RMetS
Royal Meteorological Society

The Central European limited-area ensemble forecasting system: ALADIN-LAEF

Yong Wang,[a]* Martin Bellus,[b] Christoph Wittmann,[a] Martin Steinheimer,[a] Florian Weidle,[a] Alexander Kann,[a] Stjepan Ivatek-Šahdan,[c] Weihong Tian,[d] Xulin Ma,[e] Simona Tascu[f] and Eric Bazile[g]

[a]*Department of forecasting models, Central Institute for Meteorology and Geodynamics, Vienna, Austria*
[b]*NWP division, Slovak Hydro-meteorological Institute, Bratislava, Slovakia*
[c]*NWP section, Croatian Meteorological and Hydrological Service, Zagreb, Croatia*
[d]*NMC, China Meteorological Administration, Beijing, China*
[e]*Department of atmospheric science, Nanjing University of Information Science and Technology, Nanjing, China*
[f]*NWP section, National Meteorological Administration of Romania, Bucharest, Romania*
[g]*GMAP/CNRM, Météo-France, Toulouse, France*
*Correspondence to: Y. Wang, Zentralanstalt für Meteorologie und Geodynamik, Hohe Warte 38, A-1190 Wien, Austria.
E-mail: yong.wang@zamg.ac.at

The Central European limited-area ensemble forecasting system ALADIN-LAEF (Aire Limitée Adaptation Dynamique Développement InterNational – Limited-Area Ensemble Forecasting) has been developed within the framework of ALADIN international cooperation and the Regional Cooperation for Limited-Area modelling in Central Europe (RC LACE). It was put into pre-operation in March 2007. The main feature of the pre-operational ALADIN-LAEF was the dynamical downscaling of the global ensemble forecast from the European Centre for Medium-range Weather Forecasts (ECMWF). In 2009, ALADIN-LAEF was upgraded with several methods for dealing with the forecast uncertainties to improve the forecast quality. These are: (1) the blending method, which combines the large-scale uncertainty generated by ECMWF singular vectors with the small-scale perturbations resolved by ALADIN breeding into atmospheric initial condition perturbations; (2) the multi-physics approach, wherein different physics schemes are used for different forecast members to account for model uncertainties; and (3) the non-cycling surface breeding technique, which generates surface initial condition perturbations.

This article illustrates the technical details of the updated ALADIN-LAEF and investigates its performance. Detailed verification of the upgraded ALADIN-LAEF and a comparison with its first implementation (dynamical downscaling of ECMWF ensemble forecasts) are presented for a two-month period in summer 2007. The results show better performance and skill for the upgraded system due to the better representation of forecast uncertainties. Copyright © 2011 Royal Meteorological Society

Key Words: ensemble prediction system; probabilistic forecasts

Received 31 May 2010; Revised 25 November 2010; Accepted 26 November 2010; Published online in Wiley Online Library 11 February 2011

Citation: Wang Y, Bellus M, Wittmann C, Steinheimer M, Weidle F, Kann A, Ivatek-Šahdan S, Tian W, Ma X, Tascu S, Bazile E. 2011. The Central European limited-area ensemble forecasting system: ALADIN-LAEF. *Q. J. R. Meteorol. Soc.* **137**: 483–502. DOI:10.1002/qj.751

1. Introduction

Over recent years, limited-area model ensemble prediction systems (LAMEPSs) have become more important as scientific tools for improving the prediction of high-impact weather (especially mesoscale short-range probabilistic predictions), for identifying model error sources, and for developing methods for reducing weather forecast errors.

Several LAMEPSs have been developed in the last few years (Du *et al.*, 2003; Chen *et al.*, 2005; Frogner *et al.*, 2006; Garcia-Moya *et al.*, 2010; Bowler *et al.*, 2008; Li *et al.*, 2008; Marsigli *et al.*, 2008); various approaches are employed for dealing with the uncertainties related to limited-area numerical forecasts in these ensemble forecasting systems. There are four main sources of uncertainty which should be adequately tackled in a LAMEPS:

- Uncertainties due to errors in initial conditions (IC). To quantify these uncertainties, some strategies are designed for generating the perturbation to the IC. Breeding (Toth and Kalnay, 1993) is used by short-range ensemble forecasting (SREF; Hamill and Colucci 1997; Stensrud *et al.*, 2000; Du and Tracton 2001; Du *et al.*, 2003) at the National Centers for Environmental Prediction (NCEP). A very popular method for the perturbation of IC is the dynamical downscaling of global Ensemble Prediction Systems (EPSs). The Met Office Global and Regional EPS (MOGREPS; Bowler *et al.*, 2008) downscales dynamically its global EPS counterpart based on the Ensemble Transform Kalman Filter (ETKF; Wang and Bishop, 2003) method. The IC perturbation of LAMEPS at the Norwegian Meteorological Institute (Frogner *et al.*, 2006) is provided from a version of the ECMWF EPS with dry targeted singular vectors (SVs) over northwestern Europe. In a similar way, the regional EPS at the Meteorological Service of Canada (MSC; Li *et al.*, 2008) applies the perturbations from the MSC global EPS with moist targeted SVs. The COnsortium for Small scale MOdeling Limited-area Ensemble Prediction System (COSMO-LEPS; Molteni *et al.*, 2001; Marsigli *et al.*, 2008) follows a strategy of representative members to downscale the ECMWF EPS, in which the representative members are chosen from clusters of ECMWF EPS forecasts. Other methods for generation of initial uncertainty are used as well, such as employing analyses from different forecast centres, like multi-analysis short-range EPS (SREPS) at the Spanish Met Service, INM (Garcia-Moya *et al.*, 2010), poor man's EPS (PEPS) at the German Weather Service (DWD; Denhard, pers. comm.), the University of Washington Mesoscale Ensemble (Eckel and Mass, 2005), and the approach of random initial conditions perturbation followed by Du *et al.* (1997) and Stensrud *et al.* (2000). Chen *et al.* (2005) proposed a method for initial perturbations using a Different Physical Mode Method (DPMM) for dealing with the initial uncertainties on different scales. Some research efforts on ETKF/ET (Ensemble Transform; Bishop and Toth, 1999) and convective available potential energy (CAPE) SVs in the LAM context have been devoted by Wang *et al.* (2006b), Stappers and Barkmeijer (2008), and Bowler and Mylne (2009). Bishop *et al.* (2009) have introduced an ET method for generating high-resolution initial perturbations for regional ensemble forecasts.
- Uncertainties due to errors in lateral boundary conditions (LBCs) related to the coupling with its global counterpart. Most LAMEPSs are coupled with global EPS to account for the uncertainties in LBCs. The SREPS at INM and PEPS at DWD obtain the LBC perturbations by using deterministic global forecasts from different NWP centres.
- Uncertainties arising from the description of surface conditions and corresponding physical processes in the model. There are only a few LAMEPSs which take the surface uncertainties into account, especially the surface initial conditions. The application of different values of empirical parameters in the model surface schemes is used in COSMO-DE-EPS (Gebhardt *et al.*, 2008). NCEP uses the different surface analyses and surface physics schemes for representing the surface uncertainties in its SREF system. In some global EPSs, surface conditions like sea surface temperature (SST) have been perturbed (e.g. in ECMWF-EPS, SST perturbations have been generated by using different SST analyses based on different analysis schemes), using various observational data (Leutbecher and Palmer, 2008). In MSC-EPS, surface perturbations are introduced through modifications (Gaussian perturbations) of the SST, the albedo and the roughness-length fields (Houtekamer *et al.*, 2007).
- Uncertainties caused by approximation in model formulation and physical parametrization. Multi-physics, multi-model, perturbing the empirical parameters in the physical parametrizations, and stochastic physics are the popular methods for representing model uncertainties in all the LAMEPS systems, e.g. multi-model methods in SREF at NCEP, SREPS at INM, PEPS at DWD, and multi-physics in SREF at NCEP and COSMO-SREPS. In regional EPS at MSC, model physics uncertainties are partly accounted for by stochastically perturbing the parameters in the parametrization schemes; the perturbations are obtained from first-order Markov processes (Li *et al.*, 2008). In MOGREPS, two stochastic approaches are included to represent the model uncertainties: the 'random parameters' and the 'stochastic convective vorticity' scheme (Bowler *et al.*, 2008).

In current operational LAMEPSs, some approaches are employed to address the initial uncertainties, like breeding or targeted SVs (Du *et al.*, 2003; Frogner *et al.*, 2006), but the dynamical downscaling of global EPS perturbations or of analyses from different forecast centres remain the most attractive way for simulating the initial uncertainties, mainly because of simplicity and good performance. Bowler and Mylne (2009) indicate that the dynamical downscaling is slightly more skilful than the regional perturbation ensemble with ETKF; similar conclusions on targeted SVs have been given by Charron (2008). However, one has to keep in mind that the regional small-scale uncertainty cannot be explicitly simulated in a LAMEPS with dynamical downscaling, which is usually just following the governance of the global ensemble that is driving it. As is known, there are significant forecast uncertainties at the mesoscale short range and in local details. The representation of those mesoscale uncertainties in a LAMEPS is particularly useful for the forecasting of high-impact weather, quantitative precipitation prediction, low cloud and visibility, wind gusts, etc.

Unreliability and underdispersion are well-known problems in the field of ensemble forecasting (Hamill and Colucci, 1997). This is especially true for the surface variables such as 2 m temperature and precipitation, rather than

for the atmospheric variables (Buizza *et al.*, 2000; Mullen and Buizza, 2001). One of the reasons is the lack (or inadequate representation) of uncertainties in the surface initial conditions. In many LAMEPSs that are being developed, there are hardly any methods for perturbing the initial state of the land surface.

One further issue is the possible inconsistency due to different perturbation treatments in the LAMEPS and its coupling with the global EPS. The LAM perturbation might conflict with perturbations coming from the LBCs, introducing instability or spurious gravity waves. Hence, the introduction of LAM native perturbations (ensuring that the consistency problem is addressed), and the perturbation of the surface initial conditions, are of great interest for the design of the LAMEPS.

At ZAMG (ZentralAnstalt für Meteorologie und Geodynamik, Vienna), the Central European Limited-Area Ensemble Forecasting system ALADIN-LAEF has been developed within the framework of the international cooperative ALADIN/LACE project. The purpose is to provide reliable short-range probabilistic forecasts to the national weather services of ALADIN/LACE partners, and to allow the probabilistic information be propagated into downstream models, e.g. those of hydrology and the energy industry. ALADIN-LAEF has been run at ECMWF in pre-operation mode since March 2007. The dynamical downscaling of ECMWF-EPS (Buizza and Palmer, 1995; Leutbecher and Palmer, 2008) was the main feature of ALADIN-LAEF up to 2009, in which their initial condition perturbations and lateral boundary perturbations were provided by the perturbed ECMWF-EPS members. The model uncertainty and the surface uncertainty were not taken into account in the pre-operational ALADIN-LAEF.

Some research studies suggest that the quality of short-range probabilistic forecasts, in particular for high-impact weather, can be improved by introducing LAM native small-scale perturbations (Tracton *et al.*, 1998), by perturbing surface initial conditions (Sutton *et al.*, 2006), and by accounting for model physics errors (Du *et al.*, 2003; Eckel and Mass, 2005).

In developing ALADIN-LAEF, we have tried to introduce a quantification of those uncertainties. A major upgrade to ALADIN-LAEF was applied in 2009, in which the LAM specific initial uncertainties, surface uncertainties and the model uncertainties are addressed. This article describes briefly the new design of ALADIN-LAEF and its comparison with the pre-operational ALADIN-LAEF (with dynamical downscaling of ECMWF ensemble forecasts).

The new design of ALADIN-LAEF is built on new strategy for simulating errors in the model forecast system. For dealing with the initial uncertainties, a blending method is implemented which is based on the idea of combining the large-scale perturbation from the ECMWF SVs and the small-scale perturbations from the LAM native breeding vectors. The blending method combines the advantages of the ECMWF SV perturbation, which is computed for the future uncertainties (Buizza and Palmer, 1995; Molteni *et al.*, 1996), and those of the breeding vector, which accounts for the uncertainties from the past (Descamps and Talagrand, 2007). On the other hand, it minimizes the risk of inconsistency due to the different treatments of perturbations in the global and regional EPS systems. To simulate the error in the surface initial condition, a strategy of non-cycling surface breeding (NCSB) is also implemented in ALADIN-LAEF. This uses the perturbed atmospheric forcing to generate the perturbation to the surface initial condition, such as soil moisture, etc. As in many other LAMEPSs, multi-physics and coupling with ECMWF-EPS members are used for dealing with the model and LBC uncertainties. The benefits from the introduction of the quantification of those uncertainties and the performance of the new design of ALADIN-LAEF are investigated in this study.

The article is organised as follows. In section 2 the ALADIN-LAEF configuration is introduced. Section 3 describes the perturbation methods in ALADIN-LAEF: blending for atmospheric initial conditions, NCSB for surface initial conditions and multi-physics for model uncertainty. Section 4 is dedicated to the technical implementation of ALADIN-LAEF. Section 5 presents results of a two-month verification and comparison between the new design of ALADIN-LAEF with the pre-operational version. A summary and conclusions are given in section 6.

2. ALADIN-LAEF system configuration

The key element of a LAMEPS is the underlying limited-area model. ALADIN-LAEF uses the high-resolution LAM ALADIN-Austria (Wang *et al.*, 2006a). This is the ALADIN configuration running operationally at ZAMG. ALADIN is a hydrostatic, spectral LAM. It includes a hybrid vertical coordinate, a spectral method with bi-periodic extension of the domain using elliptical truncation of double-Fourier series, a two-time-level semi-Lagrangian advection scheme, semi-implicit time-stepping, fourth-order horizontal diffusion, Davies–Kalberg type relaxation and digital filter initialisation (DFI). A brief description of ALADIN model physics will be given in section 3.2.

The ALADIN-LAEF integration domain covers the whole of Europe and a large part of the Atlantic, as shown in Figure 1. This domain is chosen since the development of many weather systems in the Atlantic is very important for the forecast over Central Europe, and it is large enough to minimize the impact of LBCs on the Central European region. Du and Tracton (1999) have given a detailed discussion on the impact of LBCs on LAMEPS.

ALADIN-LAEF is run at a resolution of 18 km in the horizontal and 37 levels in the vertical, up to 54 h ahead and twice per day at 0000 and 1200 UTC. ALADIN-LAEF is constructed with 18 members, of which the first 16 members are perturbed and LBC perturbations are provided by the first 16 ECWMF EPS members. The other two ALADIN-LAEF members take IC and LBCs from the ECMWF EPS control member with resolution of T399L62 and the ECMWF deterministic forecast with resolution of T799L91, respectively.

3. Perturbations

3.1. Initial condition perturbation: Blending global SVs and LAM breeding vector

It is crucial to quantify the initial perturbations for a skilful LAMEPS. Many studies have been done on IC generation methods for global EPS. However, there are very limited studies on methods for LAM native initial perturbation generation. SREF at NCEP (Du *et al.*, 2003) uses breeding to generate LAM initial perturbations and Bowler and Mylne

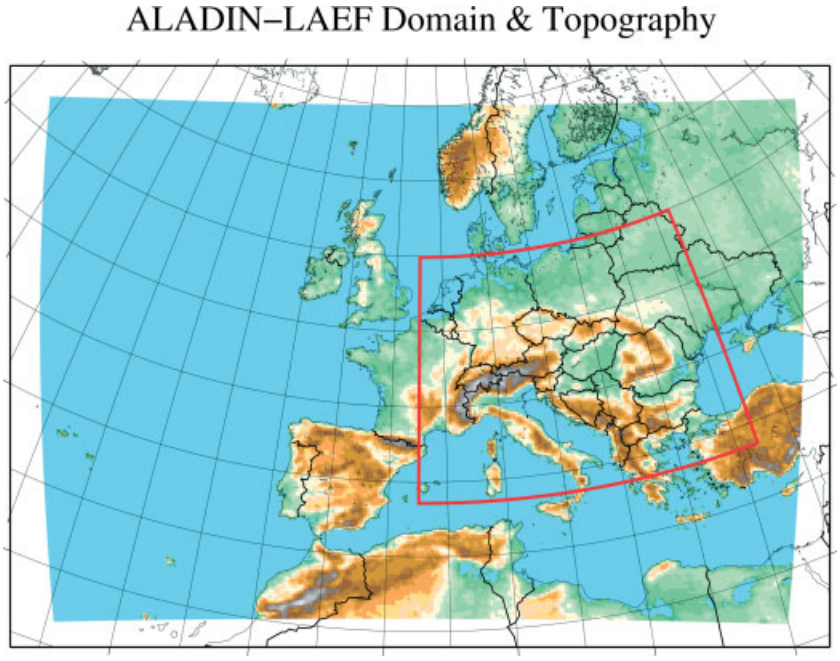

Figure 1. ALADIN-LAEF domain and model topography. The area oulined in red depicts the verification domain, covering Central Europe.

(2009) tested regional ETKF in MOGREPS. Experiments on LAM SVs (Stappers and Barkmeijer, 2008) and LAM ETKF/ET (Wang *et al.*, 2006b) are still in a very early stage. Most LAMEPS simply nest a LAM ensemble within a global ensemble or downscale the analyses from different forecast centres.

In a LAMEPS context, two requirements for generating IC perturbations have to be considered. Firstly, the IC perturbations should be effective immediately from the initial time; this means that they should focus on quantifying the uncertainties in the analysis (Bowler *et al.*, 2008). Secondly, the LAM IC perturbations should be consistent with LBC perturbations; in other words the LAM IC generation method should generate initial boundary values that are reasonably consistent with the LBC perturbations provided by the global EPS with which the LAMEPS couples. For example, it is very likely that a system using LAM breeding perturbations coupled to a global EPS using SVs, would lead to spurious gravity-wave generation at the lateral boundaries. It is unclear if an ill-posed set-up like this would be superior to the driving global EPS, which is an essential requirement for the usefulness of a LAMEPS. Meshing LAM ETKF/ET with global ETKF/ET (Bowler and Mylne, 2009), LAM breeding with global breeding (Du *et al.*, 2003) or LAM SVs with global SVs are probably good ways to avoid spurious gravity-wave generation at the lateral boundaries (Bishop, pers. comm.).

For ALADIN-LAEF, the practically and operationally available global EPS are ECMWF and PEARP/ARPEGE (Descamps *et al.*, 2009) which use the SV technique to simulate the initial uncertainty. So the obvious choice for generating the small-scale IC perturbations in ALADIN-LAEF would be to use SVs. However, LAM SV research is still at a very early stage. In addition, SVs may be inappropriate to generate perturbations very close to initial time, which are crucial for LAMEPS. This is due to the design of the SVs, which are optimized to grow over a certain forecast time (using very crude analysis-error estimates). This approach is reasonable for medium-range forecasts, but its usefulness is less clear for LAMEPS, and probably it is not suitable for short-range ensemble forecasts (Bowler *et al.*, 2008). Moreover, computation of SVs is expensive.

The use of breeding in ALADIN-LAEF was encouraged by its simplicity, cost effectiveness and the success of its application at NCEP. In order to ameliorate the aforementioned inconsistency problem of coupling with ECMWF EPS, a new idea of blending for generation of LAM native perturbations has been implemented in ALADIN-LAEF. The idea is to use the ALADIN blending technique (Brožková *et al.*, 2001), a digital filter and spectral analysis method, to combine the large-scale uncertainty generated by ECMWF SVs with the small-scale uncertainty generated by breeding with the ALADIN model. The combined perturbation has the feature that its large-scale part of the perturbation is from ECMWF SVs, and the small-scale part is from ALADIN breeding.

It is believed that the new perturbations can:

- reduce the inconsistency between the different perturbation methods used in the global ensemble system (to determine the perturbation at the LAM domain boundaries) and in the LAMEPS (to create initial state uncertainty); this is based on the fact that the large-scale part of the new initial perturbations is consistent with its global counterpart;
- represent the small-scale uncertainty in the analysis in more detail and more accurately due to the higher resolution and more balanced orographic/surface forcing of the LAM breeding; this should be a better representation of reality than interpolated large-scale perturbations from the global model. Buizza *et al.* (2005) found an advantage of breeding in short-range and small-scale forecasts; and
- contain the future uncertainty generated by SVs and the uncertainty from the past generated by breeding. As discussed by Toth and Kalnay (1997), Ehrendorfer (1997) and Buizza *et al.* (2005), breeding attempts to give the best estimate of the actual errors in the

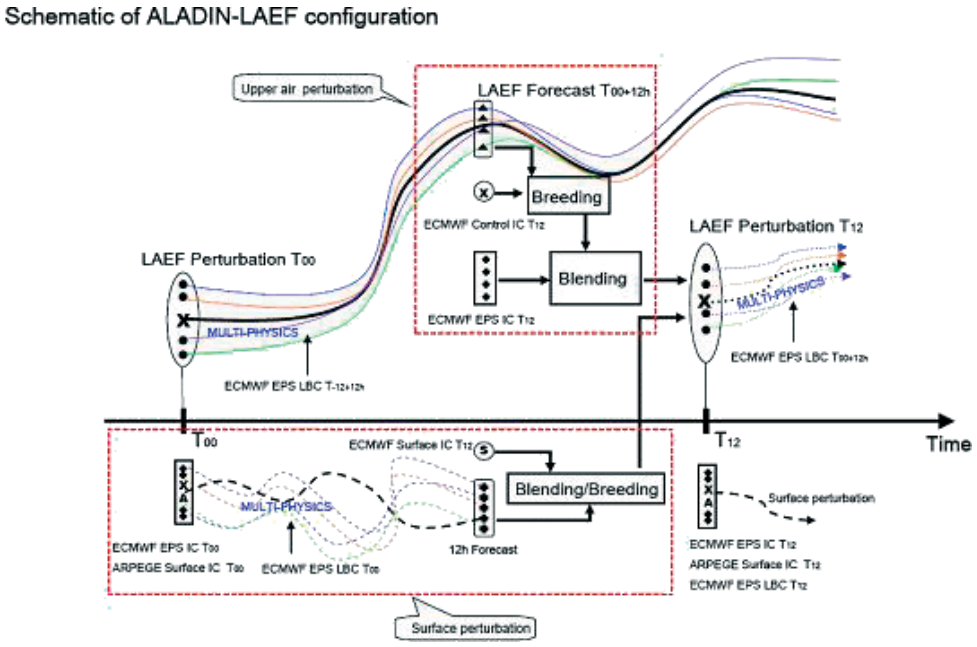

Figure 2. Schematic Figure description of the ALADIN-LAEF configuration. It shows the generation of upper-air perturbation of ALADIN-LAEF with breeding and blending (upper part) and the generation of the surface perturbation with non-cycling surface breeding (lower part), valid for the 1200 UTC run. Detailed discussion is given in the text.

initial analysis based on the past information of the flow, whereas the SVs contain future information of possible forecast error.

In the following, the methods which are applied for the breeding and blending in ALADIN-LAEF will be briefly described.

3.1.1. Breeding

Breeding (of growing vectors) is designed to simulate how the growing errors are 'bred' and maintained in a conventional analysis cycle through the successive use of short-range forecasts. The bred vectors should thus provide a good estimate of possible growing error in the analysis (Toth and Kalnay, 1993, 1997). The set-up of breeding consists of the following steps:

- the introduction of an arbitrary perturbation to the control analysis (this should be done only once);
- integrating the model with control analysis and the perturbed IC;
- building the difference between the two forecasts at a fixed time interval (to correspond better to the assimilation cycle);
- scaling down the forecast difference in amplitude to the size of the perturbation; and
- adding/subtracting the rescaled difference to the new control analysis.

Steps (ii) to (v) are then repeated, and the perturbations are bred to grow along the forecast trajectory.

In ALADIN-LAEF breeding, the perturbed initial conditions are generated in sets of positive and negative pairs around a control analysis. The breeding implementation has the features: (i) cold start, (ii) 12 h cycle, (iii) two-side and centring around the control analysis, (iv) wind components, temperature, moisture and surface pressure are perturbed at each level and model grid point, and (v) constant rescaling.

3.1.2. Blending large-scale perturbations from SVs and small-scale perturbations from breeding

The core idea of blending is to combine large-scale features of the perturbed global model analysis (represented by ECMWF EPS members generated by the SV technique) with the small-scale features provided by the perturbed LAM analysis (generated by ALADIN-LAEF breeding). This is motivated by the fact that the information on the mesoscale uncertainty is more reliably represented by the short-wave part of the LAM analysis spectrum than by the short-wave part of the corresponding perturbed ECMWF analysis obtained by interpolation, since these scales are not resolved by the global model.

The ALADIN blending technique (Brožková *et al.*, 2001) is a spectral technique using a standard non-recursive low-pass Dolph–Chebyshev digital filter. For spectral models like ALADIN, where the model prognostic variables are defined by the spectral coefficients of their Fourier expansion, the new blended state can be quite easily obtained by the combination of appropriate wave numbers over the selected part of the spectrum. Furthermore, as a result of an effective response of the digital filter in the stop-band, the spectral coefficients are progressively damped and not strictly set to zero (hence the possible shock in the transition zone can be easily avoided).

The blending procedure consists of several subsequent steps. The main principle is to apply a digital filter to the perturbed initial states from ECMWF SV and ALADIN breeding on the original ALADIN grid but at a lower spectral resolution. This lower spectral resolution is defined by the blending ratio, which according to the theory depends on the scales that can be analyzed by the driving model rather than on the ones it can predict. The difference between those filtered fields represents a large-scale increment. This increment contains almost pure low-frequency perturbation information, which is then added to the original high-frequency signal of the perturbed high-resolution LAM

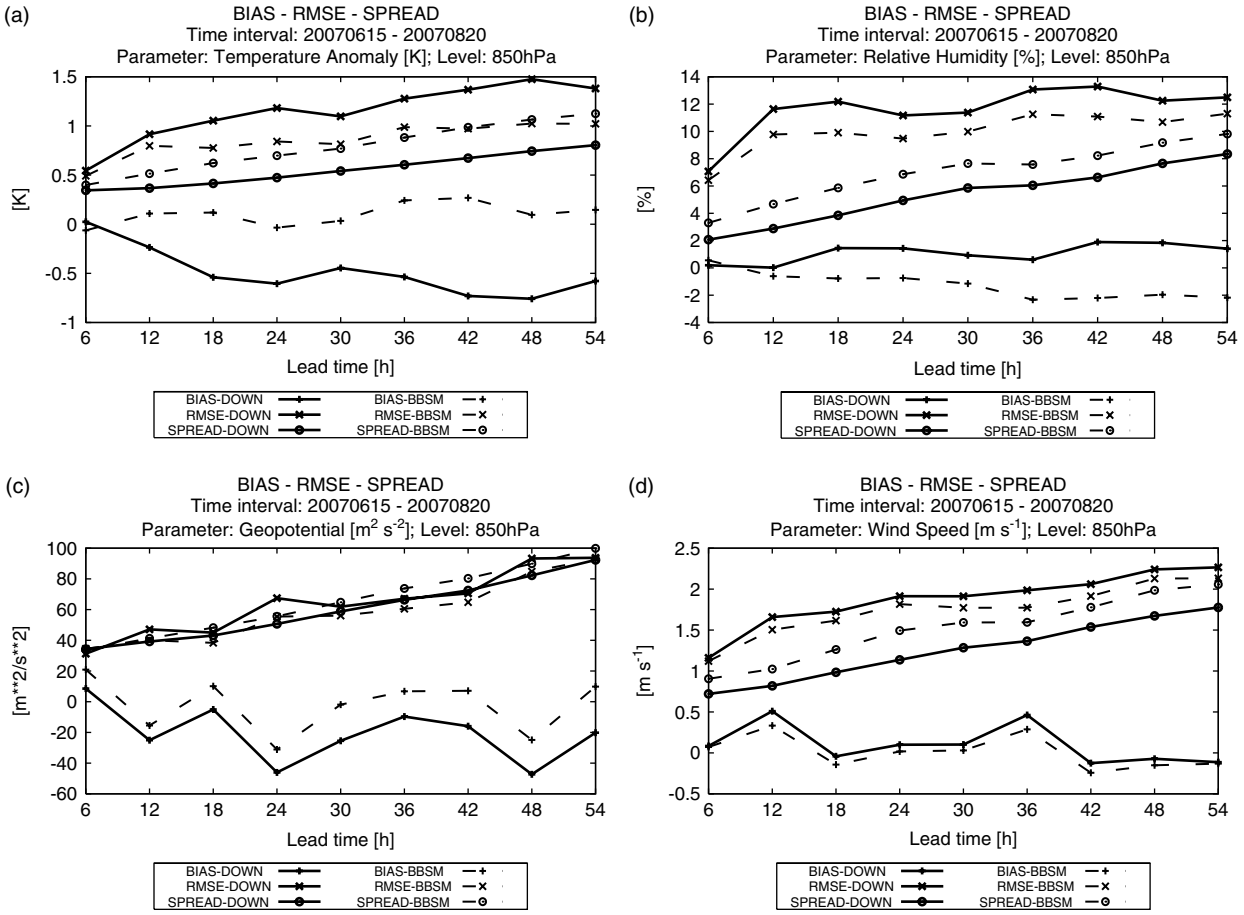

Figure 3. RMS error of the ensemble mean (×) and ensemble spread (○) for DOWN (solid lines) and BBSM (dashed lines) for (a) T850, (b) RH850, (c) Z850, and (d) V850, averaged over the verification domain (Figure 1) and over the verification period 15 June to 20 August 2007.

analysis (i.e. to the ALADIN breeding analysis). The combination (blending) of both spectra is performed in the transition zone.

The digital filter (DF) technique was originally used for the initialization of meteorological fields in NWP. Its detailed description can be found in Lynch and Huang (1992). The symbolic equation of blending can be summarized after Brožková *et al.* (2006) and Derkova and Bellus (2007):

$$IC_{\mathrm{blend}(m)} = A_{\mathrm{bred}(m)} + \left\{ \left(\overline{A}^{\mathrm{DF}}_{\mathrm{sv}(m)} \right)_{\mathrm{LOW}} - \left(\overline{A}^{\mathrm{DF}}_{\mathrm{bred}(m)} \right)_{\mathrm{LOW}} \right\}_{\mathrm{HIGH}}, \quad (1)$$

where IC_{blend} denotes initial condition after blending, A_{sv} stands for the perturbed analysis generated by ECMWF SVs, and A_{bred} for the ALADIN breeding analysis. LOW denotes the operations performed at cut-off blending truncation and HIGH the ones at ALADIN original spectral resolution. Index m indicates the mth member of the ensemble. The digital filter is applied at spectral resolution LOW to remove small-scale noise or to obtain a clean long wave state. The final result IC_{blend} is created at ALADIN spectral resolution HIGH by adding the acquired large-scale increment to the original ALADIN breeding analysis.

The detailed description and discussion of blending, in particular the technical implementation in ALADIN-LAEF and details on the tuning of the breeding–blending cycle, are given in Bellus (2008) and Bellus *et al.* (2010).

3.2. *Model perturbation: multi-physics*

Imperfection and simplification in model formulation, in particular in model parametrizations, give rise to the main forecast errors. Addressing those model uncertainties is necessary for a skilful LAMEPS. In ALADIN-LAEF, the multi-physics approach is applied to deal with the uncertainties due to model errors. Different ALADIN configurations and variations of certain parametrizations are included. The different parametrization configurations used in ALADIN-LAEF are summarized in Table I. They comprise different parametrization schemes and variations of those schemes. The physical processes addressed by these configurations are mainly large-scale and subgrid-scale precipitation, radiation, turbulent transport and diffusion processes.

As shown in Table I, five different ALADIN parametrization configurations are used. The main characteristics of these configurations are:

- ALADIN-25. This ALADIN physics set-up (Gerard, 2000, pers. comm.) can be seen as a reference or basic setting which is well tuned for ALADIN resolution around 15 km. The main physical parametrizations in ALADIN-25 are: a revised Kessler scheme for cloudiness and large-scale precipitation (Kessler, 1969); the mass-flux-type scheme of Bougeault and Geleyn (1989) for deep convection with moisture convergence closure (BGMC) or with CAPE closure (BGCP); a radiation scheme based on Ritter and Geleyn (1992),

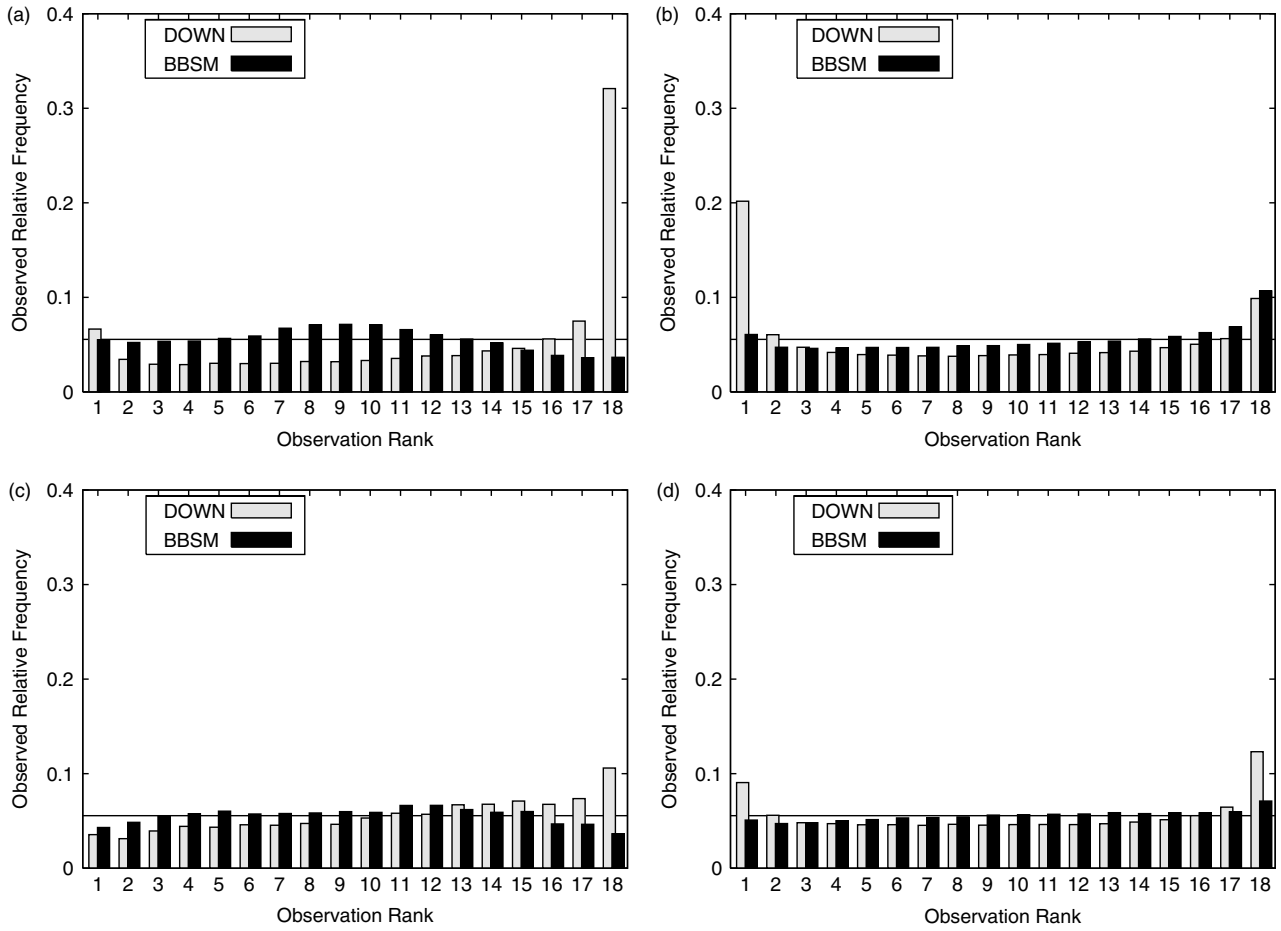

Figure 4. Talagrand diagram of DOWN and BBSM for forecast lead time +54 h for (a) T850, (b) RH850, (c) Z850, and (d) V850. The verification domain and period are as Figure 3.

Table I. Summary of ALADIN-LAEF multi-physics.

Member	Configuration	Cloud physics	Deep convection	Radiation Radiation	Turbulent transport	Shallow convection	Mixing length and entrainment rate setting
M 1	ALADIN-25	Kessler	BGMC	RG	Louis81	JFG87	0
M 2	ALADIN-25	Kessler	BGCP	RG	Louis81	JFG87	1
M 3	HARMONIE	Sundquist	STRACO	Savijärvi90	CBR+S90	JFG87	–
M 4	ALARO+3MT	ALARO	3MT	JFG05	JFG06	JFG87	–
M 5	ALADIN-32	Lopez	BGMC	ECMWF	Louis81	KFB	0
M 6	ALADIN-32	Lopez	BGCP	ECMWF	Louis81	KFB	1
M 7	ALARO	ALARO	BGMC	JFG05	JFG06	JFG87	–
M 8	ALARO	ALARO	BGCP	JFG05	JFG06	JFG87	–
M 9	ALADIN-32	Lopez	BGMC	ECMWF	CBR+B81	KFB	0
M 10	ALADIN-32	Lopez	BGCP	ECMWF	CBR+B81	KFB	1
M 11	ALADIN-32	Lopez	BGMC	ECMWF	CBR+S90	KFB	0
M 12	ALADIN-32	Lopez	BGCP	ECMWF	CBR+S90	KFB	1
M 13	ALADIN-32	Lopez	BGMC	ECMWF	CBR+S90	KFB	0
M 14	ALADIN-32	Lopez	BGCP	ECMWF	CBR+S90	KFB	1
M 15	ALARO+3MT	ALARO+XR	3MT	JFG05	JFG06	JFG87	–
M 16	ALARO+3MT	ALARO+XR1	3MT	JFG05	JFG06	JFG87	–
M 17	ALARO	ALARO	BGMC	JFG05	JFG06	JFG87	–
M 18	ALADIN-32	Lopez	BGMC	ECMWF	Louis81	KFB	0

More details, acronyms and abbreviations are given in the text (section 3.2).

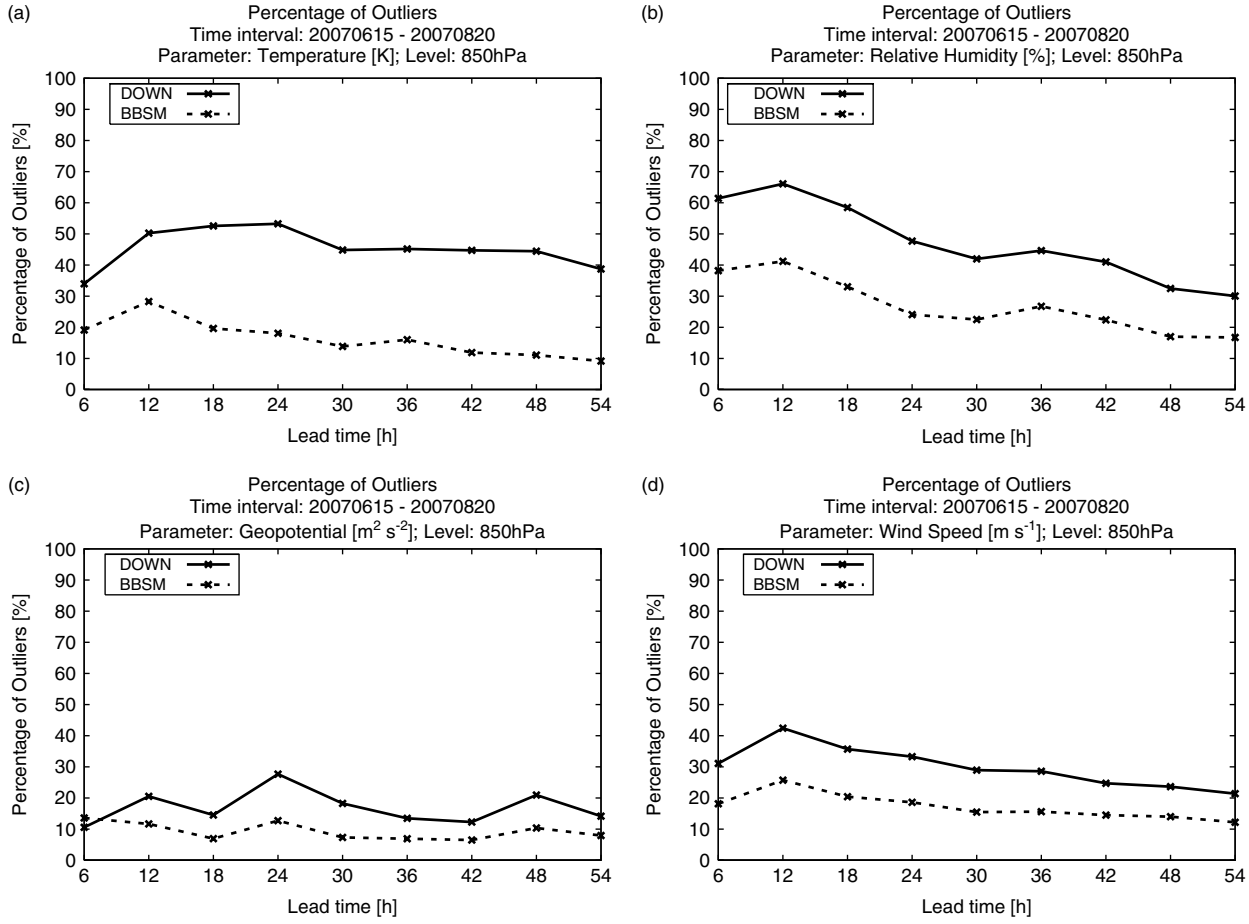

Figure 5. Percentage of outliers of DOWN and BBSM for (a) T850, (b) RH850, (c) Z850, and (d) V850. The verification domain and period are as Figure 3.

denoted RG; the computation of turbulent fluxes of heat, water vapour and momentum designed on the basis of Louis *et al.* (1981); and shallow convection parametrized using JFG87 (Geleyn, 1987). The Interactions–Soil–Biosphere–Atmosphere (ISBA) scheme (Noilhan and Planton, 1989; Giard and Bazile, 2000) is used.

- ALARO. This configuration was used as the operational one in ALADIN-Austria up to April 2009. It is an ALADIN physics package developed within the ALADIN/LACE cooperation. The main modifications from ALADIN-25 are in large-scale precipitation, turbulent transport and radiation. For resolved cloudiness and precipitation, a prognostic type parametrization ALARO is used (Catry *et al.*, 2007; Geleyn *et al.*, 2008; Gerard *et al.*, 2009), a pseudo-prognostic turbulent scheme JFG06 is introduced (Geleyn *et al.*, 2006, pers. comm.), and the radiation is calculated by using JFG05 (Geleyn *et al.*, 2005, pers. comm.). The schemes for deep convection, soil processes and mountain drag as depicted in ALADIN-25 remain mainly unchanged.
- ALARO+3MT. In addition to the components listed for ALARO, this configuration includes the new parametrization scheme 3MT (Modular Multiscale Microphysic and Transport; Gerard and Geleyn, 2005; Gerard, 2007; Piriou *et al.*, 2007; Geleyn *et al.*, 2008; Gerard *et al.*, 2009), which was mainly developed in order to deliver physically consistent results over a wide range of model resolutions, in particular for resolutions between 7 and 3 km. The use of the diagnostic convection scheme as used for ALADIN-25 and ALARO becomes obsolete. In the ALADIN-LAEF multi-physics design, three variations of configuration ALARO+3MT are employed. There are two different options for the computation of the resolved part of condensation and evaporation processes in this physical package. Variation 1 (ALARO) uses the Smith-based option (Smith, 1990; Gerard, 2007); for variation 2 (ALARO+XR) the Xu–Randall type computation (Xu and Randall, 1996) is activated; and in variation 3 (ALARO+XR1) the Xu–Randall option is used together with an option affecting the protection of convective condensate produced during the updraught computation.
- ALADIN-32. This configuration is based on the operational ALADIN set-up at Météo-France. The main differences from the configuration ALADIN-25 are: the large-scale cloudiness and precipitation scheme is the prognostic scheme proposed by Lopez (2002); the ECMWF radiation scheme used in ALADIN-32 is the Rapid Radiative Transfer Model (RRTM) for the long wave and the scheme of Morcrette (1991) for the short wave; and a prognostic Turbulent Kinetic Energy (TKE) scheme CBR (Cuxart *et al.*, 2000) is used for calculation of turbulent transport. In ALADIN-LAEF multi-physics, the CBR scheme is combined with two options (CBR+B81 and CBR+S90) for modelling the cloud in the boundary

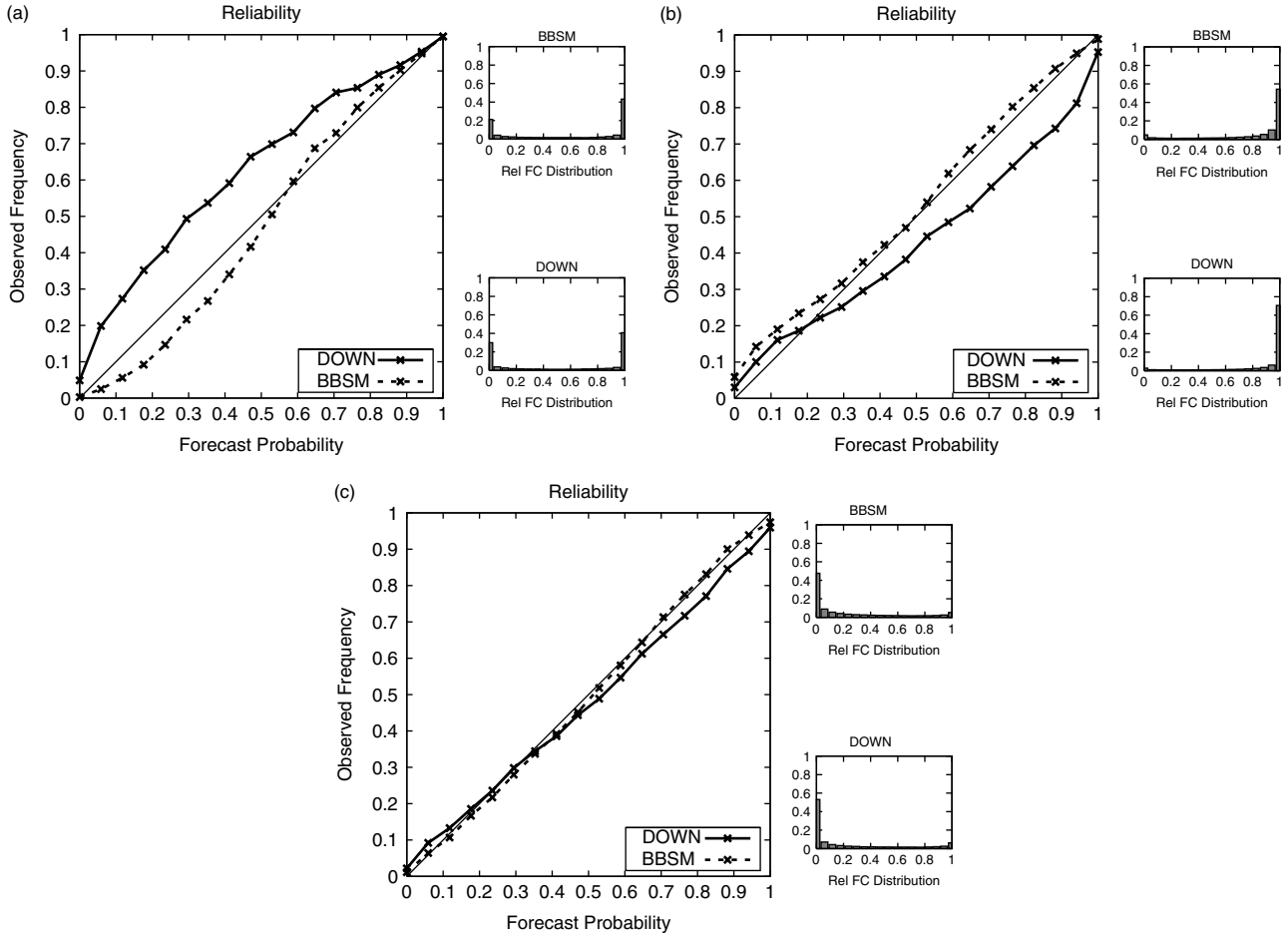

Figure 6. Reliability diagram of DOWN and BBSM for forecast lead time +54 h for (a) T850 (threshold >0°), (b) RH850 (>40%), and (c) V850 (>10 m s^{-1}).

layer, the one of Bougeault (1981) or the one of Smith (1990). Moreover, in configuration ALADIN-32, shallow convection is parametrized using the KFB scheme (Bechtold *et al.*, 2001). Those are also used as variations in the ALADIN-LAEF multi-physics.

- HARMONIE (Hirlam ALADIN Regional/Mesoscale Operational NWP in Europe) permits the use of the HIRLAM (High Resolution Limited-Area Modelling) parametrization packages within the framework of ALADIN dynamics. The combination of ALADIN dynamics with HIRLAM physics is one option in the ALADIN-LAEF multi-physics. The main elements of this physics package are: the HIRLAM radiation scheme developed by Savijärvi (1990) and described in Sass *et al.* (1994) and Wyser *et al.* (1999); and the CBR scheme (Cuxart *et al.*, 2000) with adjustments to the length-scale (Lenderink and Holtslag, 2004) for the vertical diffusion. HIRLAM convection and condensation scheme (Soft Transition Condensation, STRACO) puts emphasis on the gradual transition from convective to stratiform regimes. It is a modified Kuo-type convection scheme (Kuo, 1974). The microphysics and precipitation processes are based on the work of Sundqvist (1993).

Two different settings of some empirical and adjustable parameters for the calculation of mixing length, entrainment rate and cloud base are included in configurations ALADIN-25 and ALADIN-32. The values of those parameters are estimated/chosen by ALADIN physics experts (Geleyn, pers. comm.) to ensure that the ALADIN model keeps the same forecast quality for all settings.

3.3. *Surface initial condition perturbation*

Perturbing surface initial conditions, such as soil moisture, should have a beneficial impact on the skill of short-range probabilistic surface weather parameter forecasts (Sutton *et al.*, 2006). As mentioned in the introduction, various approaches are employed for dealing with the uncertainties that are related to the initial state of the atmosphere. But there are hardly any methods for perturbing the initial state of the land surface in ensemble systems. For example, the initial state of soil moisture and soil temperature is the same for each member of the ensemble system. In the following NCSB, a strategy to generate perturbations of the surface variables, e.g. soil moisture content and surface temperature, is briefly described. More detailed description and discussion of NCSB are given in Wang *et al.* (2010).

The idea behind NCSB is to perturb the surface initial conditions by applying short-range forecasts driven by perturbed atmosphere forcing and the breeding method. Similar to breeding, the simulation of the fast-growing 'errors of the day' (Kalnay, 2003) on the surface state is started by introducing perturbations in the atmosphere. The perturbations are not randomly seeded, but downscaled from a global EPS. The LAM model is then integrated for 6 or 12 h with the perturbed atmospheric initial and lateral

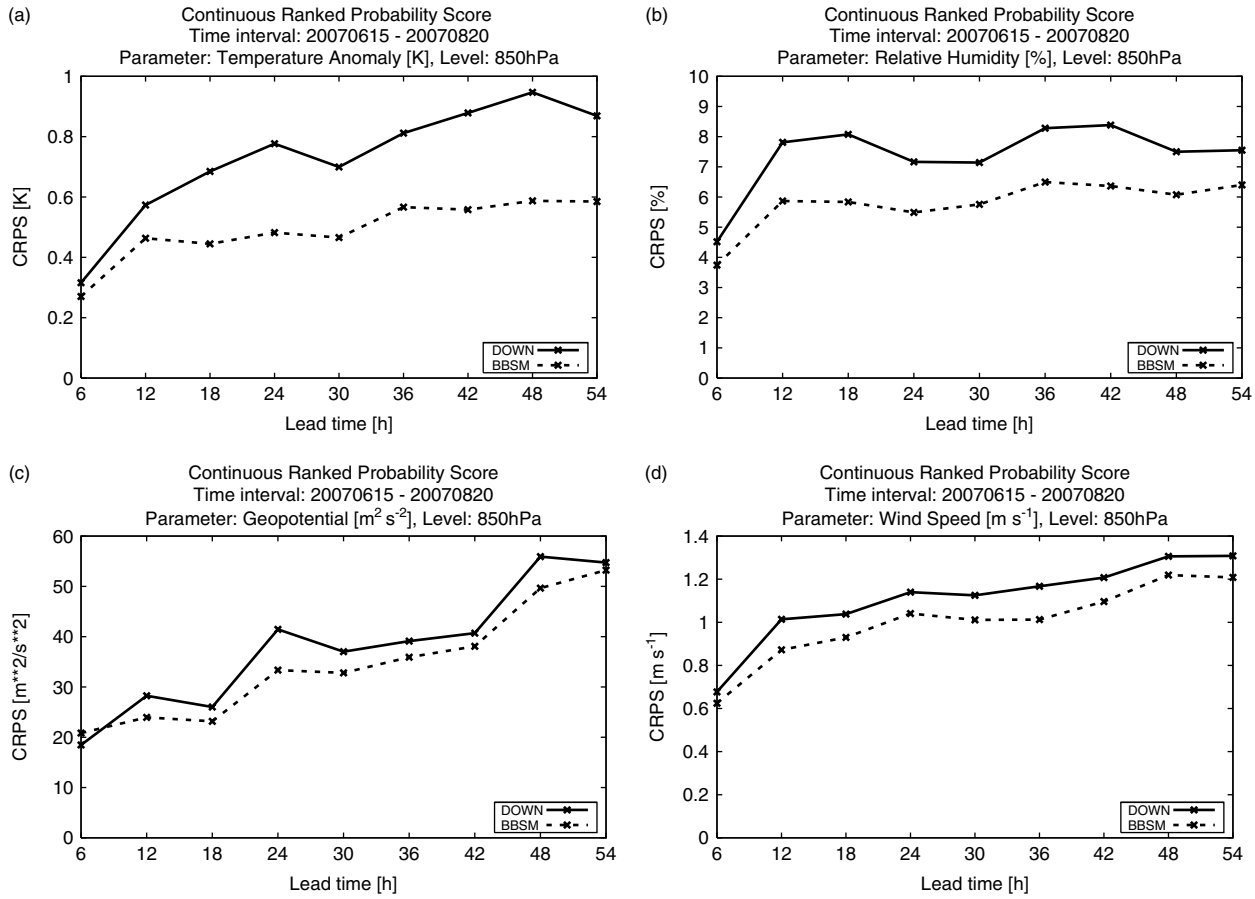

Figure 7. Continuous ranked probability scores of DOWN and BBSM for (a) T850, (b) RH850, (c) Z850, and (d) V850. The verification domain and period are as Figure 3.

boundary conditions which are provided by its global EPS counterpart. For initialization of these short-range LAM forecasts, only one common surface initial state is used. The 6 or 12 h surface forecasts are subtracted from the corresponding new surface analyses; the differences are rescaled, and added to the corresponding new surface analysis. The whole process is repeated every model run with new unperturbed surface initial conditions and atmospheric perturbations obtained from the global EPS. This non-cycling of the surface perturbations ensures that the surface initial perturbations in LAMEPS are only driven by the atmospheric perturbations from the global EPS with little impact from the LAM itself. If the surface perturbations were cycled, as in the atmospheric breeding, the surface model climate would most probably drift away from the real climate after several months (Geleyn, 1988; Brožková, 2007, pers. comm.).

In the implementation of the surface perturbation in ALADIN-LAEF, the 16 perturbed atmospheric initial conditions are downscaled from the first 16 initial perturbations of ECMWF-EPS. LBC perturbations are also obtained from the forecasts of the corresponding ECMWF-EPS members. The aforementioned multi-physics approach in section 3.2 is applied to represent model uncertainty. ARPEGE surface analysis is used instead of the ECMWF surface analysis. This is due to differences in the surface physical parametrization between the ECMWF model and ARPEGE/ALADIN. Thus, using the surface variables of ECMWF in the ALADIN integration causes inconsistencies. This deficiency can be reduced to some extent if the model surface from ECMWF is replaced by the ARPEGE surface analysis.

Mathematically, the perturbed initial conditions for the surface variables A_m, where m is the mth ensemble member, are calculated from the surface initial conditions of the control run C and the perturbed 12 h forecasts P_m as follows:

$$A_m = C + S_m(P_m - C). \tag{2}$$

The scaling factor S_m is set equal to 1 in the current implementation.

4. Technical implementation

The presented methods for generating the initial perturbations (breeding, blending, NCSB) and the multi-physics approach as model uncertainty representation, have been implemented in the upgraded ALADIN-LAEF. LBC perturbations are introduced by coupling to 12 h time-lagged ECMWF EPS forecasts. This time lag is determined for operational reasons as the forecasts should be available as early as possible. The technical realization of the upgraded ALADIN-LAEF is outlined in Figure 2.

In ALADIN-LAEF, two runs per day are performed with initial times at 0000 and 1200 UTC. For the 1200 UTC run, the production of ALADIN-LAEF forecast is as follows:

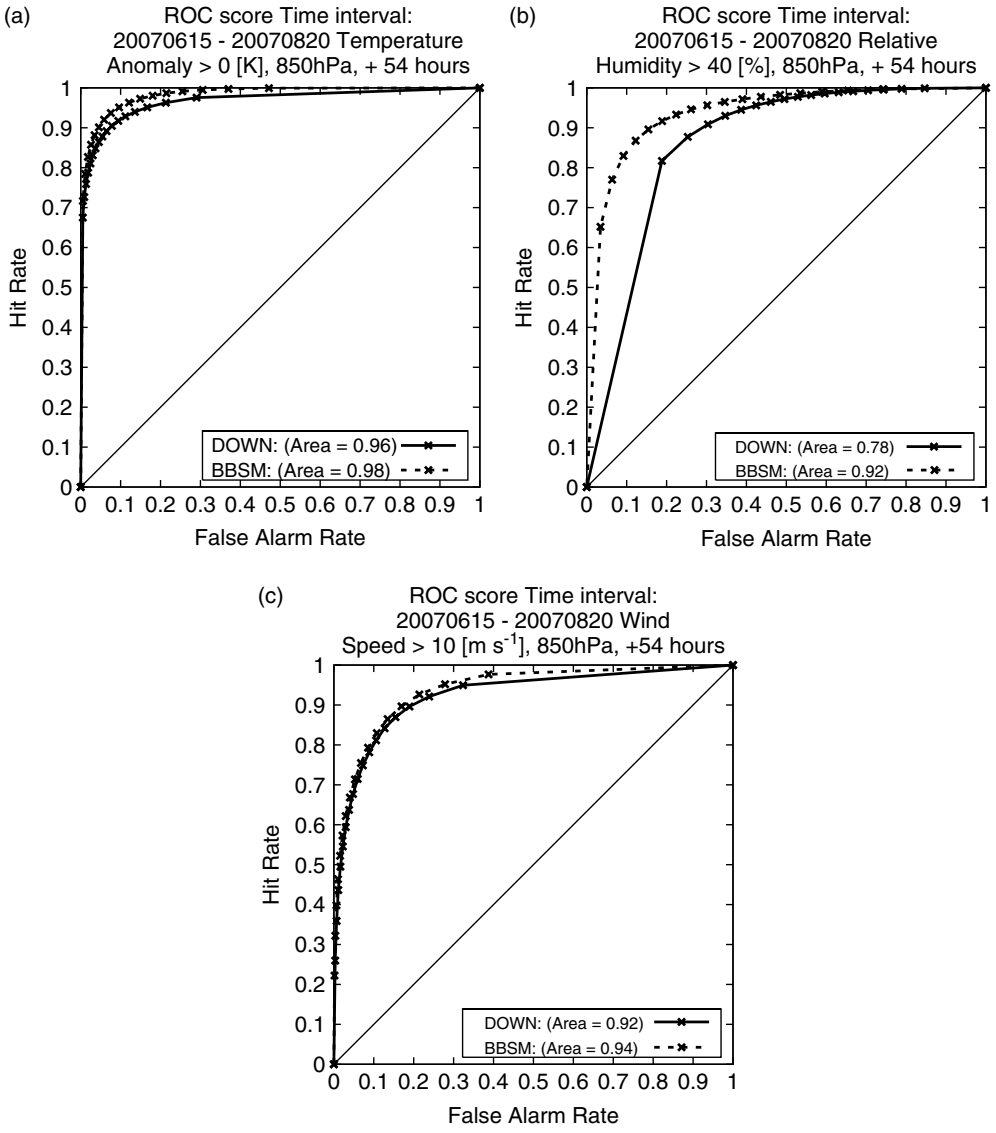

Figure 8. The relative operating characteristics (ROC) and area under ROC of DOWN and BBSM for forecast lead time +54 h for (a) T850, (b) RH850, and (c) V850. The verification domain and period are as Figure 3, and the thresholds are as Figure 6.

4.1. Data preparation

The data preparation for the ALADIN-LAEF 1200 UTC run includes two steps:

(i) Downscaling of the lateral boundary conditions (LBCs) from the ECMWF EPS resolution to the ALADIN-LAEF resolution. The LBCs given by ECMWF EPS members 01–16 of the 0000 UTC run, denoted as **ECMWF EPS LBC T_{00+12h}** in Figure 2, will be used in the following main forecast at 1200 UTC as time-lagged LBCs.

(ii) Generation of the surface initial perturbations using NCSB (summarized in the lower part of Figure 2, labelled 'Surface perturbation'):

- The ECMWF surface is replaced with ARPEGE surface (**ARPEGE Surface IC T_{00}**) in the IC at time 0000 UTC.
- ALADIN-LAEF members are started at 0000 UTC and integrated up to +12 h. The corresponding atmospheric perturbations of initial conditions and LBC are provided by ECMWF-EPS members of the 0000 UTC run (**ECMWF EPS IC T_{00}** and **ECMWF EPS LBC T_{00}**, respectively in Figure 2); the multi-physics approach is applied for representing the model error.
- The 12 h ALADIN-LAEF surface forecasts, valid at 1200 UTC, are used to specify perturbed surface initial conditions, similar to those in the breeding method.

4.2. Main forecast run

The main 1200 UTC ALADIN-LAEF forecast is described in the upper part of Figure 2 (labelled 'Upper air perturbation'). It consists of the following steps:

- Downscaling of the IC from the ECMWF-EPS 1200 UTC control run, **ECMWF Control IC T_{12}**, to ALADIN-LAEF resolution.
- Generation of the mesoscale part of the initial conditions for ALADIN-LAEF members using breeding (section 3.1.1). This step uses the +12 h forecast of the previous ALADIN-LAEF run, **LAEF Forecast T_{00+12h}**,

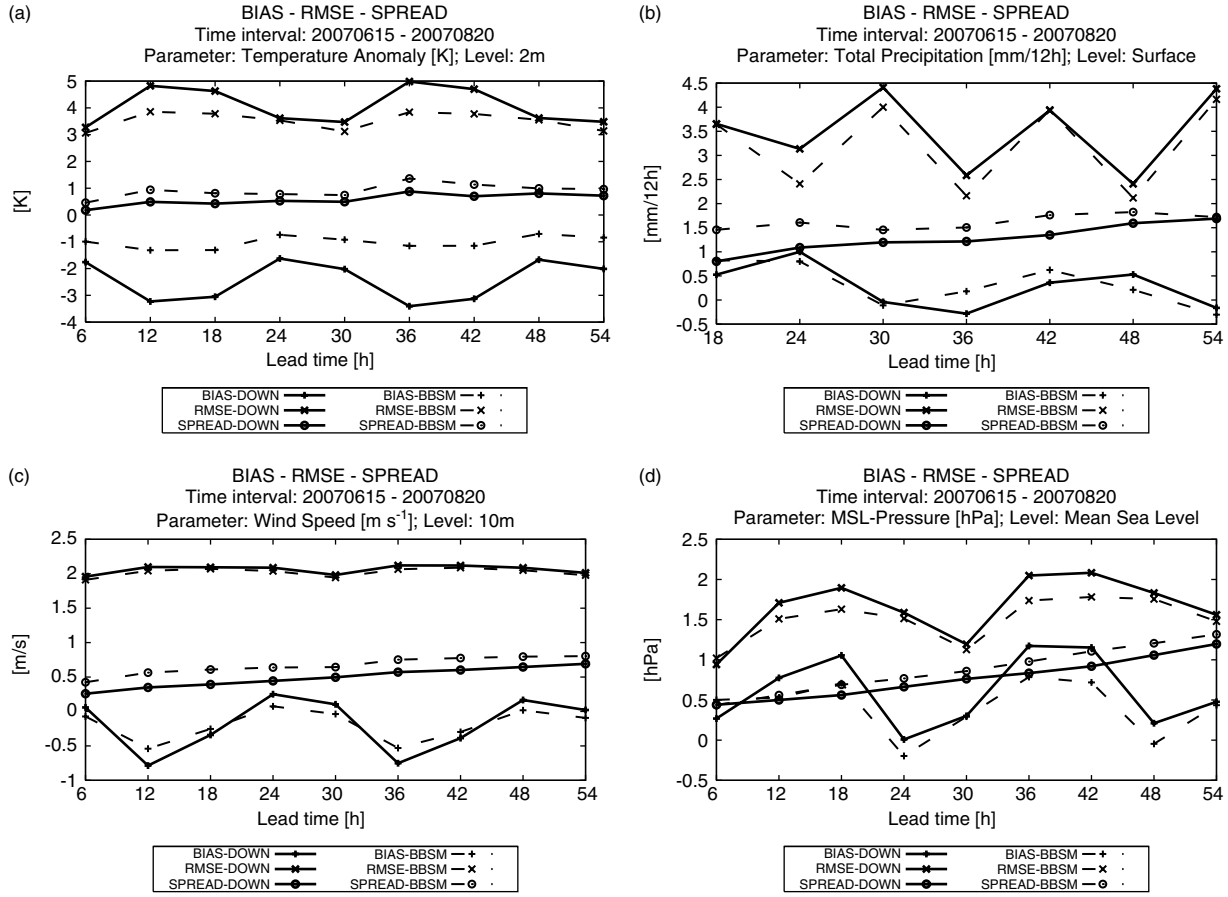

Figure 9. RMSE of the ensemble mean (×) and ensemble spread (○) of DOWN and BBSM for (a) T2m, (b) PREC, (c) W10m, and (d) MSLP. The verification domain and period are as Figure 3.

and the downscaled initial analysis of the ECMWF 1200 UTC run **ECMWF Control IC T_{12}**.

- Generation of the large-scale part of the initial conditions for ALADIN-LAEF members. This is done by downscaling of the IC from the 1200 UTC run ECMWF-EPS members, **ECMWF EPS IC T_{12}**.
- Blending (section 3.1.2). The large-scale structures of the perturbed IC from ECMWF SVs are combined with the small-scale structures from the ALADIN breeding. This gives the new atmospheric perturbations for the ALADIN-LAEF 1200 UTC run.
- Surface exchange. In this step the surface variables of all ALADIN-LAEF members are replaced by the perturbed surface variables generated by NCSB in the data preparation.
- Model integration. This uses **LAEF Perturbation T_{12}**, i.e. the perturbed atmospheric IC resulting from breeding/blending and the perturbed surface IC from NCSB, described in the previous steps. The 12 h time-lagged perturbed LBC from ECMWF-EPS and the multi-physics are applied for integration.
- Post-processing of the ALADIN-LAEF forecast to generate ALADIN-LAEF products. These are then archived in the Meteorological Archive and Retrieval System at ECMWF and provided to all the ALADIN/LACE partners.

After the main ALADIN-LAEF forecast, the preparations for the next cycle begin. These include the preparation of the current 12 h ALADIN-LAEF forecast for the next breeding cycle and performing another cycle independent 12 h ALADIN-LAEF forecast driven by atmospheric initial perturbations and LBC from ECMWF-EPS members to generate the surface perturbations for the next main forecast.

5. Results

In this study, the performance of the upgraded ALADIN-LAEF (referred to as BBSM: breeding, blending, surface perturbation, and multi-physics) and the pre-operational ALADIN-LAEF (referred to as DOWN) are compared for a two-month period (15 June to 20 August 2007). The ALADIN-LAEF runs have been initialized at 0000 UTC and are run for 54 h. ECMWF analysis is used for verification of the forecasts of upper-air weather variables, and both analysis and forecast are interpolated to a common regular $0.15^\circ \times 0.15^\circ$ latitude/longitude grid. Observations are used for the verification of surface weather variables, performed at the observation locations. Forecast values are interpolated to the observation site for smoothly varying fields, such as 2 m temperature, 10 m wind speed and surface pressure. For precipitation, the observation is matched to the nearest grid point. This matching method is commonly used in precipitation verification, but since it is rather simple, this method could lead to large errors in precipitation amount when large spatial gradients occur. The verification is performed for a limited area of the forecast domain over Central Europe, as shown in Figure 1. In the verification domain, there are 1219 SYNOP stations used in this study.

A set of standard ensemble and probabilistic forecast verification methods is applied to evaluate the performance

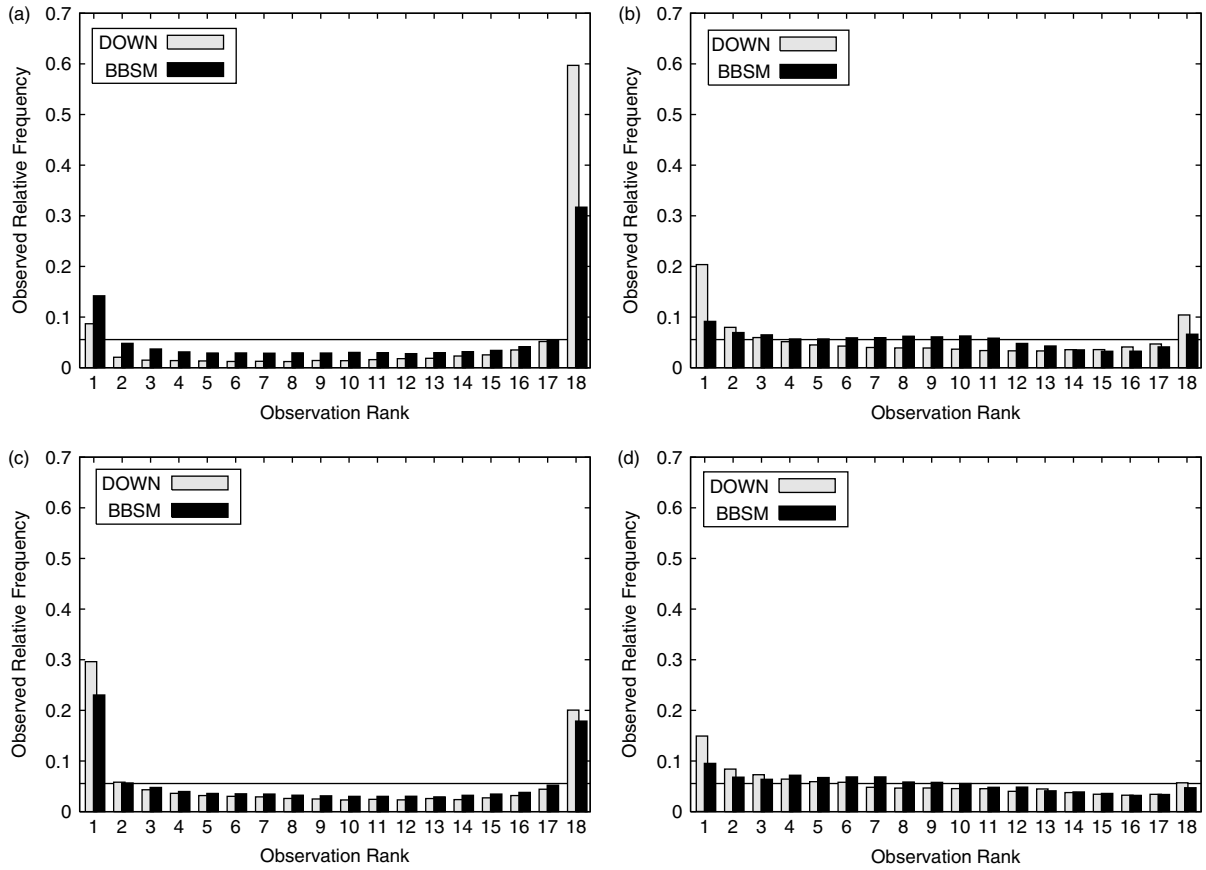

Figure 10. Talagrand diagram of DOWN and BBSM for forecast lead time +54 h for (a) T2m, (b) PREC, (c) W10m, and (d) MSLP. The verification domain and period are as Figure 3.

of the two ALADIN-LAEF configurations, BBSM and DOWN. The scores considered are ensemble spread, ensemble root-mean-square error (RMSE), the Talagrand diagram or rank histogram, continuous ranked probability score (CRPS), continuous ranked probability skill score (CRPSS), outlier statistics, relative operating characteristic (ROC) curve, the area under ROC (AROC), and the reliability diagram. A detailed description of those verification scores can be found in e.g. Mason (1982), Stanski *et al.* (1989) Anderson (1996), Hamill and Colucci (1997), Jolliffe and Stephenson (2003) and Wilks (2006). Those verification scores measure the quality of probabilistic forecasts of scalar quantities.

5.1. Verification of upper-air weather variables

Verification of upper-air weather variables (temperature, geopotential height, wind speed and relative humidity) has been carried out on different pressure levels, e.g. 500 hPa and 850 hPa. Similar results have been observed for different levels. The focus of this article is on temperature at 850 hPa (T850), geopotential height (Z850), wind speed (V850) and relative humidity (RH850).

The discrepancy between the ensemble spread and the error of the ensemble mean is a measure of the statistical reliability. The magnitude of ensemble spread should correspond to the magnitude of the RMSE of the ensemble mean. A large difference between the RMSE of the ensemble mean and ensemble spread is an indication of statistical inconsistency (Buizza *et al.*, 2005). Figure 3 shows RMSE and bias of ensemble mean, and ensemble spread of T850, Z850, V850 and RH850 for BBSM and DOWN.

For temperature, wind speed and relative humidity, BBSM shows increased spread and reduced RMSE compared to DOWN, hence a smaller discrepancy between RMSE and spread can be noticed for BBSM. BBSM performs clearly better than DOWN. For geopotential, the ensemble spread is slightly higher than RMSE for BBSM, while the ensemble spread matches the RMSE quite well for DOWN. Forecast bias of the ensemble mean of BBSM tends to be improved, except for RH850, which might be related to the model physics in the ALADIN-LAEF. The reduction of RMSE of temperature and humidity in BBSM mainly corresponds to the introduction of blending and NCSB. A requirement of a better forecast in a LAM is that the scales represented in the analysis should be in accordance with the scales resolved by the model. The dynamical downscaling method is by its nature incapable of meeting this requirement; the application of blending provides more realistic estimation of the analysis in all the scales resolved by the LAM, and this feature of blending contributes partly to the smaller RMSE in BBSM. Further, the use of NCSB ensures larger consistency in the surface coupling of ALADIN with ECMWF, which improves the forecast quality. Wang *et al.* (2010) showed that the positive impact of NCSB can be achieved for the lower-atmospheric variables, such as those at 850 hPa.

The Talagrand diagram evaluates the ability of an ensemble system to reflect the observed frequency distribution. It describes the characteristics of ensemble spread and bias (Talagrand *et al.*, 1997). A perfect ensemble system has a flat rank histogram. The uniform reference

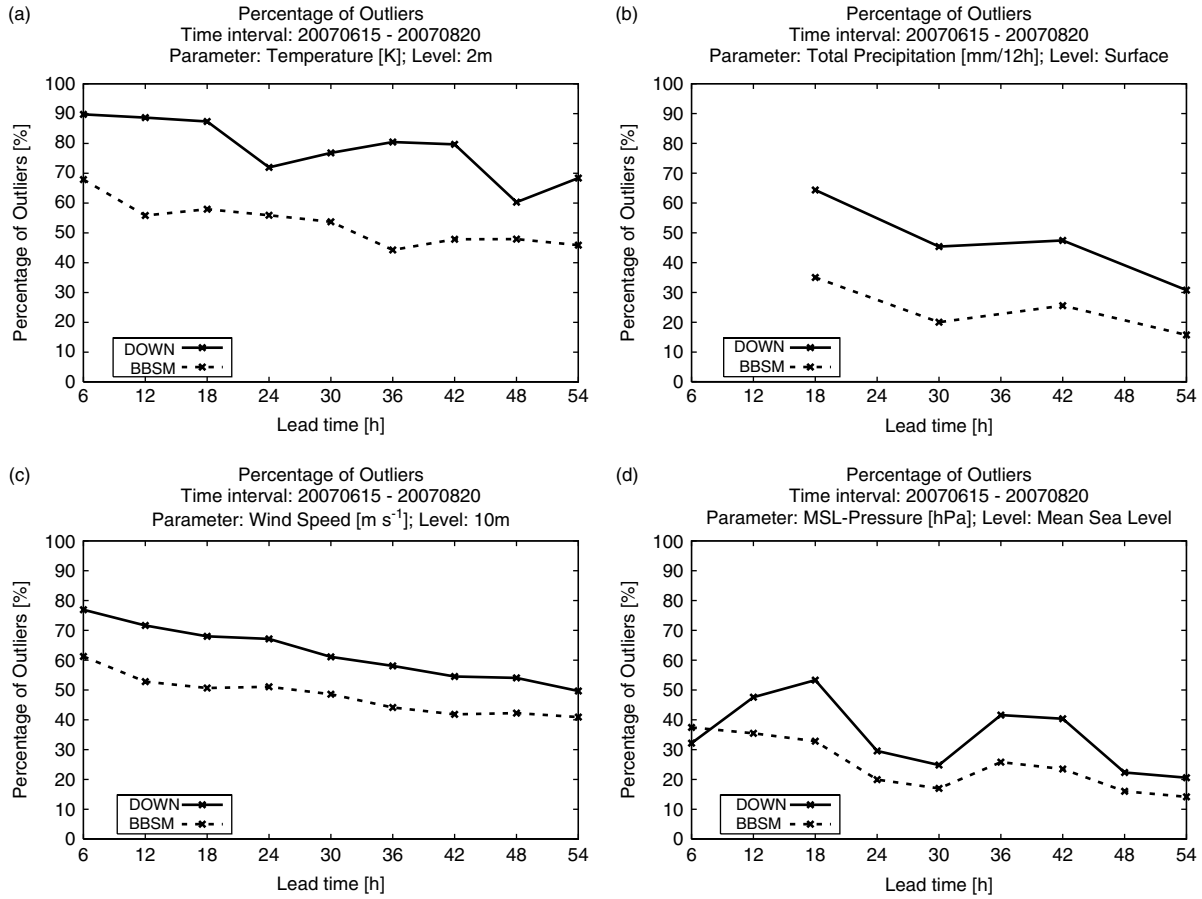

Figure 11. Percentage of outliers of DOWN and BBSM for (a) T2m, (b) PREC, (c) W10m, and (d) MSLP. The verification domain and period are as Figure 3.

rank is equal to $1/(n_{\text{ens}}+1)$, where n_{ens} is the ensemble size. Figure 4 shows the Talagrand diagrams for Z850, T850, RH850 and V850 at forecast lead time +54 h for BBSM and DOWN. The cold and moist biases in the DOWN forecast (T850 and RH850), underdispersion in V850 and a negative bias in Z850 have been improved by BBSM. The relative flatness of all rank histograms for BBSM is the indication that the observation distribution can be predicted well with BBSM. A slight overdispersion is found in Z850 for BBSM, which is consistent with the result in Figure 3.

Another measure of statistical reliability 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 $2/(n_{\text{ens}}+1)$. It is evident that BBSM has fewer outliers than DOWN for all the weather variables (Figure 5).

A comprehensive estimate of forecast quality considering statistical reliability and resolution is given by the reliability diagram. It is a graphical description of the resolution and reliability of an EPS by plotting the frequency of forecast probabilities against the related verification frequency. For a perfectly reliable EPS, the forecast and the verification probabilities should match each other (the diagonal line in the diagram). Figure 6 presents the reliability diagrams of T850, V850 and RH850 for DOWN and BBSM, valid at forecast lead time +54 h. The thresholds used in the diagrams are temperature anomaly $>0^{\circ}$C, wind speed >10 m s^{-1} and relative humidity $>40\%$. ECMWF ERA-40 data has been used to calculate climatological information, which is applied to define the thresholds in computing the reliability diagrams. As in the Talagrand diagram, a cold bias or underprediction in temperature and a moist bias or overprediction in the relative humidity are found for DOWN. BBSM has a good resolution and calibration for T850 and RH850 and shows slightly better performance for V850.

Figure 7 shows the CRPS of Z850, T850, V850 and RH850 for both BBSM and DOWN. The 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 has a negative orientation, 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. Figure 7 shows that BBSM performs clearly better than DOWN.

The ROC and AROC measure the statistical discrimination capability of an ensemble system. They give information similar to the resolution part of the Brier skill score and can be applied to access the inherent value of a prediction system more directly (an AROC score larger than 0.5 indicates skill and a perfect score is 1). As for the other scores, the ROC and AROC scores shown in Figure 8 indicate the superiority of BBSM over DOWN.

5.2. *Verification of surface weather variables*

Since the main product of ALADIN-LAEF is the short-range forecast of surface weather variables, the verification

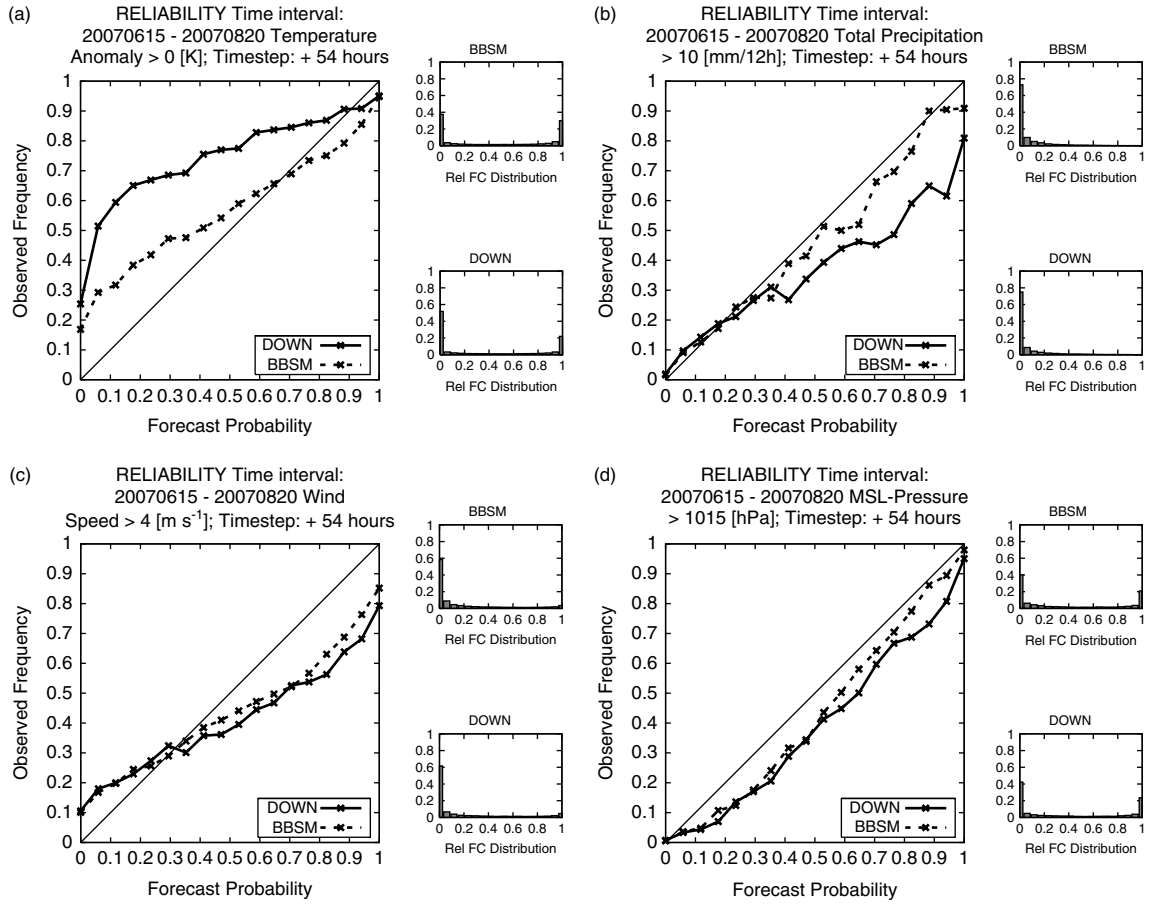

Figure 12. Reliability diagrams and sharpness of DOWN and BBSM for for forecast lead time +54 h for (a) T2m (threshold >0°C), (b) PREC (>10 mm (12 h)$^{-1}$), (c) W10m(>4 m s^{-1}), and (d) MSLP (>1015 hPa). The verification domain and period are as Figure 3.

is focused on 2 m temperature (T2m), 10 m wind (W10m), 12h accumulated rainfall (PREC) and mean sea level pressure (MSLP). Similar statistical scores to those in subsection 5.1 have been used to evaluate the forecast quality. Verification of the forecasts from DOWN and BBSM for June–August 2007 are shown in Figures 9–15.

5.2.1. Ensemble spread and RMSE of the ensemble mean

Figure 9 shows the time evolution of ensemble spread, bias and RMSE of the ensemble mean of T2m, W10m, PREC and MSLP for BBSM and DOWN. Overall, the spread of those surface variables is larger for BBSM than the spread for DOWN, and the RMSE of the ensemble mean is smaller for BBSM than for DOWN. The smaller discrepancy between ensemble spread and RMSE of the ensemble mean for BBSM implies the superiority of BBSM over DOWN concerning the reliability of the ensemble system. Clear improvements can be observed for the rainfall and near-surface temperature forecast, whereas the improvement on the RMSE of wind is almost negligible. It seems that the employment of NCSB, which is mainly related to the soil moisture and soil temperature, does not result in a clear improvement of the wind forecast. A diurnal variation of the errors of ensemble mean for both DOWN and BBSM is notable, i.e. larger error during the daytime and smaller error during the night-time, and is particularly true for T2m and MSLP (e.g. the dip in Figure 9(d)), which reveals the deficiency of ALADIN in the daytime. The strong cold bias in DOWN is removed largely in BBSM, which is due to the application of NCSB. Slightly improved bias is found in wind and surface pressure forecasts of BBSM, while the impact of BBSM on the bias of rainfall forecast is near neutral.

It is noted that the gain in statistical consistency for the surface variables from BBSM is not as large as for the upper-air variables. The ensemble spread does not match the error of the ensemble mean in surface forecasts of both BBSM and DOWN, which can be found at all lead times. The growth of the ensemble spread is determined by two factors, the initial condition perturbation and model perturbation. The obvious conclusion is that reasons are the lack of model perturbation to the relevant surface physical processes and the too small surface perturbation at initial time.

5.2.2. The Talagrand diagram

The reliability of BBSM and DOWN is evaluated in Figure 10 in terms of the Talagrand diagram for T2m, PREC, W10m and MSLP, valid at lead time +54 h. BBSM removes the bias present in DOWN, particularly in the forecast of T2m and MSLP (Figures 10(a,d)), and increases the spread of rainfall forecast compared to DOWN (Figure 10(b)). Concerning wind, small improvements in spread are found in BBSM (Figure 10(c)). Despite the larger spread of BBSM, both ALADIN-LAEF configurations are strongly underdispersive. This is different from the result for the upper-air variables shown in Figure 3, and is related to insufficient representation of surface IC perturbations and model error as discussed earlier (Figure 9).

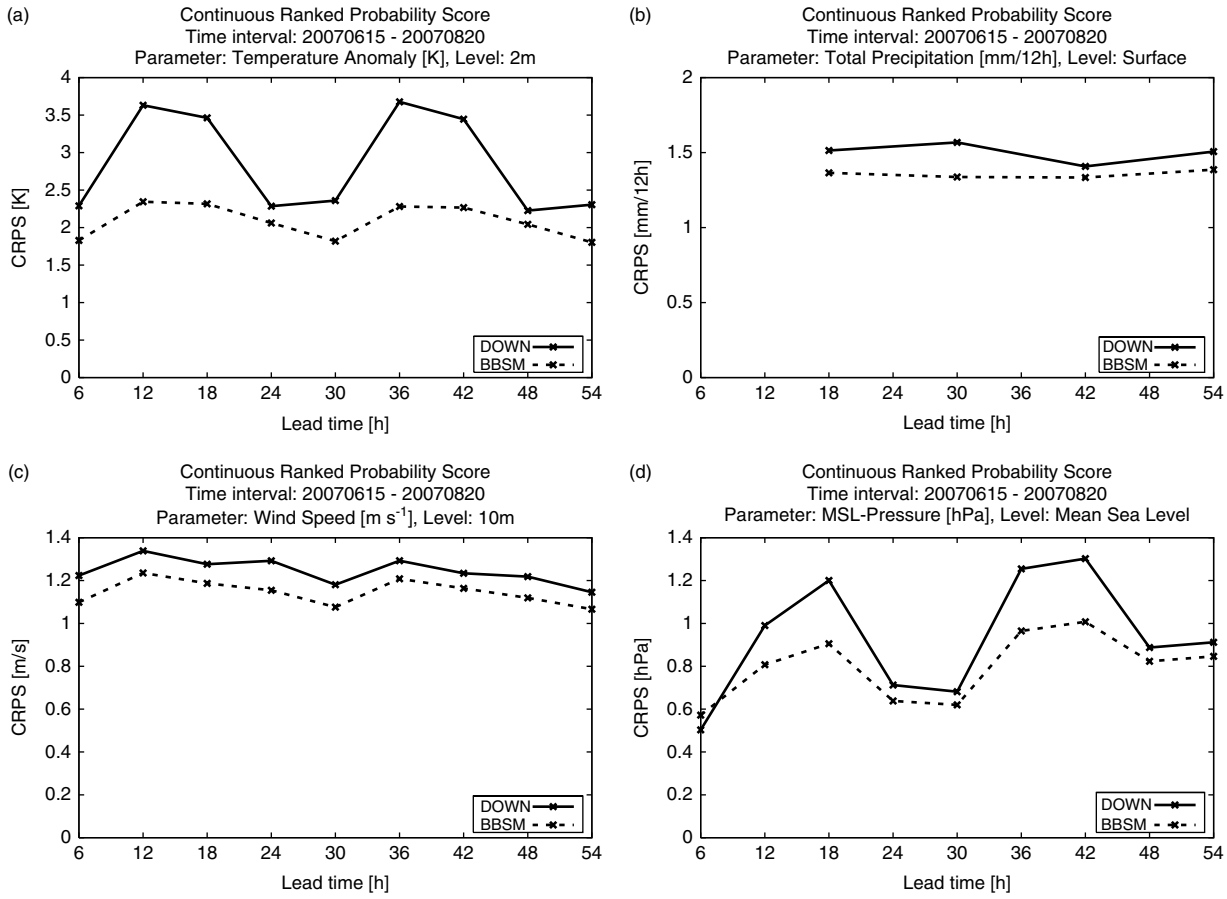

Figure 13. Continuous ranked probability score (CPRS) of DOWN and BBSM for forecasts of (a) T2m, (b) PREC, (c) W10m, and (d) MSLP. The verification domain and period are as Figure 3.

5.2.3. The outlier statistics

Figure 11 shows the outlier statistics for BBSM and DOWN. The differences between BBSM and DOWN in the outlier statistics are significant for all weather variables. BBSM forecasts are clearly more reliable than DOWN forecasts, consistent with the results in Figure 9.

5.2.4. The reliability diagram

Figure 12 presents the reliability diagrams for T2m, W10m, PREC and MSLP from the ALADIN-LAEF configurations BBSM and DOWN at lead time +54 h. The verifying events chosen are temperature anomaly $>0^{\circ}$C, wind speed $>4\,\mathrm{m\,s^{-1}}$, mean sea level pressure $>1015\,\mathrm{hPa}$ and rainfall $>10\,\mathrm{mm}\ (12\,\mathrm{h})^{-1}$. The reliability diagram of T2m (Figure 12(a)) indicates that the strong cold bias in DOWN is removed in BBSM. BBSM is more reliable and has better resolution than DOWN. This positive effect of BBSM is due to the use of surface initial condition perturbations (NCSB) in ALADIN-LAEF. The poor reliability of DOWN is caused by a cold bias in the forecast. This is a known problem in the surface coupling between ALADIN and ECMWF. The introduction of NCSB makes the surface coupling more consistent, by using the surface of a 12 h ALADIN forecast, initialized with the ARPEGE surface and driven by ECMWF EPS, as the perturbed surface analysis. Figure 12(b) shows the verification of 12 h accumulated rainfall forecasts at lead time +54 h for events larger than 10 mm $(12\,\mathrm{h})^{-1}$. The reliability curve of BBSM is closer to the diagonal than DOWN. As in the T2m forecast (Figure 12(a)), the rainfall forecasts of BBSM are more reliable, and have better resolution than DOWN. This means the combination of the IC perturbations, in particular the LAM native IC perturbations, with the multi-physics is beneficial for the probabilistic skill of the precipitation forecast.

BBSM also improves wind and MSLP, shown in Figures 12(c,d), but those improvements are not as distinct as for T2m and PREC.

5.2.5. CRPS and CRPSS

The skill of BBSM and DOWN is verified with CRPS and CRPSS, which is a proper measure of overall ensemble forecast performance. CRPSS measures the improvement of the infinite-category probabilistic forecast relative to a reference forecast. CRPSS is positively oriented, and is equal to 1 for a perfect system and it is zero for a system which has the same performance as the reference one. Negative values mean that the ensemble system is worse than the reference system. In this study the high-resolution deterministic forecast of ALADIN-Austria, which is operationally run at ZAMG with 9.6 km horizontal resolution and 60 levels in the vertical (Wang *et al.*, 2006), is used as reference in the CRPSS computation. This is based on the assumption that ALADIN-LAEF should be more skilful than the ALADIN-Austria forecast. Otherwise it would be difficult to justify the operational use of ALADIN-LAEF.

Figures 13 and 14 show the CRPS and CRPSS for T2m, PREC, W10m and MSLP, respectively. It is obvious that

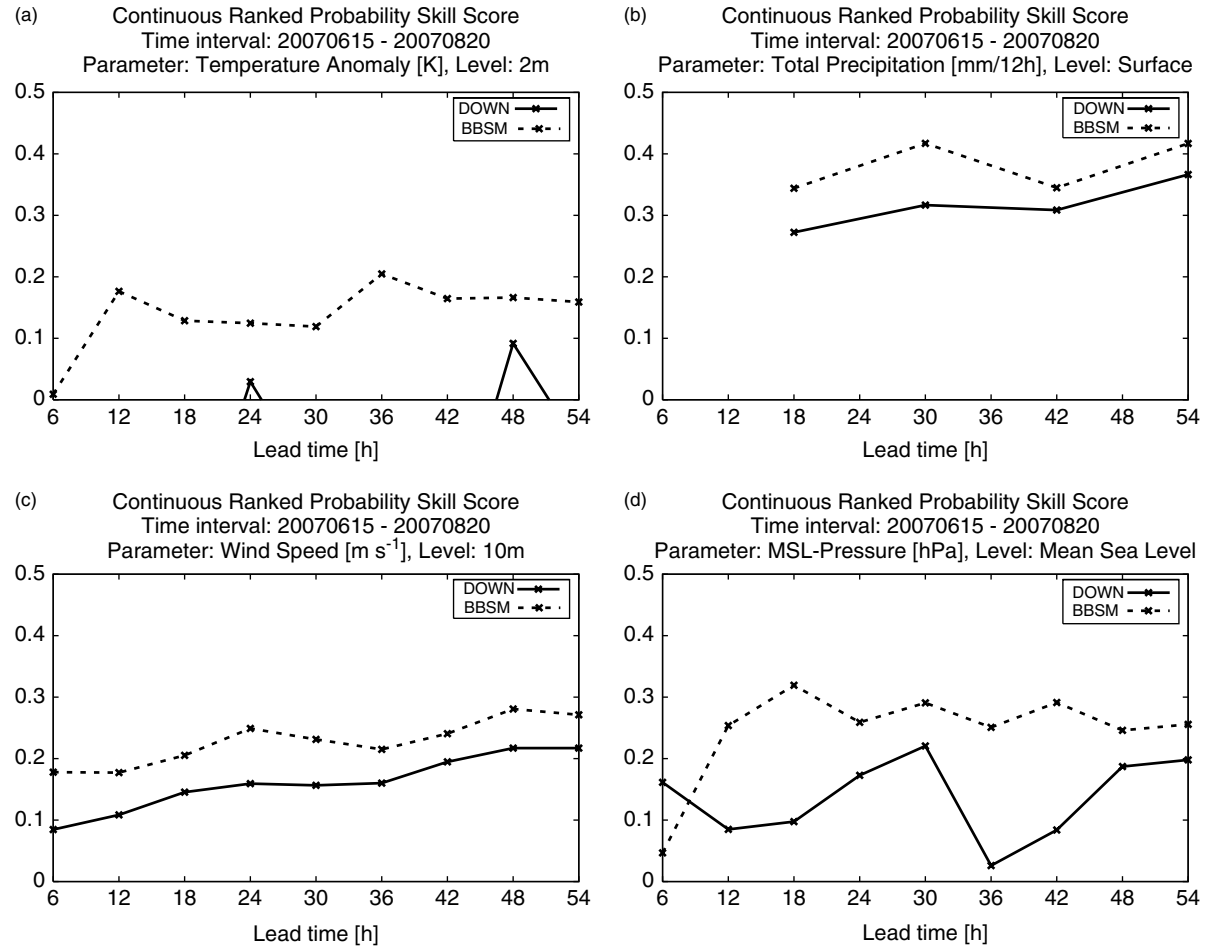

Figure 14. Continuous ranked probability skill score (CRPSS) of Down and BBSM for forecasts of (a) T2m, (b) PREC, (c) W10m, and (d) MSLP. The verification domain and period are as Figure 3.

BBSM has better performance (smaller errors) than DOWN in terms of CRPS, and is more skilful than DOWN in terms of CRPSS for all surface weather variables. Except for T2m of DOWN, all variables have higher skill than the reference for BBSM and DOWN. The lack of skill in T2m for DOWN relative to the high-resolution deterministic ALADIN-Austria forecast is mostly due to the inconsistency of the ECMWF surface IC used in DOWN with the surface parametrization schemes used in ALADIN. This inconsistency is most pronounced in the soil moisture and soil temperature and introduces the strong cold bias in T2m. In BBSM this cold bias was successfully removed by introducing NCSB.

5.2.6. ROC and AROC

The ROC score and the area under ROC for T2m, W10m, PREC and MSLP are shown in Figure 15. As for the upper-air variables, better discrimination characteristics are obtained by BBSM for the surface variables. As noted by Buizza *et al.* (2005), the AROC score provides a measure of the statistical discrimination capacity of the ensemble system, which is directly related to the inherent value of a forecasting system. The improvement of the discrimination in BBSM over DOWN is most probably caused by the introduction of the multi-physics scheme into ALADIN-LAEF.

6. Summary and conclusions

In this article, the new design of the Central European limited-area ensemble forecasting system ALADIN-LAEF has been described. Breeding–blending and NCSB methods have been applied for the generation of atmospheric initial condition perturbations and generation of surface initial condition perturbations, respectively. The multi-physics scheme has been introduced for representing the model-related uncertainty.

The blending method for atmospheric initial condition perturbation generation has a positive impact on the ALADIN-LAEF performance. The most important features of this method are:

- New initial conditions hold the large-scale uncertainty coming from the global model singular vector technique while small-scale perturbations generated by LAM breeding are retained as well.
- The short-wave part of the LAM perturbations actually contains more reliable information on mesoscale uncertainty than the pure mathematically interpolated short-wave part of the corresponding global perturbations; such scales are not resolved by the global model due to its mesh size.
- The blending method provides at least some physical consistency between the initial conditions and lateral boundary conditions used in ALADIN-LAEF.

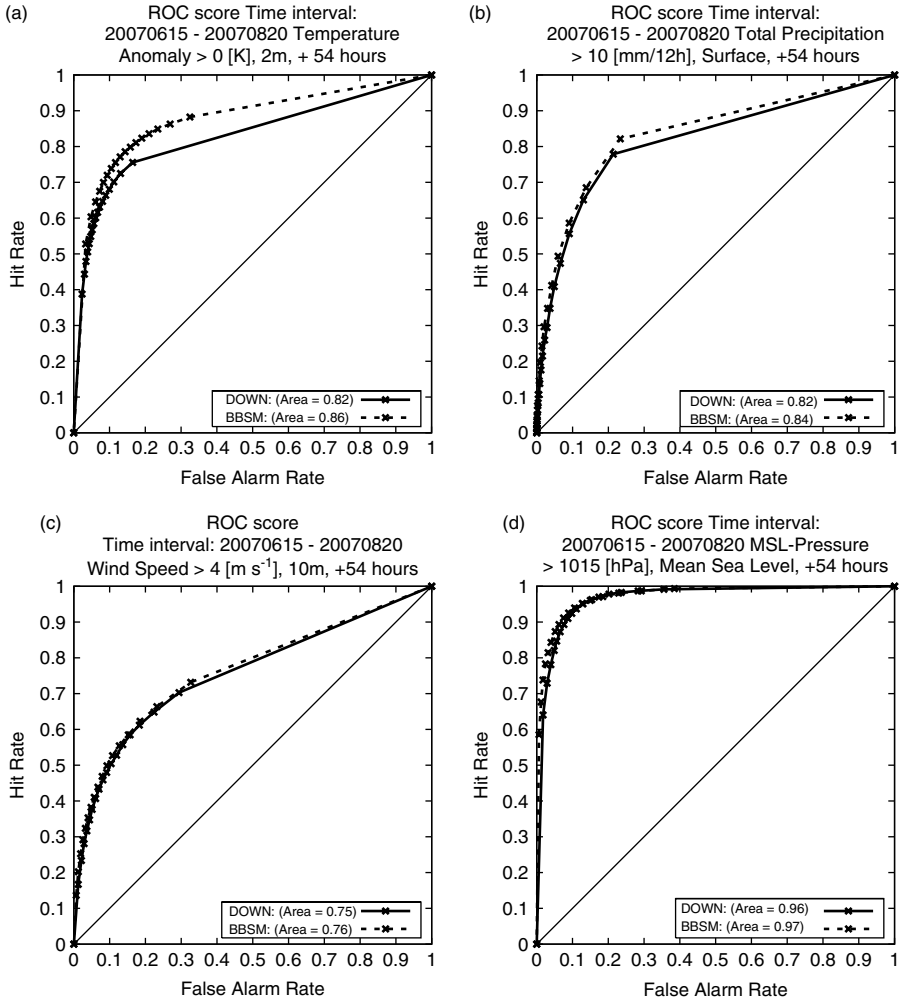

Figure 15. The relative operating characteristics (ROC) and area under ROC (AROC) of DOWN and BBSM for forecasts of (a) T2m, (b) PREC, and (c) W10m. The verification domain and period are as Figure 3, and the thresholds are as Figure 12.

NCSB is a new strategy for generating surface initial condition perturbations, in particular the perturbations of crucial parameters like soil moisture and soil temperature. It uses short-range LAMEPS forecasts driven by the global ensemble of ECMWF for generating the surface perturbation. The principle of NCSB allows its rather easy implementation within the operational LAMEPS. This task is not time-critical, but one has to keep in mind that it is not a very cheap method because it employs a series of 12 h LAM forecasts.

The implementation of different physical parametrizations and/or even different model versions for the integration of each ensemble member is an effective way for tackling the uncertainty related to the model error. Such a multi-physics approach provides generally better ensemble spread and improved reliability of the forecast, especially for rainfall. For instance, in the case of precipitation forecasts more scenarios can be realized within the ensemble, thanks to the different convection schemes.

The original ALADIN-LAEF system based on pure dynamical adaptation of ECMWF EPS was put into pre-operational mode in 2007. It was successfully upgraded with the implementation of breeding–blending, NCSB and multi-physics. The new system has been ported to the High Performance Computing Facility (HPCF) at ECMWF, where it has run operationally twice a day (0000/1200 UTC) since 2009.

The upgraded ALADIN-LAEF (denoted as BBSM experiment in the verification charts) has been verified against the pre-operational ALADIN-LAEF version (denoted as DOWN in the verification charts) over a two-month period in 2007. The verification has been carried out by comprehensive EPS verification scores considering the probabilistic forecast attributes like reliability, resolution, accuracy and discrimination. The deterministic forecast of ALADIN-Austria was used as a reference forecast in the skill score computations. The overall verification results are encouraging:

- There is a clear improvement of the quality of upper-air weather variables brought to the new system by implementation of blending and multi-physics.
- Remarkable benefits are achieved for the surface variables. This was possible due to the implementation of NCSB, which generates perturbed surface IC based on the ARPEGE surface analysis.
- Positive skill scores against deterministic ALADIN-Austria forecasts help to justify ALADIN-LAEF as a useful tool for supporting short-range weather forecasting services.
- The upgraded ALADIN-LAEF has generally more skill, better accuracy and is more reliable than the dynamical adaptation of global EPS.

However, unreliability and underdispersion (especially for some surface parameters) still remain the main

problems of ALADIN-LAEF. Possible reasons are inadequate representation of initial condition uncertainties at the surface and maybe in the lower troposphere. Another possible reason is the missing representation of uncertainties related to the model surface physics.

Hence, more effort will be put into addressing the underdispersion problems of surface weather variables in the near future. Work will be focused on better representation of uncertainties related to the model surface physics, possibly by introduction of stochastic surface parametrizations in ALADIN. Different strategies for generating surface initial condition perturbations will be investigated. Optimization of the multi-physics approach and use of ETKF/ET instead of breeding for generating the small-scale perturbations in the blending technique will be also carried out.

Acknowledgements

We gratefully acknowledge all the ALADIN/LACE/HIRLAM colleagues who have contributed to the development of ALADIN-LAEF. Edit Hagel and Richard Mladek have been involved in the development of the LAEF verification package. Special thanks go to Jun Du, Jean-François Geleyn, Craig Bishop, Xuguang Wang, Zoltan Toth, Jean Nicolau and Mozheng Wei, for fruitful discussions. Two anonymous reviewers provided constructive comments that helped improve the manuscript. ECMWF has provided the computer facilities and technical help to implement 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 29/2007).

References

Anderson JL. 1996. A method for producing and evaluating probabilistic forecast from ensemble model integrations. *J. Climate* **9**: 1518–1530.

Bechtold P, Bazile E, Guichard F, Mascart P, Richard E. 2001. A mass flux convection scheme for regional and global models. *Q. J. R. Meteorol. Soc.* **127**: 869–886.

Bellus M. 2008. '*Combination of large-scale initial conditions uncertainty with small-scale initial perturbations obtained by breeding cycling using blending technique in LAEF experiments*'. RC LACE internal report. Available online: http://www.rclace.eu/File/Predictability/2008/report_mbell_VIENNA_2008.PDF

Bishop CH, Toth Z. 1999. Ensemble transformation and adaptive observations. *J. Atmos. Sci.* **56**: 1748–1765.

Bishop CH, Holt TR, Nachamkin J, Chen S, McLay JG, Doyle JD, Thompson WT. 2009. Regional ensemble forecasts using the ensemble transform technique. *Mon. Weather Rev.* **137**: 288–298.

Bougeault P. 1981. Modeling the trade-wind cumulus boundary layer. Part I: Testing the ensemble cloud relations against numerical data. *J. Atmos. Sci.* **38**: 2414–2428.

Bougeault P, Geleyn JF. 1989. Some problems of closure assumption and scale dependency in the parameterization of moist deep convection for numerical weather prediction. *Meteorol. Atmos. Phys.* **40**: 123–135.

Bowler NE, Mylne KR. 2009. Ensemble transform Kalman filter perturbation for a regional ensemble prediction system. *Q. J. R. Meteorol. Soc.* **135**: 757–766.

Bowler NE, Arribas A, Mylne KR, Robertson KB, Beare SE. 2008. The MOGREPS short-range ensemble prediction system. *Q. J. R. Meteorol. Soc.* **134**: 703–722.

Brožková R, Klaric D, Ivatek-Šahdan S, Geleyn JF, Casse V, Siroka M, Radnoti G, Janousek M, Stadlbacher K, Seidl H. 2001. 'DFI Blending, an alternative tool for preparation of the initial conditions for LAM'. *PWRP Report Series No. 31* (CAS/JSC WGNE Report), *WMO-TD*, No. 1064, p 1.7.

Brožková R, Derkova M, Bellus M, Farda A. 2006. Atmospheric forcing by ALADIN/MFSTEP and MFSTEP oriented tunings. *Ocean Science* **2**: 113–121.

Buizza R, Palmer TN. 1995. The singular vector structure of the atmospheric general circulation. *J. Atmos. Sci.* **52**: 1434–1456.

Buizza R, Barkmeijer J, Palmer TN, Richardson DS. 2000. Current status and future development of the ECMWF ensemble prediction system. *Meteorol. Appl.* **7**: 163–175.

Buizza R, Houtekamer PL, Toth Z, Pellerin G, Wie M, Zhu Y. 2005. A comparison of the ECMWF, MSC, and NCEP Global Ensemble Prediction Systems. *Mon. Weather Rev.* **133**: 1076–1097.

Catry B, Geleyn JF, Tudor M, Benard P, Trojakova A. 2007. Flux-conservative thermodynamic equations in a mass-weighted framework. *Tellus* **59A**: 71–79.

Charron M. 2008. 'Development of a regional ensemble prediction system at Enviroment Canada'. In *Proceeedings of the First TIGGE-LAM Workshop,* 19–21 January 2008, Bologna, Italy. http://www.smr.arpa.emr.it/tiggelam/?document&1

Chen J, Xue JS, Yan H. 2005. A new initial perturbation method of ensemble mesoscale heavy rain prediction (in Chinese). *Chinese J. Atmos. Sci.* **5**: 717–726.

Cuxart J, Bougeault P, Redelsperger J-L. 2000. A turbulence scheme allowing for mesoscale and large-eddy simulations. *Q. J. R. Meteorol. Soc.* **126**: 1–30.

Derkova M, Bellus M. 2007. Various applications of the blending by digital filter technique in ALADIN numerical weather prediction system. *Meteorol. J. SHMU* **10**: 27–36.

Descamps L, Talagrand O. 2007. On some aspects of the definition of initial conditions for ensemble prediction. *Mon. Weather Rev.* **135**: 3260–3272.

Descamps L, Labadie C, Joly A. 2009. 'Status of the short-range ensemble prediction system at Météo-France'. *EMS Annual meeting 2009 and European conference on application of meteorology (ECAM)*, 28 Sept–2 Oct. 2009, Toulouse, France. http://www.emetsoc.org/annual_meetings/documents/2009/NWP2_EMS2009-212.pdf

Du J, Mullen SL, Sanders F. 1997. Short-range ensemble forecasting of quantitative precipitation. *Mon. Weather Rev.* **125**: 2427–2459.

Du J, Tracton MS. 1999. 'Impact of lateral boundary conditions on regional-model ensemble predicion. Research Activities in Atmospheric and Oceanic Modelling'. Ritchie H. (ed.) Report 28, CAS/JSC Working Group on Numerical Experimentation (WGNE), TD No. 942. World Meteorological Organization: Geneva. 67–68.

Du J, Tracton MS. 2001. 'Implementation of a real-time short-range ensemble forecasting system at NCEP: an update'. In *Preprints of 9th Conference on Mesoscale Processes,* Ft. Lauderdale, Florida. Amer. Meteorol. Soc: Boston. 355–356.

Du J, DiMego G, Tracton MS, Zhou B. 2003. 'NCEP short-range ensemble forecasting (SREF) system: multi-IC, multi-model and multi-physics approach'. In *Research Activities in Atmospheric and Oceanic Modeling*, J. Cote (ed.) Report 33, CAS/JSC Working Group on Numerical Experimentation, TD No. 1161. 509–510.

Eckel FA, Mass CF. 2005. Aspects of effective mesoscale, short-range ensemble forecasting. *Weather Forecasting* **20**: 328–350.

Ehrendorfer M. 1997. Predicting the uncertainty of numerical weather forecasts: A review. *Meteorol. Z. NF* **6**: 147–183.

Frogner IL, Haakenstad H, Iversen T. 2006. Limited-area ensemble predictions at the Norwegian Meteorological Institute. *Q. J. R. Meteorol. Soc.* **132**: 2785–2808.

García-Moya JA, Callado A, Escribà P, Santos C, Santos-Muñoz D, Simarro J. 2010. Predictability of short-range forecasting: A multi-model approach. *Tellus A.* Submitted.

Gebhardt C, Theis S, Krahe P, Renner V. 2008. Experimental ensemble forecast of prediction based on a convection-resolving model. *Atmos. Res. Lett.* **9**: 67–72.

Geleyn J-F. 1987. Use of a modified Richardson number for parameterizing the effect of shallow convection. *J. Meteorol. Soc. Japan WMO/IUGG NWP Symposium special issue:* 141–159.

Geleyn J-F. 1988. Interpolation of wind, temperature and humidity values from model levels to the height of measurement. *Tellus* **40A**: 347–351.

Geleyn J-F, Catry B, Bouteloup Y, Brožková R. 2008. A statistical approach for sedimentation inside a micro-physical precipitation scheme. *Tellus* **60A**: 649–662.

Gerard L. 2007. An integrated package for subgrid convection, clouds and precipitation compatible with the meso-gamma scales. *Q. J. R. Meteorol. Soc.* **133**: 711–730.

Gerard L, Geleyn J-F. 2005. Evolution of a subgrid deep convection parametrization in a limited-area model with increasing resolution. *Q. J. R. Meteorol. Soc.* **131**: 2293–2312.

Gerard L, Piriou JM, Brožková R, Geleyn J-F, Banciu D. 2009. Cloud and precipitation parameterization in a meso-gamma scale operational weather prediction model. *Mon. Weather Rev.* **137**: 3960–3977.

Giard D, Bazile E. 2000. Implementation of a new assimilation scheme for soil and surface variables in a global NWP model. *Mon. Weather Rev.* **128**: 997–1015.

Hamill TM, Colucci SJ. 1997. Verification of Eta/RSM short-range ensemble forecasts. *Mon. Weather Rev.* **125**: 1312–1327.
Hersbach H. 2000. Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather Forecasting* **15**: 559–570.
Houtekamer PL, Charron M, Mitchell HL, Pellerin G. 2007. 'Status of the Global EPS at Environment Canada'. In *Proceedings of Workshop on Ensemble Prediction*. ECMWF: Reading, UK. 57–68.
Jolliffe IT, Stephenson DB. 2003. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. J. Wiley and Sons: Chichester, UK.
Kalnay E. 2003. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press: Cambridge, UK.
Kessler E. 1969. On distribution and continuity of water substance in atmospheric circulation. *Meteorol. Monogr.* **10**.
Kuo HL. 1974. Further studies of the parameterization of the influence of cumulus convection on the large-scale flow. *J. Atmos. Sci.* **31**: 1232–1240.
Lenderink G, Holtslag AAM. 2004. An updated length-scale formulation for turbulent mixing in clear and cloudy boundary layers. *Q. J. R. Meteorol. Soc.* **130**: 3405–3427.
Leutbecher M, Palmer TN. 2008. Ensemble forecasting. *J. Comput. Phys.* **227**: 3515–3539.
Li X, Charron M, Spacek L, Candille G. 2008. A regional ensemble prediction system based on moist targeted singular vectors and stochastic parameter perturbation. *Mon. Weather Rev.* **136**: 443–462.
Lopez P. 2002. Implementation and validation of a new prognostic large-scale cloud and precipitation scheme for climate and data-assimilation purposes. *Q. J. R. Meteorol. Soc.* **128**: 229–257.
Louis JF, Tiedtke M, Geleyn J-F. 1981. 'A short history of the PBL parameterization at ECMWF'. In *Proceedings of Workshop on Planetary Boundary-layer Parametrization,* 25–27 November 1981, ECMWF: Reading, UK. 59–80.
Lynch P, Huang XY. 1992. Initialization of the HIRLAM model using a digital filter. *Mon. Weather Rev.* **120**: 1019–1034.
Marsigli C, Montani A, Paccagnella T. 2008. A spatial verification method applied to the evaluation of high-resolution ensemble forecasts. *Meteorol. Appl.* **15**: 125–143.
Mason I. 1982. A model for assessment of weather forecasts. *Austral. Meteorol. Mag.* **30**: 291–303.
Molteni F, Buizza R, Palmer TN, Petroliagis T. 1996. The ECMWF Ensemble Prediction System: Methodology and validation. *Q. J. R. Meteorol. Soc.* **122**: 73–119.
Molteni F, Marsigi C, Montani A, Nerozzi F, Paccagnella T. 2001. A strategy for high-resolution ensemble prediction. Part I: Definition of representative members and global-model experiments. *Q. J. R. Meteorol. Soc.* **127**: 2069–2094.
Morcrette J-J. 1991. Radiation and cloud radiative properties in the ECMWF operational forecast model. *J. Geophys. Res.* **96D**: 8121–9132.
Mullen SL, Buizza R. 2001. Quantitative precipitation forecast over the United States by the ECMWF ensemble prediction system. *Mon. Weather Rev.* **129**: 638–663.
Noilhan J, Planton S. 1989. A simple parameterization of land surface processes for the meteorological models. *Mon. Weather Rev.* **117**: 536–549.
Piriou JM, Redelsperger J-L, Geleyn J-F, Lafore J-P, Guichard F. 2007. An approach for convective parameterization with memory, in separating microphyscis and transport in grid-scale equations. *J. Atmos. Sci.* **64**: 4127–4139.
Ritter B, Geleyn J-F. 1992. A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. *Mon. Weather Rev.* **120**: 303–325.
Sass BH, Rontu L, Savijärvi H, Räisänen P. 1994. '*HIRLAM-2 Radiation scheme: Documentation and tests*'–. *HIRLAM Technical Report 16. http://hirlam.org/publications/TechReports/TR16ab.html*.
Savijärvi H. 1990. Fast radiation parameterization schemes for mesoscale and short-range forecast models. *J. Appl. Meteorol.* **29**: 437–447.
Smith RNB. 1990. A scheme for predicting layer clouds and their water content in a general circulation model. *Q. J. R. Meteorol. Soc.* **116**: 435–460.
Stanski HR, Wilson LJ, Burrows WR. 1989. *Survey of common verification methods in meteorology*. WWW Tech. Report 8, TD No. 358. World Meteorological Organization: Geneva.
Stappers R, Barkmeijer J. 2008. 'HIRLAM singular vectors: First results with CAPE-SVs'. In *Proceedings of Joint 18th ALADIN Workshop and HIRLAM ASM*, Brussels, Belgium. 7–10 April 2008. http://www.cnrm.meteo.fr/aladin/IMG/pdf/ASM2008stappers.pdf.
Stensrud DJ, Bao JW, Warner TT. 2000. Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems. *Mon. Weather Rev.* **128**: 2077–2107.
Sundqvist H. 1993. Inclusion of ice-phase of hydrometeors in cloud parameterization of large-scale and mesoscale models. *Contrib. Atmos. Phys.* **66**: 137–147.
Sutton C, Hamill TM, Warner TT. 2006. Will perturbing soil moisture improve warm-season ensemble forecasts? A proof of concept. *Mon. Weather Rev.* **134**: 3174–3189.
Talagrand O, Vautard R, Strauss B. 1997. 'Evaluation of probabilistic prediction systems'. In *Proceedings of Workshop on Predictability*. ECMWF: Reading, UK.
Toth Z, Kalnay E. 1993. Ensemble forecasting at NMC: The generation of perturbation. *Bull. Am. Meteorol. Soc.* **74**: 2317–2330.
Toth Z, Kalnay E. 1997. Ensemble forecasting at NCEP and the breeding method. *Mon. Weather Rev.* **125**: 3297–3319.
Tracton MS, Du J, Toth Z, Juang H. 1998. 'Short-range ensemble forecasting (SREF) at NCEP/EMC'. In *Proceedings of 12th Conference on Numerical Weather Prediction,* Phoenix, Arizona. Amer. Meteorol. Soc: Boston. 269–272.
Wang XG, Bishop CH. 2003. A comparison of breeding and ensemble Kalman filter ensemble forecast schemes. *J. Atmos. Sci.* **60**: 1140–1158.
Wang Y, Haiden T, Kann A. 2006a. 'The operational limited-area modelling system at ZAMG: ALADIN-Austria'. *Österreich. Beitr. Meteorol. Geophys.* **37**: ISSN 1016–6254.
Wang Y, Kann A, Ma X, Tian W. 2006b. 'Dealing with uncertainties in the initial conditions in ALADIN-LAEF'. In *Proceedings of 1st MAP D-PHASE scientific meeting*, 6–8 November 2006, Vienna, Austria. 10–13.
Wang Y, Kann A, Bellus M, Pailleux J, Wittmann C. 2010. A strategy for perturbing surface initial conditions in LAMEPS. *Atmos. Sci. Lett.* **11**: 108–113. DOI: 10.1002/asl.260.
Wilks DS. 2006. *Statistical Methods in the Atmospheric Sciences*. Academic Press: London.
Wyser K, Rontu L, Savijärvi H. 1999. Introducing the effective radius into a fast radiation scheme of a mesoscale model. *Contrib. Atmos. Phys.* **72**: 205–218.
Xu M, Randall D. 1996. A semi-empirical cloudiness parameterization for use in climate models. *J. Atmos. Sci.* **53**: 3084–3102.

