J} \sum_{j=0}^{J-1} G_{LB}^{j} e^{-2\pi i (jm/J)} \quad \text{and} \quad R_{m,LOW}^{LB} = \frac{1}{J} \sum_{j=0}^{J-1} R_{LB}^{j} e^{-2\pi i (jm/J)}.
\]

The blending spectral truncation is determined by the resolution of the global ensemble, the resolution of initial perturbations of the global ensemble, and the resolution of the regional ensemble. The estimation of the blending spectral truncation (number of equivalent waves resolved around a great circle of the earth) \( T_R^{cut} \) may be obtained as follows:

\[
T_R^{cut} = \sqrt[3]{T_G^{2} \times T_R^{\phi}} \quad \text{and} \quad T_G^{\phi} = \sqrt{T_G^{f} \times T_G^{\phi}}.
\]
a blending large-scale increment (Giard 2001), which is then added on the original regional model spectra of the regional ensemble. The choice of the filter determines the implicit blending interval and weights. Hence, the blending spectral truncation is not used here to simply cut off part of signal. It is used to define the transition zone where the spectral coefficients are progressively damped by the digital filter. The smooth transition between the spectra also means that the method is not very sensitive to the exact choice of the blending spectral truncation, provided that the spectral transition is sufficiently wide in the DFI.

The DFI uses the digital filter applied to time series of the model variables generated by short-range adiabatic backward and diabatic forward integrations from the initial time in order to ensure the appropriate state of balance between the mass and wind initial fields (Lynch and Huang 1992; Lynch et al. 1997).

For any model state \(G_m^{L_{LOW}}\) and \(R_m^{L_{LOW}}\), denoted in the following as \(f_n\), known at the time step \(n\) \(\{f_{-N}, \ldots, f_{-1}, f_0, f_1, \ldots, f_N\}\), it may be regarded as the Fourier coefficients of a function \(F(\theta)\):

\[
F(\theta) = \sum_{n=-N}^{n=N} f_n e^{-i n \theta},
\]

where \(\theta\) is the digital frequency. The filtering of \(f\) could be conducted by multiplying \(F(\theta)\) by a function \(H(\theta)\):

\[
H(\theta) = \begin{cases} 1 & \text{if } |\theta| \leq \theta_c \\ 0 & \text{if } |\theta| \geq \theta_c \end{cases},
\]

where \(\theta_c\) is the cutoff frequency. Let \(f_n^L\) denote the low-frequency part of \(f_n\), clearly

\[
H(\theta) \times F(\theta) = \sum_{n=-N}^{n=N} f_n^L e^{-i n \theta}
\]

and

\[
f_n^L = \sum_{k=-N}^{k=N} (H \times F)_k f_{n-k} = \sum_{k=-N}^{k=N} h_k f_{n-k},
\]

where \(h_k = h_{-k}\) are filter weights defined by Dolph–Chebyshev polynomials. We apply

\[
H(\theta) = h_0 + 2 \sum_{k=1}^{N} h_k \cos(k \theta \Delta t)
\]

in the DFI, where \(\Delta t\) is the time step and \(2N + 1\) is the order of the filter. The detailed description of DFI can be found in Lynch and Huang (1992) and in Lynch et al. (1997). Since the “ideal low-pass” filter can produce unpleasant oscillations in the stop band and also some damping in the pass band (the attenuation of amplitudes of the long waves), the weights of low-pass filter are usually modulated by a “window.” This trick can significantly reduce the Gibbs oscillations in the stop band and at the same time minimizes waves dumping in the pass band, too (Lynch et al. 1997).

Although DFI is by definition a time filter, it acts like a space filter as well. It is because of the fact that in meteorology, where a wave’s propagation speed is limited, high frequencies are usually associated with horizontally short waves. As a consequence, the properties of the DFI can be used besides its original purpose—the data initialization, also for the scale selection in the spectral blending by digital filter technique (Derková and Belluš 2007). One could apply a cheap low-pass filter directly in space as in Caron (2013), as it was tried in ALADIN some years ago. But the disadvantage of such a direct blending is an unknown degree of balance between the spectral fields, and a very delicate choice of both transition zone and the weights (Giard 2001). An advantage of DFI blending is that it could provide a slight local adaptation to atmospheric conditions through spectral mapping and remapping. In case of rapidly changing situations (in principle mostly linked to smaller scales) more confidence will be given to the regional ICs, which is supposedly more accurate than the interpolated global ICs.

The solution of the model, integrated from \(-t_N\) to \(t_N\), is weighted averaged:

\[
f^L(0) = \sum_{n=-N}^{N} h_n f_n,
\]

so that at the end of the DFI a balanced initial state is achieved. Supposing that the results of DFI on \(G_m^{L_{LOW}}\) and \(R_m^{L_{LOW}}\) are \(G_m^{L_{LOW}}\) and \(R_m^{L_{LOW}}\), respectively, we obtain the large-scale part of the perturbed ICs of global and regional model by

\[
G^{L_{SP}}(x) = \sum_{m=-M}^{M} G_m^{L_{LOW}} e^{2i\pi(mx/L_x)} \quad \text{and}
\]

\[
R^{L_{SP}}(x) = \sum_{m=-M}^{M} R_m^{L_{LOW}} e^{2i\pi(mx/L_x)}.
\]

The symbolic equation of blending can be summarized after Brožková et al. (2006) and Derkova and Bellus (2007):

\[
\text{IC}_{\text{blend}} = \frac{R - R^{L_{SP}}}{L_{\text{small}}^L} + \frac{G^{L_{SP}}}{L_{\text{large}}^L},
\]
where IC\textsubscript{blend} denotes initial condition after blending. After blending, the field itself keeps the large scale from global IC perturbations, and the small scale from regional ensemble IC perturbations with a smooth and fully model compatible transition between them both.

To demonstrate the effect of the blending, we compute the kinetic energy spectra for perturbed ICs of global ensemble, regional ensemble and their blending. Figure 2 shows an example of kinetic energy spectra for ICs of the first ECMWF-EPS member (solid line in blue), the first ALADIN breeding member (solid line in green), and the first blending member (dotted line in red) at model level 31 (around 250 hPa) valid at 0000 UTC 10 Aug 2007. The kinetic energy spectra are averaged over the whole ALADIN-LAEF domain.

**Figure 2.** An example of kinetic energy spectra for perturbed ICs of the first ECMWF-EPS member (solid line in blue), the first ALADIN breeding member (solid line in green), and the first blending member (dotted line in red) at model level 31 (around 250 hPa) valid at 0000 UTC 10 Aug 2007. The kinetic energy spectra are averaged over the whole ALADIN-LAEF domain.

3. Experimental setup

a. Configuration of ALADIN-LAEF

ALADIN-Austria is the operational limited-area model run at ZAMG (Wang et al. 2006). In ALADIN-LAEF, we use ALADIN-Austria with a horizontal resolution of 18 km and 37 vertical levels. ALADIN-Austria is based on hydrostatic equations and uses a hybrid vertical coordinate, a spectral method with bi-periodic extension of the domain using elliptical truncation of the double-Fourier series, a two-time level semi-Lagrangian advection scheme, a semi-implicit time stepping, fourth-order horizontal diffusion, and Davies–Kalberg-type relaxation and DFI.

The ALADIN-LAEF integration domain for this study is shown in Fig. 4. It covers all of Europe and a large part of the Atlantic. ALADIN-LAEF comprises 17 ensemble members, of which the first 16 members are perturbed. Their large-scale IC perturbations used in blending and LBC perturbations are provided by the first 16 members of the ECMWF-EPS (Leutbecher and Palmer 2008). The 17th ALADIN-LAEF member contains ICs and LBCs from the ECMWF-EPS control forecast with resolution of T399L62.

b. ALADIN breeding

Breeding (breeding of growing vectors) was proposed by Toth and Kalnay (1993). It is to simulate how the growing errors are “bred” and maintained in a conventional analysis cycle by using short-range forecasts successively. The bred vectors should thus provide a good estimate of possible growing error in the analysis (Toth and Kalnay 1993, 1997).

ALADIN breeding is implemented as follows: (i) adding a random perturbation to the control analysis for the very first ensemble, this should be done only once; (ii) integrating the model ALADIN-Austria with control analysis and the perturbed ICs; (iii) building the difference between the two 12-h forecasts; and (iv) scaling down the forecast difference in amplitude to the size of the perturbation. We use a constant rescaling. The rescaling factor is computed empirically by using the root-mean-square error of the 12-h temperature forecast near 850 hPa (J. Du 2006, personal communication). (v) Adding/subtracting the rescaled difference and their blending. The blended ICs apparently have the large-scale perturbations from the global ensemble, and the small-scale uncertainty provided by the regional ensemble, which cannot be found in the global EPS. As a consequence, the blended ICs should be more consistent with the LBCs provided by the driving global EPS.
to the new control analysis. The perturbed ICs are generated in sets of positive and negative pairs around the control analysis. Steps (ii) to (v) are then repeated, and the perturbations are bred to grow along the forecast trajectory. After a few cycles, the breeding vector structure should be statistically independent of the very first arbitrary perturbations (Magnusson et al. 2008). In ALADIN breeding, wind components, temperature, moisture, and surface pressure are perturbed at each level and model grid point.

c. Experiments

To evaluate the blending, comparisons with dynamical downscaling of ECMWF-EPS and ALADIN breeding have been conducted. In Table 1 the setup of the comparison is described. In the experiment, we did not apply any other perturbations, for example, the land surface perturbation and the multiphysics for the model perturbation in ALADIN-LAEP. This makes possible to have a clean comparison between blending, dynamical downscaling and breeding.

In our experiments, the resolution of ECMWF-EPS is T1399L62 (spectral triangular T399 with 62 vertical levels, corresponding to 50-km resolution). Initial conditions are perturbed using a combination of initial time and evolved singular vectors (Buizza and Palmer 1995) computed at T42L62 resolution, with a 48-h optimization time interval and a total-energy norm. Singular vectors are computed to maximize the final-time norm over different areas (Barkmeuer et al. 1999), then
combined and scaled to have an initial amplitude comparable to an estimate of the analysis error. Model uncertainties due to physical parameterizations are simulated using a stochastic scheme (Buizza et al. 1999). The ensemble control analysis is obtained by interpolating the \(T_1,799L91\) analysis to the ensemble \(T_1,399L62\) resolution.

For all those three experiments, the LBC perturbations are provided by ECMWF-EPS forecasts in a 6-h interval.

The blending implementation in this study is practically a breeding-blending cycle. Based upon the ALADIN-LAEF 12-h forecasts, ALADIN breeding is used to generate the IC perturbations, which are then blended with the corresponding 16 perturbed ICs of ECMWF-EPS. The new blended IC perturbations for each of the 16 ALADIN-LAEF members are used as ICs for the ALADIN-LAEF forecast; the ALADIN-LAEF 12-h forecasts are employed for the next breeding-blending cycle. In short, the blending system takes its small scales from a separate breeding cycle based on the ALADIN-LAEF 12-h forecasts started with the blended IC perturbations.

In our implementation of blending, the spectral resolution of ECMWF ensemble \(T^f_G\) is equal to 399; and the spectral resolution of its IC perturbations (i.e., the resolution of ECMWF SV \(T^p_G = 42\)); the ECMWF equivalent spectral resolution for ALADIN-LAEF \(T^f_R\) is \(L/(3 \times 18\text{ km}) = 741\), where \(L = 40\text{000 km}\) is the Earth’s perimeter and 3 represents the denominator for ALADIN-LAEF quadratic grid. Therefore, the blending spectral truncation is computed as

\[
T^f_G = \sqrt{T^f_G \times T^p_G} \simeq 129; \quad T^\text{cut}_R = \sqrt{\frac{T^2_G}{T^f_R}} \simeq 231,
\]

and the blending ratio \(T^f_R/T^\text{cut}_R\) is about 3.2. In ALADIN-LAEF the original spectral resolution, the maximum zonal and meridional wavenumbers are 107 and 74, the corresponding blending spectral truncation implemented in this study are then 33 and 23, respectively. In the implementation of DFI, we used the standard Dolph–Chebyshev filter. The DFI is applied by integrating ALADIN backward adiabatically and forward diabatically. The time span of the filter is 4 h with eight integration time steps in each direction.

4. Results

In this study the ALADIN-LAEF runs with blending, breeding, and dynamical downscaling have been initialized at 0000 UTC and integrated to 54 h for a 2-month period (15 June–20 August 2007). The forecasts of upper-air weather variables were verified against ECMWF analysis, and both analysis and forecast are interpolated to a common regular 0.15° × 0.15° latitude–longitude grid. Verification of the forecasts of surface weather variables were performed against the surface observations at the observation location. Forecast values are interpolated to the observation site for smoothly varying fields, such as 2-m temperature and 10-m wind speed. For precipitation, which has strong spatial gradients, the observation is matched to the nearest grid point. The verification is performed for a limited area of the forecast domain over central Europe, as shown in Fig. 4. There are 1219 synop stations in the verification domain being used in this study.

A set of standard ensemble and probabilistic forecast verification scores is applied to evaluate the performance of the ALADIN-LAEF experiments. The scores considered are ensemble spread, root-mean-square error (RMSE) of ensemble mean, continuous ranked
Those verification scores measure the quality of probabilistic forecasts of scalar quantities. A detailed description of those verification scores can be found in Hamill and Colucci (1997), Mason (1982), Jolliffe and Stephenson (2003), Stanski et al. (1989), and Wilks (2006).

In this study, we focus on verification of the most concerned variables in regional EPS: 2-m temperature (T2m), 10-m wind speed (W10m), 12-h accumulated rainfall (PREC), and three representative upper-air variables: temperature at 850 hPa (T850), wind speed at 850 hPa (V850), and relative humidity at 850 hPa (RH850).

a. Ensemble mean RMSE and spread

RMSE of the ensemble mean describes the deterministic skill of the ensemble and gives good measures of overall forecast performance. The ensemble spread is the standard deviation of the ensemble members with respect to its mean. It reflects the diversity of the ensemble forecast. The RMSE should basically match the spread if the ensemble is correctly predicting the distribution of potential outcomes.

Figures 5a–c show the RMSE of ensemble mean and ensemble spread of T850, V850, and RH850 for ALADIN-LAEF with blending (dotted), breeding (dashed), and downscaling (solid) as function of forecast range. The RMSE of the ensemble mean forecasts is very similar between blending, breeding, and downscaling for temperature, wind, and relative humidity at 850 hPa for the whole forecast range. The spread of blending, breeding, and downscaling for all those three weather variables are less than the error of the ensemble mean, which means that the ensemble forecasts are underdispersive. Among the three ensembles, downscaling of ECMWF SV based perturbation shows the largest growth of spread. Since all the three ensembles use the same LBCs of ECMWF-EPS forecasts, one may speculate that it might be because of the greater consistency between the IC perturbations in downscaling with the perturbed LBC perturbations from ECMWF-EPS forecasts. As seen from Figs. 5a–c, the magnitude of the spread for downscaling is generally remarkable smaller than for breeding and blending in the early forecast range, in other words, more underdispersive.
Very probably this is related to the lack of small scales in the downscaled ECMWF initial perturbations.

The spread for breeding performs similar to blending and downsampling in the early forecast range (up to 18 h) for T850, but in the later forecast range the spread for breeding grows clearly slower than for blending and downsampling, in which the spread of blending and downsampling keeps growing almost equally. For V850 and RH850 breeding has much larger spread than downsampling up to 24–30 forecast hours. In the later forecast range the spread of breeding grows slower and becomes smaller than downsampling and blending.

As expected, blending has the best overall performance. It takes the advantage of breeding and downsampling. In the early forecast range, the spread for blending is similar (T850 and V850) or very closed (RH850) to breeding, which is considerably better than downsampling. This competitive performance of blending shows the dominated effect of breeding in the blending perturbation. During the later forecast range the spread of blending keeps increasing rapidly as the downsampling, which is superior to breeding. This might explain that the blending of large-scale uncertainty from its driving global model in regional IC perturbations makes it more consistent with the LBC perturbations coming from the driving global EPS.

The performance of blending, breeding, and downsampling for the surface weather variables has been also compared in Figs. 5d–f, which show the RMSE of the ensemble mean and the ensemble spread of T2m, W10m, and PREC for ALADIN-LAEF with blending, breeding, and downsampling. The RMSE of blending, breeding, and downsampling performs very similar for T2m and W10m; a slightly better RMSE of blending and breeding than downsampling can be observed for PREC in the early forecast hours (up to 30 h), then they reach the RMSE of downsampling and give the same result in the late forecast hours (30–54 h).

Like the upper-air variables, the spread of all the surface weather variables, T2m, W10m, and PREC for blending, breeding, and downsampling are remarkably underdispersive. The positive impact of blending can also be seen in the growth of ensemble spread for the surface variables. Blending is close to breeding, which performs better than downsampling in the first 24 forecast hours. In the late forecast period downsampling is superior; this could be due to the matched IC perturbations of downsampling with the LBC perturbations from ECMWF-EPS forecasts. Blending is similar to downsampling, as it has the same large-scale perturbations as in downsampling, which are more consistent with the LBC perturbations from ECMWF-EPS forecasts. Breeding has a larger spread in the early forecast range, but the growth of spread is slower than blending and downsampling.

### b. Outlier statistics

The measure of the statistical reliability discussed in this subsection is the percentage of outliers. This is the statistic of the number of cases when the verifying analysis at any grid point lies outside the whole ensemble. A more reliable EPS should have a score closer to \(2/(n_{\text{ens}} + 1)\), where \(n_{\text{ens}}\) is the ensemble size.

Figures 6a–c give the outlier statistics for the forecasts of temperature, wind speed, and relative humidity at 850 hPa for blending, breeding, and downsampling; while Figs. 6d–f show the outlier statistics for the forecasts of 2-m temperature, 10-m wind speed, and 12-h accumulated precipitation.

The benefit of using blending can be found in the outlier statistics for both upper-air and surface weather variables. For T850 and V850 in Figs. 6a and 6b, blending gets the best reliability in terms of outlier statistics. Superior results for blending can be found for W10m and PREC in Figs. 6d and 6f. Blending is close to breeding for RH850 in Fig. 6c and for T2m in Fig. 6d at a short lead time, they have less outliers than downsampling; and it turns too close to downsampling at the later time, and they have less outliers than breeding.

The result of outlier statistics and the result of the ensemble spread in Fig. 5 are consistent with each other for blending, breeding, and downsampling.

### c. The continuous ranked probability score

The CRPS is an overall measure of the skill of a probabilistic forecast, measuring the skill of the ensemble mean forecast as well as the ability of the perturbations to capture the deviations around it (Bowler and Mylne 2009). CRPS is the generalized form of the discrete ranked probability score, simulating the mean over all possible thresholds. As noted by Hersbach (2000), CRPS is analogous to an integrated form of the Brier score, which can be decomposed into reliability, resolution, and uncertainty. The CRPS orientates negatively, so smaller values are better; and it rewards concentration of probability around the step function located at the observed value. A perfect CRPS score is zero, as with the Brier score (Wilks 2006).

Figure 7 shows CRPSs of upper-air variables T850, V850, and RH850 for blending, breeding, and downsampling; and of surface variables T2m, W10m, and PREC, respectively. The statistical significant tests for CRPS are computed by using the bootstrap method (Wilks 2006). The 95% and 5% confidence intervals for the three experiments and all the verifying upper-air and surface variables are also shown in Fig. 7. The differences in the
skill score are statistically not significant, but noticeable. CRPS score supports the findings with spread and outliers in the previous subsections. The blending has the best performance, while breeding is better only at the short lead time.

d. A case study

The impact of blending is evaluated by using several standard statistical scores over a 2-month period in the last subsections. To better understand the blending method and its outperformance over the downscaling and the regional breeding, a case study is investigated.

Figure 8 shows the time evolution of ensemble spread of T850 of ALADIN-LAEF with the blending, the breeding and the downscaling from initial time to 48-h forecasts in a 12-h interval. The ensemble spread was calculated on each grid over the whole ALADIN-LAEF domain. It is a widely used measure to describe the perturbation in the ensemble forecast. The blending has more local small-scale information than downscaling at initial time as expected; these scales cannot be properly resolved by the downscaling. The ensemble spread of blending has a similar distribution to the breeding, but is smaller. The blending spread grows faster than the breeding and becomes close to downscaling in the later forecast range. The larger spread in the breeding at initial time keeps the same magnitude and does not grow during the forecast hours. It is noted that the local small-scale perturbation in the blending and the breeding decays after 12 forecast hours. The difference of perturbations in the blending and the breeding at the large scale keeps visible in the earlier and later forecast range. These results are consistent with the result of the 2 months of statistical scores discussed previously.

The ensemble spread of surface pressure has been also investigated. They show the similar results (not shown).

Figure 9 shows the time evolution of kinetic energy spectra for perturbed ICs and forecasts of the ALADIN-LAEF member with downscaling, breeding, and blending at around 700 hPa. The kinetic energy spectra are averaged over the whole ALADIN-LAEF domain. Such a spectral analysis at different lead times for this selected case can help us to understand “how” different
perturbations evolve with time and, therefore, help to explain their forecast skills.

The effect of the blending is obviously at the initial time and in the earlier forecast range. After 12–18-h forecasts the kinetic energy spectra of blending, breeding, and downscaling at the small scale get closer to each other. The spectra of the blending remains the same as downscaling, while the spectra of the breeding differs from the blending and downscaling at large scale clearly in the earlier forecast range, and still keeps visible in the later forecast range. This confirms the results discussed before.

5. Summary and conclusions

In this paper a new blending method for generating IC perturbations in regional EPS has been proposed and described. Blending combines the large-scale IC perturbations from the global EPS with the small-scale IC perturbations from the regional EPS by using the digital filter and spectral analysis techniques. The blending method is implemented in ALADIN-LAEF, in which the large-scale IC perturbations are provided by ECMWF-EPS, and the small-scale IC perturbations are generated by breeding in ALADIN-LAEF. Ensemble forecasts with blending are compared with ensembles with regional breeding and dynamical downscaling of ECMWF-EPS. Results are evaluated by using some standard verification scores over central Europe and for a 2-month summer period in 2007. A case study is also conducted to have an in-depth analysis of the results. We verified the surface weather variables: 2-m temperature, 10-m wind, and 12-h accumulated rainfall; and the upper-air weather variables: wind, temperature, and relative humidity at 850 hPa.

The IC perturbations generated by blending implemented in ALADIN-LAEF can provide a better estimate of the actual errors in the initial analysis based on the past information about the flow through breeding. The small-scale uncertainty in the analysis is more detailed and accurate due to the higher resolution of ALADIN, and it is more in balance with the orographic
FIG. 8. Ensemble spread of T850 of ALADIN-LAEF with blending (blend), breeding (breed), and downscaling (down) from initial time to 48-h forecasts in a 12-h interval. The forecast started at 1200 UTC 23 May 2011.
and surface forcing that is used in the regional model. This is a better representation of reality than the interpolated large-scale perturbations from the global model, for example, through dynamical downscaling, since these scales are not resolved in the global model. On the large scale, the regional IC perturbations generated by blending are the uncertainties from ECMWF-EPS, which can better sample the large-scale feature.

Fig. 9. Kinetic energy spectra for perturbed ICs and forecasts of first member of ALADIN-LAEF with downscaling (down, solid line in blue), breeding (breed, solid line in green), and blending (blend, dashed line in red) at model level 20 (around 700 hPa). The kinetic energy spectra are averaged over the whole ALADIN-LAEF domain. The forecast started at 1200 UTC 23 May 2011.
because of its global geographical extent. As the LBC perturbations are provided by ECMWF-EPS forecasts, the IC perturbations of blending on the larger scale are now consistent with the LBC perturbations.

Those features in the IC perturbations of blending have been confirmed by the experiments presented in this study, and it leads to the conclusion that blending has the best overall performance in comparison with breeding and downscaling.

The RMSE of the ensemble mean for blending, breeding, and downscaling is very close to each other. All three ensembles are underdispersive; their spread is less than the errors of their ensemble mean, this is even more evident for the surface variables. One may speculate that the strong underdispersion for the surface weather variables may be due to lack of the perturbations in the model physics and land surface conditions. It should be also noted that the little spread might be due to some extent to observation error (including the error of representativeness, which may be particularly large for station observations of surface variables) rather than a deficiency in the ensemble (Saetra et al. 2004).

Downscaling underperforms at the short lead time. The ensemble spread, statistical reliability, and the skill score of CRPS for downscaling is noticeable inferior to breeding and blending. Very probably this is related to the lack of small scales in the downscaled ECMWF initial perturbations. Downscaling of ECMWF SV-based perturbation shows the largest growth rate of spread among the three ensembles, and in the later forecast period (after 24 h roughly) performs well, which is comparable to blending. The relative rapid growth of spread for downscaling might correspond to the greater consistency between the IC perturbations in downscaling with the LBC perturbations from the ECMWF-EPS forecast, since all three ensembles use the same perturbed LBCs from ECMWF-EPS forecasts.

Breeding is superior during the first 24 forecast hours, it has larger spread, which is better matched RMSE, has fewer outliers, and slightly higher skill. A plausible explanation for less underdispersion of breeding at the short lead time may be due to the more detailed small-scale IC perturbations generated by the regional ALADIN breeding. In the later forecast range the spread of breeding grows slower and becomes smaller than downscaling and blending. This would be speculated to be due to the mismatch between the IC perturbations generated by breeding with the large-scale forcing through LBC perturbations from ECMWF-EPS forecasts, and/or the properties of breeding method itself (Wang and Bishop 2003). However, more in-depth studies are needed to verify if it is really true or not. This is left to be our future research to better understand the mechanism of blending.

One would expect that breeding performs better than the downscaling, since it is attempting to simulate the errors in the regional forecast. This would be true for the short lead times since the IC perturbations are important in this period; as the influence of the IC perturbations decrease with increasing leading time, the impact of the mismatch between IC and LBC perturbations becomes more dominant. This would suggest that regional ensemble is best if it considers not only the uncertainties in the regional analysis, but also the consistency between the IC and LBC perturbations.

Blending has the best overall performance. In the early forecast hours, blending is similar or much closer to breeding, which is considerably better than downscaling. This competitive performance of blending shows that the blending perturbations, like breeding, contain more detail on the small scale than downscaling, which persists to 24 h roughly. During the later forecast range the spread of blending keeps increase rapidly as the downscaling, which is superior to breeding. This is due to the impact of the LBC perturbations. Blending has the same large-scale perturbations as in downscaling, which are consistent with the LBC perturbations.

As expected from the design of the blending method, blending ensures that the scale of the IC perturbations matches the scale resolved by the regional model, and ensures that the IC perturbations match the LBC perturbations. It makes it possible that blending has the best overall performance.

The case study confirms the results from the 2-month averaged verification. It helps to better comprehend how the blending works. The time evolution of the ensemble spread and the kinetic energy spectra of the blending, breeding, and downscaling prove that the effect of the blending is at the initial time and in the earlier forecast range. The blending becomes close to the downscaling at all the scales with the time, while the breeding differs from the blending and downscaling at the large scale clearly in the earlier forecast range, and keeps visible in the later forecast range.

It has been noted that there are some issues that are not investigated in this study, for example, what signal in IC perturbations generated by breeding and downscaling is good for the error growth in the ensemble, why does the mismatch between IC and LBC cause the slower growth rate of spread in the ensemble in the later forecast hours, and what is the effect of spinup process on IC perturbations for blending, breeding, and downscaling? These questions will be addressed in future studies to better understand the mechanism of the blending.

Furthermore, we are going to apply the blending idea in convection-allowing ensembles. For such an ensemble,
it would need to have its own initial perturbations, which represent the uncertainties in the corresponding resolution. Further, because the model domain of such an ensemble is usually rather small, the possible impact of the LBC perturbations and the mismatch between IC and LBC perturbations might be more significant. These possibilities will be investigated in the future, too.

Acknowledgments. We gratefully acknowledge Dominique Giard, Radmila Brozkova, and Dijana Klaric who have contributed to the ALADIN spectral blending. Special thanks to Jun Du, Craig Bishop, Xuguang Wang, and Zoltan Toth for valuable discussions during the ALADIN-LAEF implementation. ECMWF has provided the computer facilities and technical help implementing ALADIN-LAEF on the ECMWF HPCF. The work has been partly funded by RC LACE and ÖAD (Austrian Agency for International Cooperation in Education and Research, Project CN 13/2010).

REFERENCES


Machenhauer, B., and J. E. Haugen, 1987: Test of a spectral limited area shallow water model with time dependent lateral boundary conditions and combined normal mode/semi-Lagrangian time integration schemes. Proc. ECMWF Workshop on Techniques for Horizontal Discretization in Numerical Prediction Models, Reading, United Kingdom, ECMWF, 361–377.

Magnusson, L., M. Leutbecher, and E. Källen, 2008: Comparison between singular vectors and breeding vectors as initial


