

# The impact of downscaled initial condition perturbations on convective-scale ensemble forecasts of precipitation

C. Kühnlein,<sup>a,b\*</sup> C. Keil,<sup>c</sup> G. C. Craig<sup>c</sup> and C. Gebhardt<sup>d</sup>

<sup>a</sup>Hans-Ertel Centre for Weather Research, Ludwig-Maximilians-Universität München, Germany

<sup>b</sup>European Centre for Medium-Range Weather Forecasts, Reading, UK

<sup>c</sup>Meteorologisches Institut, Ludwig-Maximilians-Universität München, Germany

<sup>d</sup>Deutscher Wetterdienst, Offenbach, Germany

\*Correspondence to: C. Kühnlein, European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading RG2 9AX, UK.  
E-mail: christian.kuehnlein@ecmwf.int

A major issue in convective-scale ensemble prediction systems (EPSs) is the specification of effective initial condition perturbations (ICPs). The present work considers the suitability of downscaled ICPs from a multi-model global EPS for short-range regional ensemble forecasts of convective precipitation at convection-permitting resolution. Previous studies indicated the importance of convective-scale initial condition uncertainties, with the most pronounced impact in weather conditions characterised by weak versus strong synoptic-scale forcing of convection. However, the downscaled ICPs do not explicitly represent small-scale uncertainty, which questions their effectiveness in convective-scale EPSs. To investigate the issue, the high-resolution ensemble system of the Deutscher Wetterdienst, COSMO-DE-EPS, which includes physics perturbations and lateral boundary condition perturbations in addition to ICPs, is employed. Forecasts are compared with a second EPS, identical but without ICPs, for a period of 3.5 months in the central European warm season. Weakly forced conditions are considered separately from strongly forced conditions, using an objective classification based on the area-averaged convective adjustment time-scale.

Generally for all EPSs, forecast quality measures show a distinct behaviour in strong versus weak forcing conditions. However, the impact of the ICPs is found to be similar in the two regimes. The impact of the ICPs is clearly largest and positive (consistently in terms of ensemble variance and probabilistic forecast quality, but negative for the equitable threat score) in the first six forecast hours when the ICPs dominate the physics perturbations and lateral boundary condition perturbations. The ICPs then decay relatively quickly with lead time as the physics perturbations and lateral boundary condition perturbations start to become important and later dominant. Probabilistic precipitation forecasts by the EPSs outperform the deterministic COSMO-DE at the same convection-permitting resolution, and this more strongly in the first nine forecast hours with the EPS applying the ICPs.

*Key Words:* limited-area ensemble prediction system; convection-permitting NWP; downscaled initial condition perturbations; COSMO model; convective adjustment time-scale

Received 30 October 2012; Revised 23 May 2013; Accepted 14 August 2013; Published online in Wiley Online Library 9 January 2014

## 1. Introduction

Effectively accounting for the various sources of uncertainty in high-resolution regional ensemble prediction systems (EPSs), particularly in the newest generation of models that explicitly simulate (deep) moist convection using  $\mathcal{O}(1\text{ km})$  horizontal grid spacings  $\Delta_h$  (Kong *et al.*, 2006; Clark *et al.*, 2009; Gebhardt *et al.*, 2011; Migliorini *et al.*, 2011), is a crucial but largely open issue. One question in such convection-permitting EPSs is the specification of ensemble perturbations to account for initial condition uncertainty. Established techniques in synoptic-scale global EPSs (e.g. Buizza *et al.*, 2005) require reconsideration due

to the fundamentally different dynamical processes (i.e. most notably baroclinic versus moist convective instabilities), rapid error propagation and growth, and higher nonlinearity which explicitly enter the problem at convective-scale numerical resolution (Zhang *et al.*, 2003; Hohenegger and Schär, 2007a,b). Moreover, the high-resolution EPSs are run as limited-area models (LAMs). The nested LAM configuration necessitates the consideration of lateral boundary condition uncertainty to avoid underdispersion of the ensemble, i.e. overconfident forecasts, associated with the boundary constraints on error growth (discussion in Saito *et al.*, 2012, and references therein). Therefore, lateral boundary condition perturbations (BCPs) are

implemented in LAM EPSs, typically by using the lateral boundary conditions from different members of a driving EPS run with coarser resolution. One resulting effect is that the impact of initial condition perturbations (ICPs) in the LAM EPSs can be overridden by the impact of the BCPs with lead time (e.g. Hohenegger *et al.*, 2008). Besides ICPs and BCPs, another major source of uncertainty, which ought to be considered especially in convection-permitting EPSs due to the numerous small-scale sensitivities (associated with turbulence, cloud microphysics, radiation, among others), is model error (Bouttier *et al.*, 2012).

One common approach to account for initial condition uncertainty in operational LAM EPSs is based on dynamical downscaling of an ensemble of coarser-resolution driving (global) model forecasts. An example is the Consortium for Small-scale Modelling–Limited-area Ensemble Prediction System (COSMO-LEPS; Marsigli *et al.*, 2005; Montani *et al.*, 2011), which interpolates forecast fields from a set of representative members of the European Centre for Medium-Range Weather Forecasts (ECMWF) global EPS to obtain different initial conditions for the regional domain with  $\Delta_h \sim 10$  km. Within the Met Office Global and Regional EPS (MOGREPS; Bowler *et al.*, 2008; Bowler and Mylne, 2009), the regional EPS members with  $\Delta_h \sim 24$  km obtain ICPs (here differences of global EPS members from the control, in contrast to the full fields) from short-range forecasts of its global EPS counterpart. Although methods specific to the global EPSs are adopted to generate the ICPs in these forecasting systems, small-scale perturbations are allowed to grow in the higher-resolution LAM EPS (Bowler and Mylne, 2009). Within MOGREPS, Bowler and Mylne (2009) compared the downscaled ICPs against ICPs generated by means of an ensemble transform Kalman filter run specifically for the regional ensemble, and found similar ensemble forecast skill in their tests with the different methods. Recently, downscaled ICPs have also been successfully applied in experimental convection-permitting LAM EPSs with  $\Delta_h \sim \mathcal{O}(1)$  km (Hohenegger *et al.*, 2008; Peralta *et al.*, 2012). In these convection-permitting ensembles, an intermediate-resolution LAM EPS is typically used to transfer the information in a chain of forecasts from the coarse-resolution global EPS to the high-resolution LAM EPS. Generally, the downscaling approach to generate ICPs is attractive due to its relative simplicity and practicality of implementation. An advantageous aspect is that the ICPs and BCPs typically originate from the same driving model forecast for each ensemble member, which avoids errors due to inconsistencies that may occur if the ICPs are applied independently from the BCPs (Caron, 2013).

However, the downscaling approach cannot comprehensively address initial condition uncertainties at the scales represented by the high-resolution LAM EPSs. Specifically, small-scale initial condition uncertainties in meso- and convective scales not considered by the coarse-resolution driving EPS may be important to the short-range convection-permitting forecasts of highly relevant parameters such as precipitation, low-level wind and temperature. In fact, recent results from studies by Vié *et al.* (2011) and Craig *et al.* (2012) indicate that convective-scale initial condition uncertainty is significant and accounting for this uncertainty has a clearly positive impact on the precipitation forecasts with convection-permitting EPSs. Moreover, both of these studies showed a much larger impact of the convective-scale ICPs in weather conditions lacking strong large-scale forcing of convection, when the formation of convective precipitation depends crucially on small-scale processes. Similar conclusions with regard to the impact of small-scale uncertainties were drawn earlier by Stensrud and Fritsch (1994a,b), who considered a case-study of mesoscale convective systems developing in a weakly forced large-scale environment. In another case-study, Leoncini *et al.* (2010) reported a significant impact of convective-scale model-state perturbations applied within the atmospheric boundary layer on the convection-permitting ensemble forecasts of precipitation. Hohenegger and Schär (2007b) showed that error propagation and growth in response to unbalanced ICPs

occurs through both sound and gravity waves (and/or numerical noise) which can relatively quickly ( $\sim 1$  h) impact the entire LAM domain compared to advective processes, followed by subsequent changes in moist convective instability.

The present work investigates the suitability of ICPs generated by means of a downscaling approach for the LAM EPS forecast of convective precipitation at convection-permitting numerical resolution. The underlying EPS for the present study is COSMO-DE-EPS (Gebhardt *et al.*, 2011; Peralta *et al.*, 2012) of the Deutscher Wetterdienst (DWD) which now provides operational ensemble forecasts for central Europe. In the current implementation of this system, a multi-model global EPS is downscaled via an intermediate-resolution ( $\Delta_h \sim 7$  km) LAM EPS, whose short-range forecasts are then used to create the ICPs around the high-resolution analysis ( $\Delta_h \sim 2.8$  km) which initialises the convection-permitting LAM EPS (Peralta *et al.*, 2012). Furthermore, the downscaled ICPs in COSMO-DE-EPS are combined with BCPs and with varying tunable parameters in selected physics parameterisation schemes (Gebhardt *et al.*, 2011), henceforth simply denoted as physics perturbations (PYPs). The approach of the present work is to study the ICP impact by comparing the full operational COSMO-DE-EPS (including the ICPs, BCPs, and PYPs) against the same EPS but without the ICPs (including only BCPs and PYPs). In addition, each of these two EPSs is compared to the deterministic forecast system COSMO-DE running at the same resolution.

The ICPs in COSMO-DE-EPS involves interpolation from the coarser-resolution driving model forecasts. Because the resulting ICPs are anticipated to represent variability of the larger-scale flow instead of individual convective clouds and generally do not address uncertainty specific to the high-resolution LAM analysis, their overall performance for the convective-scale precipitation forecasts is unclear. Particularly under weather conditions when convective-scale details of the atmospheric state are known to be more relevant (i.e. weak large-scale forcing), it is an open question whether the downscaled ICPs are effective in providing ensemble spread and improving the probabilistic forecast quality of the EPS in terms of convective precipitation. In order to thoroughly address these issues, the present study investigates the impact of the ICPs and the performance of the convection-permitting LAM EPS in a weather-regime-dependent way, by considering forecasts under weak forcing conditions of convective precipitation separately from forecasts performed under strong forcing conditions, using forecast data for a period of 3.5 months in the central European warm season 2011.

Note that, in contrast to downscaled ICPs investigated in the present work, techniques that explicitly quantify initial condition uncertainty within high-resolution LAM EPSs have been developed as well, and Saito *et al.* (2011) and Wang *et al.* (2011) provide recent overviews and discussions. With regard to convection-permitting LAM EPSs, the subject represents a vital field of research (e.g. Vié *et al.*, 2011; Migliorini *et al.*, 2011; Caron, 2013).

The outline of this article is as follows. Section 2 describes the ensemble forecasting system, the structure of the ICPs, and the investigation set-up. Section 3 presents the results from the evaluation of various ensemble forecast quality measures with a focus on total precipitation. Area averages and deterministic scores are used to show the different behaviour and performance of the ensemble forecasts in the distinct forcing regimes of convective precipitation. Afterwards, the magnitude and duration of the ICP impact (together with the PYPs and BCPs) on measures for ensemble variance and probabilistic forecast quality is investigated. A summary and discussion of the obtained results is provided in section 4.

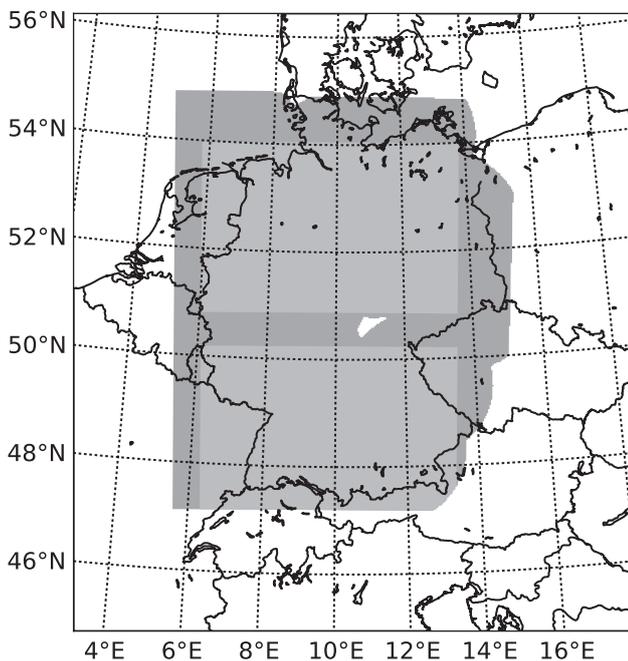
## 2. Data and methods

This section describes COSMO-DE-EPS, the spatial structure of the ICPs, and configurations of COSMO-DE-EPS applied in the

investigation. In addition, the section defines the investigation region and period, and also reviews briefly the convective adjustment time-scale which is used to separate weakly from strongly forced regimes of convective precipitation.

### 2.1. Convection-permitting ensemble prediction system

COSMO-DE-EPS is based on the regional convective-scale weather forecast system COSMO-DE which has been operational at DWD since 2007 (Baldauf *et al.*, 2011). COSMO-DE in turn represents a configuration of the non-hydrostatic COSMO model solving fully compressible equations with the split-explicit integration approach. In its current version, COSMO-DE covers central Europe with a rotated latitude–longitude grid of spacing  $\Delta_h \sim 2.8$  km ( $421 \times 461$  grid points over a domain of size  $1200$  km  $\times$   $1300$  km; Figure 1) and 50 vertically stretched model layers from 10 m above the ground to 22 km above mean sea level. COSMO-DE simulates deep convection explicitly but represents



**Figure 1.** The COSMO-DE domain (full) with the investigation region (dark grey shading). Investigation subregions over northern and southern Germany are shown by lighter grey shading. Political boundaries and coastlines are black solid lines.

**Table 1.** Configuration of COSMO-DE-EPS with 20 ensemble members. Rows denote four different subgroups of the EPS sharing ICPS+BCPs from one of the four members of the driving multi-model global EPS. Columns represent subgroups using the same perturbed physics parameter. The members of the particular subgroup with perturbed mixing length of turbulence (*tur\_len*) in the last column apply an additional perturbation of the latent heat nudging coefficient (*lhn\_coef*).

		IB1	IB2	IB3	IB4	IB5
		<i>entr_sc</i>	<i>q_crit</i>	<i>rlam_heat</i>	<i>rlam_heat</i>	<i>tur_len</i> + <i>lhn_coef</i>
PY1	IFS (ECMWF)	m1	m2	m3	m4	m5
PY2	GME (DWD)	m6	m7	m8	m9	m10
PY3	GFS (NCEP)	m11	m12	m13	m14	m15
PY4	GSM (JMA)	m16	m17	m18	m19	m20

**Table 2.** Full names, default values in COSMO-DE, and perturbed values in the IBP and BP EPSs of the varied parameters appearing in Table 1. The COSMO model manual (Schättler *et al.*, 2009) gives a more detailed description of these parameters.

Parameter	Description	Default	Perturbed
<i>entr_sc</i>	Entrainment rate for shallow convection	$0.0003 \text{ m}^{-1}$	$0.002 \text{ m}^{-1}$
<i>q_crit</i>	Critical value for normalised oversaturation	1.6	4.0
<i>rlam_heat</i>	Scaling factor of the laminar boundary for heat	1.0	0.1
<i>rlam_heat</i>	Scaling factor of the laminar boundary for heat	1.0	10.0
<i>tur_len</i>	Asymptotic mixing length of turbulence	150 m	500 m
<i>lhn_coef</i>	Latent heat nudging coefficient	1.0	0.5

effects of shallow convection through a parameterisation scheme (Baldauf *et al.*, 2011, and references therein give a thorough description of all model components).

COSMO-DE-EPS considers uncertainty from initial conditions, lateral boundary conditions and tunable parameters in certain physics parameterisation schemes (denoted as PYPs). As illustrated in Table 1, the EPS applies four different ICPS + BCPs combined with five different PYPs, resulting in 20 ensemble members. The PYPs alter selected parameters from their default value in COSMO-DE in a straightforward non-stochastic and spatially uniform manner. The PYP characterising a certain member is kept constant for all forecast start and lead times. The perturbed parameters named in Table 1 are described in Table 2. The parameters and their perturbed values are chosen with the aim to maximise warm season precipitation variability (Gebhardt *et al.*, 2011).

The four different lateral boundary conditions to COSMO-DE-EPS are provided by a four-member LAM EPS (named BC-EPS) based on the COSMO forecast model, using a grid spacing of  $\Delta_h \sim 7$  km and parameterised deep convection on the COSMO-EU domain (encompassing Europe, northern Africa, and the eastern Atlantic region; Baldauf *et al.*, 2011). All BC-EPS members employ the identical COSMO forecast model, i.e. no model perturbations are applied. The BC-EPS in turn is driven by a multi-model global EPS based on the deterministic forecasts of the Global Forecasting System (GFS) of the National Center of Environmental Prediction, the global model GME of DWD, the Integrated Forecast System (IFS) of ECMWF, and the Global Spectral Model (GSM) of the Japan Meteorological Agency. Global forecasts of GFS, GME, and GSM started at 0000, 0600, 1200, and 1800 UTC are considered except for the IFS forecasts which exist for start times 0000 and 1200 UTC only. The four-member BC-EPS is run every 6 h and the 20-member COSMO-DE-EPS every 3 h.

The COSMO-DE-EPS run started at time  $t$  is driven by the BC-EPS run started at  $t - 6$  h or  $t - 9$  h: the 0000/0300 UTC runs of COSMO-DE-EPS use the BC-EPS forecasts of 1800 UTC on the day before, the 0600/0900 UTC runs of COSMO-DE-EPS use the 0000 UTC run of BC-EPS, and so forth. Three members of the BC-EPS are driven by the GME, GFS, and GSM forecasts started at the same time, whereas one of the four BC-EPS members started at 0600/1200 UTC and 1800/0000 UTC is driven by the preceding 0000 and 1200 UTC IFS forecasts, respectively. The update frequency of the LBCs in the BC-EPS is 3 h, and in COSMO-DE-EPS it is 1 h, with linear interpolation in between. The BC-EPS initialisation is performed by spatial interpolation of the global forecast fields.

The BC-EPS forecasts are also used to generate the ICPs (Peralta *et al.*, 2012). The basic procedure is to calculate differences of the BC-EPS forecasts with the operational COSMO-EU forecasts (using the same grid spacing  $\Delta_h \sim 7$  km) always started 3 h before the initialisation of COSMO-DE-EPS. The differences are computed for the main prognostic variables: zonal and meridional velocity, temperature, specific humidity, and pressure perturbation. Then, these difference (or perturbation) fields are interpolated from the 7 km to the 2.8 km horizontal grid and finally added to the high-resolution COSMO-DE analysis fields at the initialisation time of the COSMO-DE-EPS forecasts. The resulting fields represent the same perturbed initial conditions within four subgroups of the 20 COSMO-DE-EPS members, consistent with the BCPs (rows in Table 1), i.e. ICPs and BCPs resulting from different global models are not intermixed.

A number of further steps are involved in the implementation of the ICPs, which are explained in detail in section 2.2 of Peralta *et al.* (2012). These include the application of a low-pass exponential filter on the perturbation fields which gradually increases from zero filtering outside the atmospheric boundary layer to complete damping of the ICPs near the surface (section 2.2 below explains how this affects spectra of perturbations), hydrostatic balancing of the perturbed pressure field in order to avoid spurious vertical accelerations, and the correction of possibly created unphysical values in the perturbed humidity fields. The low-pass exponential filter with strong damping of the lowest model layers is applied as the interpolation from the 7 to the 2.8 km horizontal grid can cause strong unrealistic gradients (especially close to the surface where terrain-following model levels are deformed due to orography) which then trigger spurious surface fluxes and humidity adjustments (Peralta *et al.*, 2012).

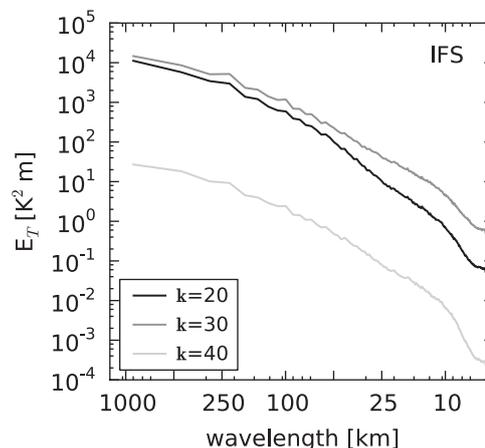
Finally, COSMO-DE-EPS also includes a perturbation in the latent heat nudging scheme employed in COSMO-DE for assimilation of high-resolution radar observations (Stephan *et al.*, 2008). However, the perturbation is restricted to the 'nudgecast mode' that relaxes the model forecast to radar observations made in the first hours after initialisation, i.e. the COSMO-DE analysis is not perturbed. It is implemented by means of a 50% reduction in the weight of the latent heat nudging coefficient and applied to a subset of four members (Tables 1 and 2); section 2.2.5 of Peralta *et al.* (2012) gives discussion. As noted in Peralta *et al.* (2012), the effect of the latent heat nudging perturbation on the (precipitation) forecasts is generally small.

As described, COSMO-DE-EPS became operational at DWD in May 2012.

## 2.2. Structure of initial condition perturbations

The generation of the downscaled ICPs in COSMO-DE-EPS involves interpolation of coarser-resolution driving model forecasts to the high-resolution LAM domain plus subsequent filtering of the perturbations at lower atmospheric levels (section 2.1). To illustrate the resulting spatial structure of the ICPs, power spectra, standard deviation and maximum amplitude of the perturbations are provided.

Figure 2 shows spatial spectra of the perturbation field of temperature  $T$  (i.e. full field minus the COSMO-DE analysis) on individual  $\sigma$  model levels at initialisation time as an example. The spectra are computed over a square region given by removing 50/70 grid points in longitude/latitude on each side of the COSMO-DE domain. One-dimensional spectra  $E_T(k)$  are obtained from 2D Fourier transforms on linearly detrended perturbation fields, and subsequent summation of the Fourier coefficients over annuli in wavenumber space. The individual spectra are averaged over the 3.5 month investigation period (section 2.4). The results in Figure 2 are shown only for one ensemble member where the ICPs stem from the IFS global model, as the computed spectra for other driving models appear almost identical. The spectra demonstrate that amplitudes at scales  $\sim 10$  km are two to three orders of magnitude smaller



**Figure 2.** Spectral variances of the ICP fields of temperature  $T$  on the COSMO-DE grid at the 0600 UTC initialisation time. The spectra are computed on various  $\sigma$  vertical model levels with index  $k$  ( $k=20, 30, 40$  correspond to heights  $\sim 7, 3, 0.8$  km). The horizontal grid spacing of the IFS global model is equivalent to  $\sim 16$  km, that of the BC-EPS (and COSMO-EU) is  $\sim 7$  km, and that of the COSMO-DE-EPS is  $\sim 2.8$  km.

than at  $\sim 100$  km. When comparing vertical model levels, the largest amplitudes of the ICPs are found around  $\sim 3$  km altitude. For model levels lower than  $\sim 1$  km altitude, the power decreases strongly throughout the entire spectrum (and becomes zero closer to the surface), as a result of the filtering of the ICPs.

The standard deviation (as an average over each model level) and the maximum amplitude (on each model level) of the ICPs have also been computed. As with the spectra already discussed, both quantities are calculated as an average over the investigation period. For model level  $k=30$ , where the largest power is present in the spectra, the standard deviation is within 0.5 and 0.6 K for the different COSMO-DE-EPS members. The average maximum amplitudes of the ICPs are between 2.7 and 3.1 K for the different members and thus are quite substantial.

From the perspective of the underlying convection-permitting forecast model, the ICPs can be characterised as large scale and are also strongly damped in the atmospheric boundary layer. The results indicate that the ICPs do not directly perturb the flow on scales of individual convective clouds, but might considerably modify the cloud environment due to their large amplitude.

## 2.3. EPS configurations used in this study

The two different EPSs compared in this work will be denoted as IBP (Initial, Boundary, Physics perturbations) and BP (Boundary, Physics perturbations). Note that the IBP and BP EPSs differ by the application of the ICPs, but also the latent heat nudging perturbation had to be switched off due to technical requirements in the BP EPS—however, the impact of the latter is negligible for the results in this article. Apart from that, the two EPSs are identical, i.e. all other aspects of COSMO-DE-EPS described in section 2.1 apply equally to the IBP and BP EPSs. In addition to the comparison of the two EPSs, the study evaluates the performance of both the IBP and the BP EPSs versus the deterministic forecast of COSMO-DE. Table 3 summarises the abbreviations used to describe the ensembles.

## 2.4. Region, period and verification data

This study considers a continuous investigation period in the European warm season from 1 May to 15 August 2011. During this time, COSMO-DE-EPS in the IBP configuration was pre-operational at DWD; the BP EPS was run additionally. Except for two days (12 and 13 July 2011) when technical problems did not allow completion of all computations, our dataset comprises 105 days of forecasts with the full 20 members for both the IBP and the BP EPSs. Data for COSMO-DE forecasts were routinely

Table 3. Summary of abbreviations used to describe the ensembles.

Abbreviation	Full name and description
EPS	Ensemble prediction system
ICP	Initial condition perturbation
BCP	Lateral boundary condition perturbation
PYP	Physics perturbation
DE	Deterministic forecast model COSMO-DE
BP EPS	EPS using lateral boundary condition and physics perturbations
IBP EPS	EPS using initial condition, lateral boundary condition, and physics perturbations

Table 4. Statistics of the classification into forecasts under strong and weak forcing conditions. The fourth and fifth columns give, respectively, the number of entire forecasts not included in the evaluation due to the absence of significant precipitation and the number of missing days due to technical problems of the EPSs. Numbers in columns two to four are for COSMO-DE with initialisation time 0600 UTC.

	Strong: $\langle \tau_c \rangle_{\max} < 6 \text{ h}$	Weak: $\langle \tau_c \rangle_{\max} > 6 \text{ h}$	'Dry'	Missing	Total
Number of forecasts	83	16	6	2	107

available from the operational system. The present study focuses on forecasts initialised at 0600 UTC (0800 local time).

The investigation region for the forecasts is chosen consistent with the area coverage of DWD's high-resolution radar network used for the verification. Figure 1 shows that this subregion of the full COSMO-DE domain covers most of Germany, parts of the neighbouring countries and adjacent seas. The investigation region is far away from the COSMO-DE domain boundaries, so that issues associated with the boundary forcing (e.g. damped convection/precipitation associated with interpolation and relaxation zones) should not affect the evaluation too strongly.

The verifying observations are obtained from DWD's network of 16 Doppler radars. For each radar, near-surface reflectivities are available every 5 min with a spatial resolution of 1 km in range and a resolution of 1° in azimuth (Stephan *et al.*, 2008). An efficient quality control for the single scans has been recently implemented (Helmert *et al.*, 2012). The reflectivities are converted to precipitation rates by empirical  $Z-R$  relations. Radar estimates at the individual grid points possibly affected by bright-band effects are filtered.

In section 3, an evaluation of the ensemble precipitation forecasts is presented using area averages, deterministic scores, as well as measures for ensemble variance and probabilistic forecast quality. All these quantities are consistently evaluated at the COSMO-DE grid points of the investigation region (Figure 1) where valid radar observations are available at a particular lead time. This includes also the quantities which do not directly involve the radar observations, e.g. measures for ensemble variance or the computation of the convective adjustment time-scale which is used as the indicator for the prevailing weather regime (section 2.5).

Throughout this article, hourly model output data of total precipitation is considered in the evaluation of the forecasts. Output times of the deterministic COSMO-DE forecasts were completely excluded from the evaluation when a precipitation rate  $P$  of  $1 \text{ mm h}^{-1}$  was not attained by a minimum number of 100 grid points in the investigation region, simultaneously for all forecasting systems. In total, these amount to  $\sim 8\%$  of the underlying data.

### 2.5. Indicator for prevailing weather regime

The convective adjustment time-scale  $\tau_c$ , defined as the ratio of convective available potential energy (CAPE) over the rate of change of CAPE, has been shown to be a useful indicator for the prevailing type of forcing of convective precipitation and

different predictability (Done *et al.*, 2006; Molini *et al.*, 2011; Keil and Craig, 2011; Zimmer *et al.*, 2011; Craig *et al.*, 2012; Keil *et al.*, 2013). The idea is that the time-scale  $\tau_c$  provides a measure of how rapidly conditional instability is removed from the atmosphere through the release of moist convection. If the time-scale  $\tau_c$  is much smaller than the time-scale over which the synoptic environment evolves, convection (considered as an average over an ensemble of clouds) is in equilibrium with, and predominantly controlled by, the synoptic-scale forcing. In contrast, if  $\tau_c$  is large, then convection is supposedly inhibited due to the lack of synoptic-scale forcing (i.e. non-equilibrium with the environment), and is only released subject to the availability of some local triggering mechanism.

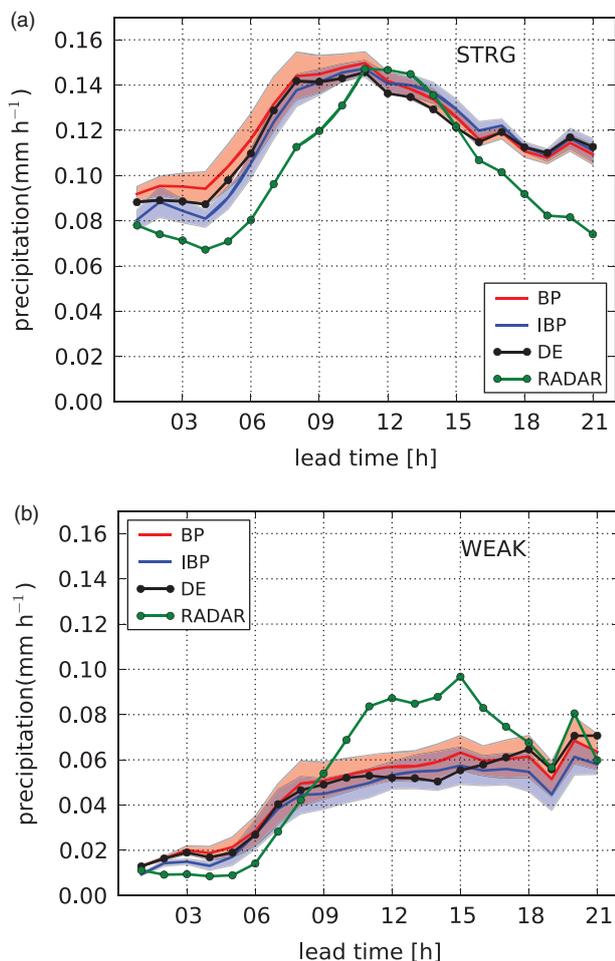
In section 3, the convective adjustment time-scale  $\tau_c$  is employed to distinguish between forecasts performed in different meteorological conditions. Specifically, the interest is in isolating typical warm season weak forcing days with a pronounced diurnal cycle in convective precipitation from the days when the latter is more controlled by large-scale weather systems, and in investigating how the EPSs perform during weak versus strong forcing conditions. Similar to previous studies (e.g. Done *et al.*, 2006; Craig *et al.*, 2012; Keil *et al.*, 2013), the spatial average of  $\tau_c$  is considered, here over the specified investigation region (Figure 1). If the spatial average of  $\tau_c$  for the hourly model output, denoted as  $\langle \tau_c \rangle$ , exceeds once a day the threshold of 6 h, then this day is denoted as *weak forcing*. This criterion is motivated on the observation that during weak forcing conditions the diurnal cycle typically shows large values of  $\langle \tau_c \rangle$  initially, i.e. non-equilibrium, but these relax towards equilibrium values in the course of the day as conditional instability is released (Done *et al.*, 2006; Molini *et al.*, 2011). Contrary to weak forcing, if  $\langle \tau_c \rangle$  shows no value exceeding the 6 h threshold then the respective day is classified as *strong forcing*.

All further details about the computation of  $\tau_c$  from the hourly forecast fields of CAPE and precipitation rate  $P$  (the rate of change of CAPE is estimated from the column-integrated latent heat release, which is proportional to  $P$ ) are as described in Craig *et al.* (2012) and Keil *et al.* (2013), and not repeated in this work. Here,  $\tau_c$  is derived from the deterministic forecast of COSMO-DE, and this value of  $\tau_c$  is then used to perform consistently the weak/strong forcing classification for the forecasts of COSMO-DE as well as the two considered EPSs IBP and BP.

Table 4 provides the resulting statistics according to the weather-regime-dependent classification for the entire investigation period. It shows that, for the majority of forecasts (or days), precipitation processes are linked to strong large-scale forcing. The number of forecasts with weak forcing conditions are  $\sim 20\%$  of the strong forcing cases. This relative number of strong to weak forcing cases is equal to the study of Keil *et al.* (2013), who applied a similar approach based on the convective adjustment time-scale to study the weather-regime-dependent predictability of convective precipitation using an experimental COSMO-DE-EPS (with BCPs and PYPs) for 88 days in the warm season of 2009. As given in Table 4, six forecasts showed no significant precipitation at all (i.e. the threshold of 100 grid points of the investigation region with  $P > 1 \text{ mm h}^{-1}$  defined at the end of section 2.4 was not exceeded for all lead times of these forecasts), which thus did not allow a reasonable estimate of  $\langle \tau_c \rangle$  to be computed. Correspondingly, these 'dry' forecasts are not considered in the evaluation.

### 3. Weather-regime-dependent evaluation of the EPS forecasts

In this section, results are presented from the evaluation of COSMO-DE-EPS over the 3.5 month investigation period. Because the improvement of quantitative precipitation forecasts is one of the major objectives of COSMO-DE-EPS and short-range ensemble systems in general, the focus is on the forecasted hourly precipitation rates  $P$ . The area-averaged convective adjustment time-scale is employed to objectively distinguish between forecasts (or days) dominated either by strong or weak large-scale forcing



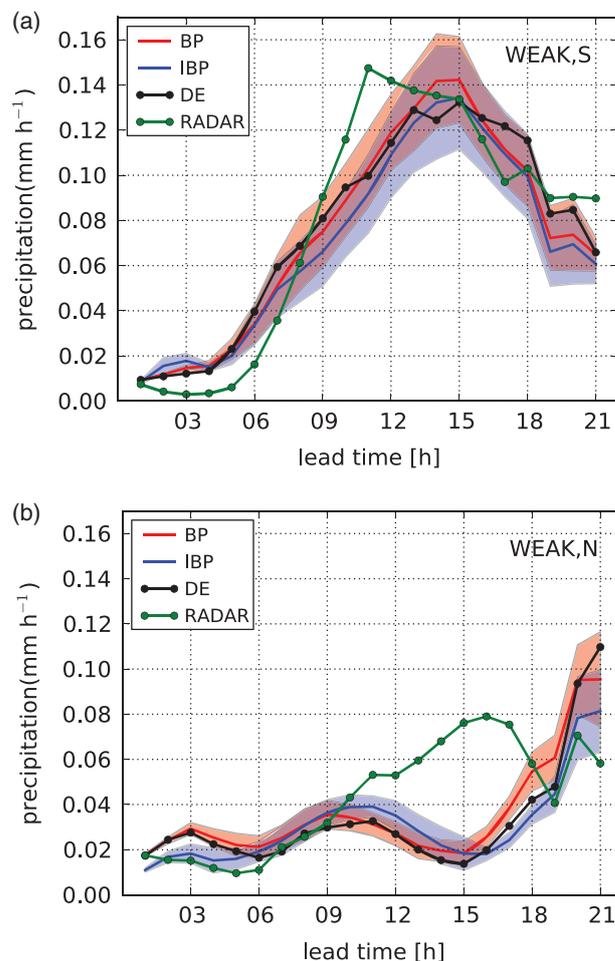
**Figure 3.** Spatially averaged hourly precipitation rate as function of lead time from initialisation time 0600 UTC (0800 local time). The spatial average is over the investigation region (Figure 1). In addition, averages over forecasts under (a) strong and (b) weak forcing conditions of the investigation period are shown. The red and blue solid lines represent, respectively, the ensemble mean of the BP and IBP EPSs, with the shaded regions in lighter colours the corresponding ensemble standard deviation. The black line is for COSMO-DE and the green line for the verifying radar observations.

weather regimes of convective precipitation, as explained in section 2.5. Regime-independent results will not be presented because these appear very similar to the results obtained for strong forcing, explained by the relatively high number of samples in the latter case (Table 4).

### 3.1. Spatially averaged hourly precipitation

The spatially averaged hourly precipitation rate as a function of lead time is displayed in Figure 3. The classification into days with strong or weak forcing conditions reveals the different characteristics of precipitation. Throughout the day the observed radar-derived precipitation rates are larger on strong than on weak forcing days. Both regimes show the maximum values in the local afternoon and evening hours (the 12 h forecast lead time in Figure 3 corresponds to 2000 local time or 1800 UTC). As expected, the diurnal cycle is especially pronounced during the weak forcing days, with a low precipitation rate in the morning hours and a substantial relative increase of the precipitation rate throughout the period of major convective activity in the afternoon and evening hours. Recall, however, that forecast hours (and observations) showing no significant precipitation are excluded from the evaluation (section 2.4 gives the applied threshold criterion). For example, this can explain the increase of precipitation during the night hours under weak forcing.

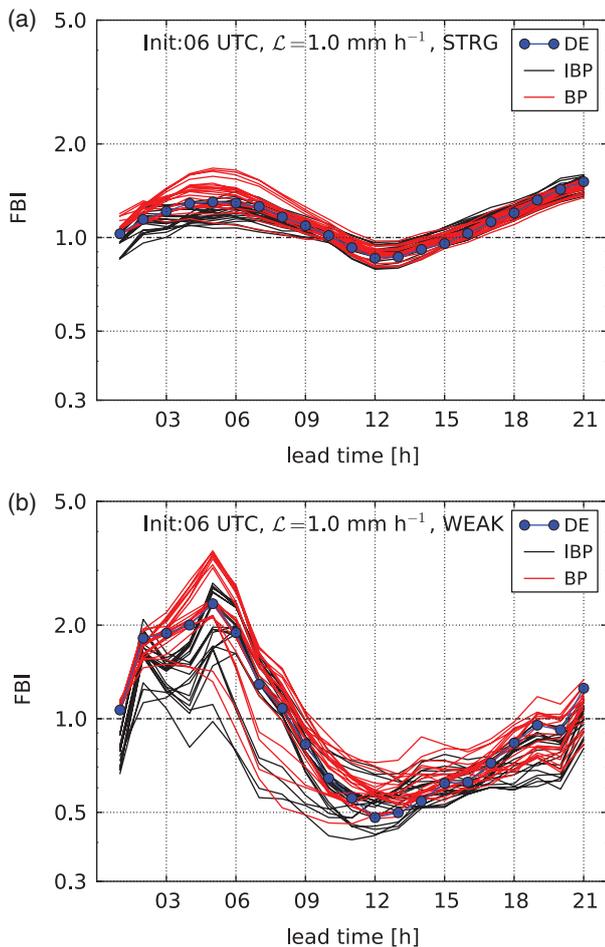
The forecasted ensemble mean of the average precipitation rate by the two EPSs (IBP and BP) and the deterministic COSMO-DE show qualitatively similar behaviour in both regimes: they



**Figure 4.** Spatially averaged hourly precipitation rate as function of lead time from initialisation time 0600 UTC (0800 local time). The spatial average is over (a) southern and (b) northern Germany. The temporal average is over all forecasts under weak forcing conditions. Other details are as Figure 3.

overestimate the observed values under strong forcing conditions and underestimate them in the weak forcing conditions. Generally, the BP EPS tends to produce more precipitation than the IBP EPS. Under strong forcing, this results in a lower overestimation for the IBP versus the BP EPS especially in the first 6–9 h of the forecast. The difference between the IBP and BP EPSs is less clear under weak forcing. The representation of the precipitation maximum in the diurnal cycle is poor (i.e. ~30–40% underestimation) with all forecasting systems.

The spread of the spatially averaged precipitation for the IBP ensemble members (Figure 3) shows larger values under weak versus strong forcing conditions with the onset of major convective activity in the diurnal cycle from ~6 h lead time (1400 local time) onwards. The spread for the BP EPS unexpectedly indicates larger values than for the IBP EPS during the first forecast hours under strong forcing conditions (Figure 3(a)). For the BP EPS, the larger spread (and ensemble mean) can be attributed particularly to a subgroup of the BP ensemble members driven by the same global model (GSM) that produces significantly more precipitation than the other members of the EPS; the respective ensemble subgroup can be seen in Figure 5(a). We presume that this behaviour is related to the inconsistent boundary condition forcing from the driving coarser-resolution model forecasts (e.g. Warner *et al.*, 1997), with larger effects for certain driving models and under strongly forced conditions. The ICPs improve the consistency with the driving model forecasts, and our results suggest that this provides smaller errors in the predicted precipitation amounts. Note that the lower precipitation amounts in the IBP EPS cannot be explained by a straightforward impact of the ICPs on the vertical thermal structure. For example, the initial CAPE values are about 5–10% larger in the IBP EPS than in the BP EPS on average, which supports the opposite effect.

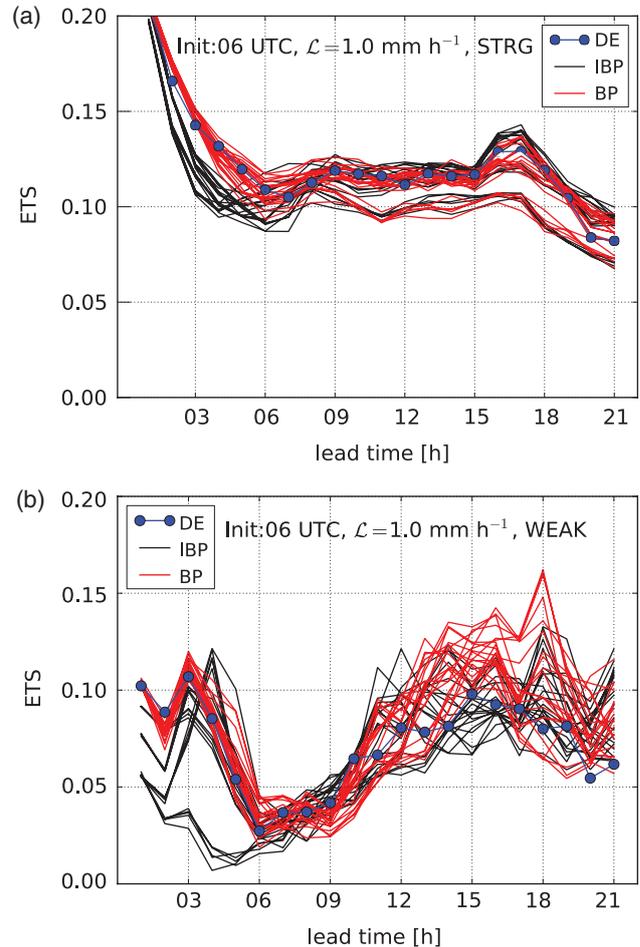


**Figure 5.** Frequency bias index (precipitation rate threshold is  $\mathcal{L} = 1 \text{ mm h}^{-1}$ ) as an average over forecasts under (a) strong and (b) weak forcing conditions. The black and red lines represent the individual members of the IBP and BP EPSs, respectively, and the blue line with circles gives the result for COSMO-DE. The  $y$ -axis for the FBI uses a logarithmic scale.

Some insight into the nature of the systematic underestimation of precipitation in weak forcing conditions can be obtained by considering the role of orography. For this purpose, the spatially averaged precipitation has additionally been computed for two smaller subregions of the main investigation region located in the southern and northern part of Germany (Figure 1). The respective results for southern Germany shown in Figure 4(a) depict a fairly good diurnal cycle representation of the averaged precipitation, in particular the evening precipitation maximum. In contrast, the results for northern Germany in Figure 4(b) show an even worse representation than Figure 3(b), with a wrong precipitation minimum in the evening (i.e.  $\sim 60\text{--}70\%$  underestimation). The forecasting system fails to simulate the diurnal variation of convective precipitation over the relatively flat terrain characterising the northern region, but benefits from enhanced predictability most likely associated with orographic forcing (Anthes *et al.*, 1985) in the relatively mountainous southern region. The same behaviour of the precipitation forecasts has been observed in the predictability case-study by Walser *et al.* (2004) using the Canadian Mesoscale Compressible Community (MC2) model at similar resolution. This suggests a systematic problem of diurnal cycle representation of convective precipitation (i.e. biased towards significantly underestimating precipitation) in regions away from significant orography by current convection-permitting forecast models.

### 3.2. Deterministic scores

Two standard, widely used deterministic scores, the frequency bias index (FBI) and the equitable threat score (ETS) (Wilks,



**Figure 6.** As Figure 5, but for equitable threat score.

2006), are used to evaluate the relative skill of all IBP and BP ensemble members as well as the single COSMO-DE forecast, again as a function of lead time. For both the FBI in Figure 5 and the ETS in Figure 6, the behaviour is again very different in the different meteorological regimes. The FBI shows much larger error amplitudes (a perfect FBI would be 1; FBI values larger than 1 indicate an overestimation of the forecasted to the observed area where the precipitation rate  $P > \mathcal{L}$  with threshold  $\mathcal{L} = 1 \text{ mm h}^{-1}$ , and vice versa) and also a much larger spread between the individual ensemble members under weak compared to strong forcing conditions. The ETS depicts a considerably lower skill (the higher the ETS the better) under weak forcing conditions. The latter applies primarily in the first 9–12 h of the forecast when moist convection starts to develop in the diurnal cycle. The particularly low values of the ETS between 6 and 9 h are likely associated with the forecasted convection occurring at wrong locations, because the spatially averaged hourly precipitation (Figure 3(b)) and FBI (Figure 5(b)) indicate relatively good representation of precipitation during this time.

Differences between the IBP and BP ensemble members are generally most pronounced until 6–9 h of forecast lead time. The results for the FBI are similar to the results of the spatially averaged hourly precipitation in Figure 3, showing a positive impact of the ICPs in the first hours in both regimes, and a slightly negative impact at forecast times  $> 12$  h under weak forcing conditions.

The ICP impact on the ETS is mostly negative for forecast times  $< 6$  h. This result, however, could be expected because the perturbed initial conditions of the IBP ensemble members differ from the single 'optimal' initial conditions of the COSMO-DE analysis by which the COSMO-DE forecast and all BP ensemble members are initialised. The results also reveal a relatively quick and strong response to the large-scale ICPs under weak forcing conditions. Already during the first forecast hour, a spreading of the IBP ensemble members into the four subgroups sharing the

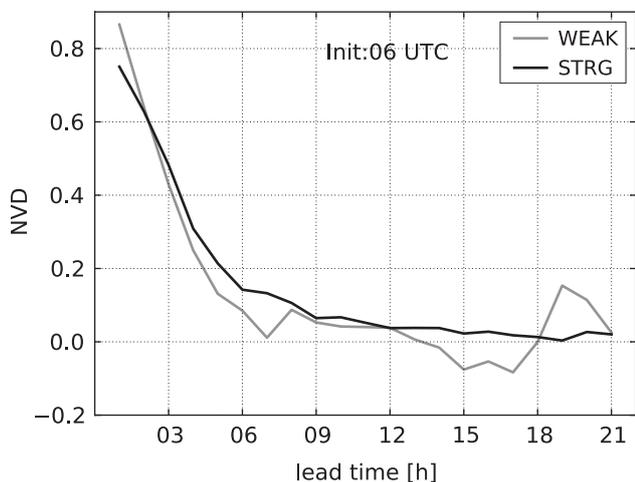


Figure 7. Normalised variance difference defined in Eq. (1) as an average over forecasts with strong (black line) and weak (grey line) forcing conditions.

same ICPs is apparent for the FBI and the ETS. The spreading between the ensemble members for both measures tends to be smaller in strong forcing conditions, and is not observed for the BP ensemble members and COSMO-DE which all lie closely together initially.

3.3. Initial condition perturbation impact on ensemble variance

A main motivation for the inclusion of ICPs in COSMO-DE-EPS is to improve its ensemble variance (Peralta *et al.*, 2012). In the following, we examine whether the ICPs fulfil this purpose considering the different forcing regimes, using the normalised variance difference (NVD; Gebhardt *et al.*, 2011). Here, the NVD considers the difference between the variances (in terms of the hourly precipitation rate  $P$ ) of the IBP  $\sigma_{IBP}^2(P)$  and BP  $\sigma_{BP}^2(P)$  EPSs, normalised by the sum of the variances:

$$NVD(P, \tau) = \frac{\sigma_{IBP}^2(P, \tau) - \sigma_{BP}^2(P, \tau)}{\sigma_{IBP}^2(P, \tau) + \sigma_{BP}^2(P, \tau)}. \quad (1)$$

The NVD is calculated as a function of forecast lead time  $\tau$ , applying averaging over all grid points of the investigation region plus separate averaging over all strong and weak forcing days of the investigation period. A positive (negative) NVD indicates a positive (negative) impact of the ICPs on the ensemble variance. The NVD is zero if the ICPs have no impact on the variance.

As shown in Figure 7, a positive impact of the ICPs on the ensemble variance is generally found in both forcing regimes. The gain in variance through the inclusion of the ICPs is clearly largest in the first forecast hours and then decays with lead time. In contrast to the quantities discussed in the previous sections, which showed a very different behaviour in the distinct forcing regimes, the ICP impact on the variance of the considered EPS is similar under strong and weak forcing conditions, only the decay with lead time tends to occur somewhat faster in weakly forced conditions from an initially slightly larger impact. At around 16 h into the forecast, a small negative impact is present under weak forcing.

The more variable NVD in the weak forcing might be attributed to some extent to the lower number of samples than in the strong forcing conditions (Table 4), but also to the different characteristics of precipitation in the distinct regimes, i.e. typically lower overall amounts with precipitating convective cells appearing more localised and intermittent in weakly forced situations.

3.4. Relative impact of different perturbations on ensemble variance

The particular design of COSMO-DE-EPS with the fixed arrangement of ensemble perturbations in certain subgroups, either using

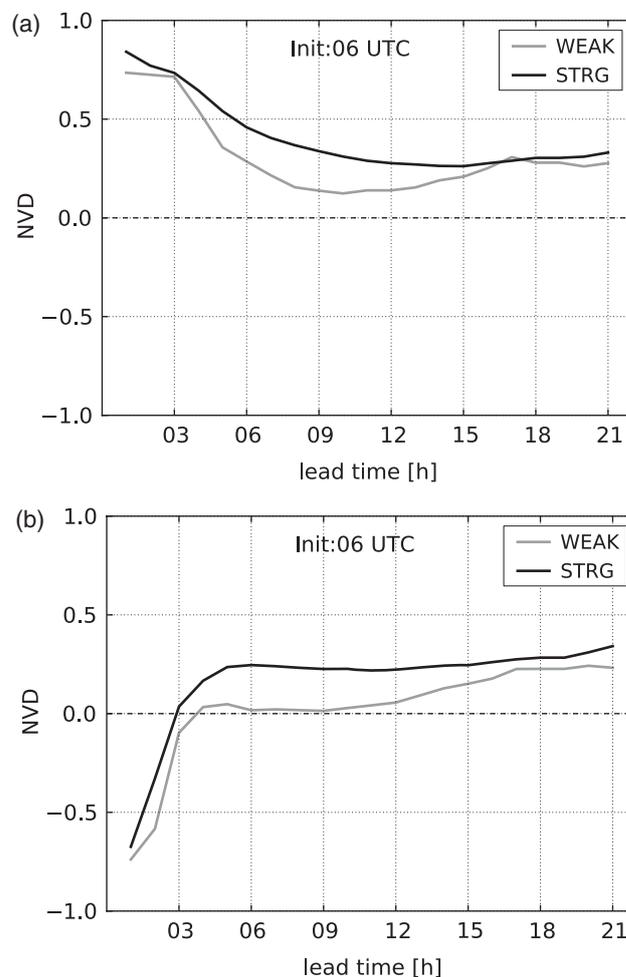


Figure 8. Normalised variance difference defined in Eq. (2) for the (a) IBP and (b) BP EPSs.

the same initial and boundary condition perturbations (ICPs + BCPs) from the same driving global forecast or the same physics perturbations (PYPs) (Table 1), allows some indication of the relative impact of the different kinds of variations in the EPSs to be obtained.

The importance of the ICPs + BCPs compared to the PYPs in generating ensemble variance can be addressed using the IBP EPS and is shown in Figure 8(a). As before in section 3.3, the NVD is employed, but redefined as

$$NVD(P, \tau) = \frac{\sigma_{IB}^2(P, \tau) - \sigma_{PY}^2(P, \tau)}{\sigma_{IB}^2(P, \tau) + \sigma_{PY}^2(P, \tau)}, \quad (2)$$

with  $\sigma_{IB}^2 := 1/5 (\sigma_{IB1}^2 + \sigma_{IB2}^2 + \sigma_{IB3}^2 + \sigma_{IB4}^2 + \sigma_{IB5}^2)$  and  $\sigma_{PY}^2 := 1/4 (\sigma_{PY1}^2 + \sigma_{PY2}^2 + \sigma_{PY3}^2 + \sigma_{PY4}^2)$  denoting the averaged variances over the individual subgroups, respectively (cf. Table 1)\*. Thus, the variance  $\sigma_{IB}^2$  stems from the ICPs + BCPs, whereas the variance  $\sigma_{PY}^2$  stems from the PYPs. Here, a positive NVD indicates a relatively larger impact on the ensemble variance through the ICPs + BCPs, whereas a negative NVD indicates the relatively larger impact of the PYPs. A zero NVD shows that the ICPs + BCPs are equally important as the PYPs in generating ensemble variance.

Similarly, the importance of the BCPs alone versus the PYPs in generating ensemble variance can be studied using the BP EPS and results are presented in Figure 8(b). Note that in this case the NVD in Eq. (2) simply uses  $\sigma_B^2$  instead of  $\sigma_{IB}^2$ , because in the BP EPS there are no ICPs.

\*We have verified that comparing four versus five ensemble subgroups does not significantly influence the results.

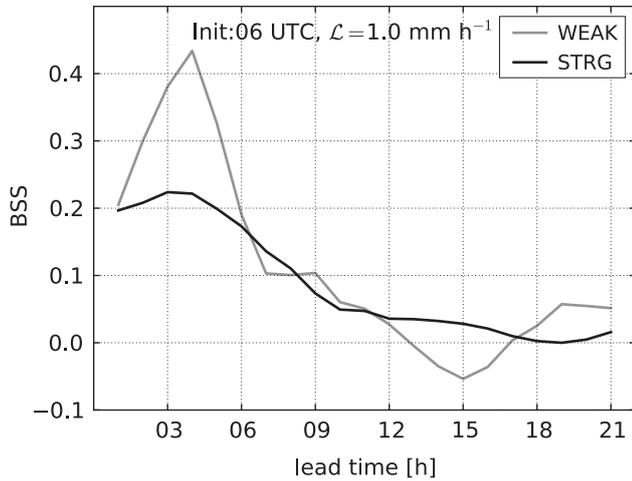


Figure 9.  $BSS_{IJ}$  for  $(I, J) = (IBP, BP)$  in Eq. (4), using the threshold  $\mathcal{L} = 1.0 \text{ mm h}^{-1}$  in Eq. (3).

The comparison of Figure 8(a) and (b) again indicates the considerable impact of the ICPs on the ensemble variance in the first forecast hours: in the BP EPS, the PYPs dominate over the BCPs at lead times  $< 3$  h, whereas in the IBP EPS the ICPs then dominate over the PYPs in this range. The BCPs become increasingly important at the later forecast times.

Significant differences in the impact of the various perturbations exist under the different meteorological regimes. Both EPSs in Figure 8(a) and (b) show a relatively larger impact of the PYPs on the ensemble variance in weak forcing conditions. The latter applies particularly in the convectively most active part of the day between about 5 and 12 h of the forecast (1300 to 2000 local time). This suggests a larger sensitivity to model error in weak versus strong forcing conditions, in agreement with other studies (e.g. Stensrud *et al.*, 2000; Groenemeijer and Craig, 2011; Keil *et al.*, 2013).

### 3.5. Probabilistic forecast quality

A main advantage of ensemble forecasts is the ability to provide probabilistic guidance on the occurrence of certain weather events. However, this requires the limited number of EPS members to sample as comprehensively as possible the uncertainty of the forecasted atmospheric state.

In this section, a weather-regime-dependent evaluation of the probabilistic forecast quality is presented. The aim is to evaluate the IBP against the BP EPSs, and the two different EPSs relative to the COSMO-DE forecast. The probabilistic forecast quality is measured using the Brier score

$$BS(\mathcal{L}, \tau) = \frac{1}{M} \sum_{i=1}^M [p(\mathcal{L}, \tau) - \hat{p}(\mathcal{L}, \tau)]^2. \quad (3)$$

In Eq. (3),  $p$  and  $\hat{p}$  represent, respectively, the forecasted and observed probabilities to exceed the precipitation rate threshold  $\mathcal{L}$ . In the case of the single deterministic forecast or the radar observations, the probabilities are either unity for exceeding the threshold or zero otherwise. The forecasts probabilities from the EPSs are computed as the fraction of the complete set of equally likely ensemble members. The summation in Eq. (3) is over the number of grid points  $M$  included at forecast lead time  $\tau$ .

With  $BS_I$  and  $BS_J$  denoting the Brier score in Eq. (3) from two different forecasting systems, the Brier skill score defined as

$$BSS_{IJ} = 1 - \frac{BS_I}{BS_J} \quad (4)$$

is used to evaluate the relative probabilistic forecast quality. The three pairs  $(I, J) = (IBP, BP)$ ,  $(IBP, DE)$ ,  $(BP, DE)$  are discussed

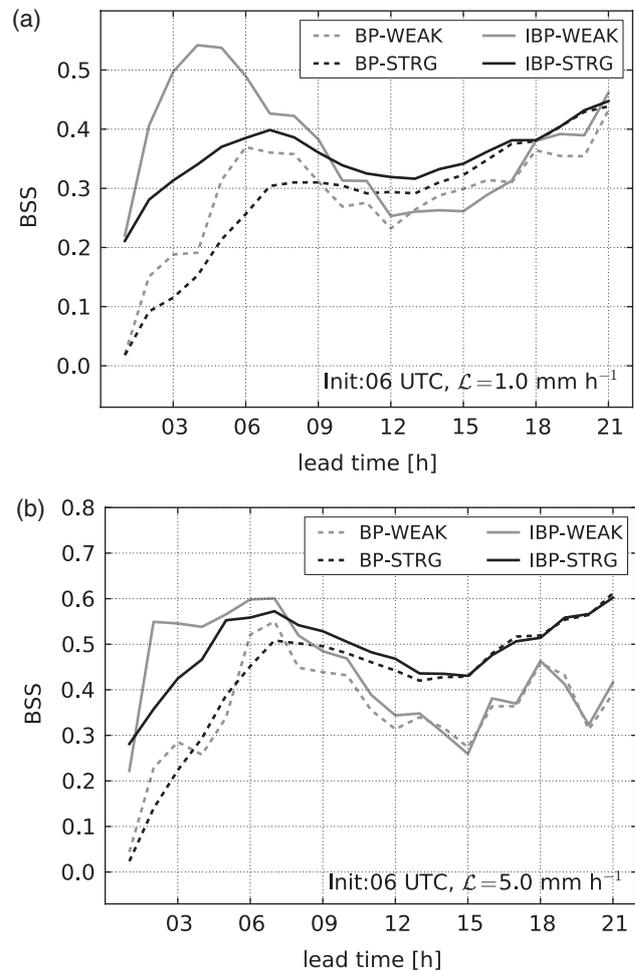


Figure 10.  $BSS_{IJ}$  for  $(I, J) = (IBP, DE)$  and  $(I, J) = (BP, DE)$  in Eq. (4), denoted with solid and dashed lines, respectively. It is shown for the thresholds (a)  $\mathcal{L} = 1.0 \text{ mm h}^{-1}$  and (b)  $\mathcal{L} = 5.0 \text{ mm h}^{-1}$  in Eq. (3). Note the different y-axis scale in (a) and (b).

below. A positive  $BSS_{IJ}$  indicates an improvement of the probabilistic forecast with system  $I$  over the reference system  $J$ , and vice versa.

The  $BSS_{IJ}$  for  $(I, J) = (IBP, BP)$  using the threshold  $\mathcal{L} = 1.0 \text{ mm h}^{-1}$  is depicted in Figure 9. The results show that the ICP impact on the BS is positive and largest in the first forecast hours, and then decays with forecast lead time. This behaviour is consistent with the impact of the ICPs on ensemble variance discussed in sections 3.3 and 3.4. Yet, for the  $BSS_{IJ}$  more larger differences are found during the first forecast hours under weak forcing when the positive impact on the  $BSS_{IJ}$  is up to twice as large. However it is noted that the initially larger impact in weak forcing conditions observed for the 0600 UTC forecast is not as pronounced for other initialisation times, e.g. for the forecast initialised at 1200 UTC, both the magnitude and duration of the impact is similar in the different regimes (not shown).

The  $BSS_{IJ}$  in terms of both the IBP and the BP EPSs versus the deterministic COSMO-DE as a reference, i.e.  $(I, J) = (IBP, DE)$  and  $(I, J) = (BP, DE)$  in Eq. (4), is shown in Figure 10 for two thresholds  $\mathcal{L} = 1.0 \text{ mm h}^{-1}$  and  $5.0 \text{ mm h}^{-1}$ . Initially, the BP EPS has a  $BSS_{IJ}$  close to zero. This is again because all BP ensemble members by design share the same initial conditions with COSMO-DE, namely the unperturbed COSMO-DE analysis. However, the  $BSS_{IJ}$  increases strongly with forecast lead time. In contrast to the BP EPS, the perturbed initial conditions in the IBP EPS results in a significant positive  $BSS_{IJ}$  immediately in the first hour of the forecast. In addition, the IBP EPS subsequently outperforms the BP EPS in terms of the  $BSS_{IJ}$  until both EPSs converge at lead times  $> 9$  h. These results apply under both strong and weak forcing conditions in qualitatively the same manner as shown before for other diagnostics. Moreover, the

described behaviour is qualitatively similar for the two different precipitation rate thresholds  $\mathcal{L}$ , but a larger  $BSS_{II}$  is generally attained for the higher threshold in Figure 10(b). In conclusion, the results in Figure 10 demonstrate the better performance of the ensemble precipitation forecasts compared to the deterministic forecast at the same convection-permitting resolution in strong as well as weak forcing conditions.

#### 4. Summary and discussion

The performance of downscaled initial condition perturbations from a driving multi-model global EPS in a LAM EPS for forecasting warm season precipitation at convection-permitting resolution has been studied. For this, COSMO-DE-EPS considering three major sources of uncertainty through ICPs, lateral BCPs and PYPs was compared against a second identical EPS but without the ICPs. The ensemble forecasts were investigated over Germany for a period of 3.5 months in the warm season of 2011. As convective-scale initial condition sensitivity is known to matter particularly in weather conditions when convective precipitation is controlled by small-scale processes, the ensemble forecasts and the impact of the downscaled ICPs were evaluated in a weather-regime-dependent way, separating forecasts performed under conditions of weakly forced from strongly forced convective precipitation. The respective classification was performed using the area-averaged convective adjustment time-scale as an objective indicator.

For all EPSs, forecast quality measures (e.g. spatially averaged precipitation, FBI, ETS) show a distinct behaviour in strong versus weak forcing conditions. However, the impact of the downscaled ICPs (e.g. as measured by the normalised variance difference between the EPS with and without ICPs), is found to be similar in strong versus weak forcing conditions. Generally, the impact of the ICPs is largest and dominates the PYPs and BCPs in the first  $\sim 6$  forecast hours. Afterwards, the impact of the ICPs decays relatively quickly with forecast lead time as the BCPs and PYPs in the EPS overwhelm the initial condition differences. The inclusion of the ICPs leads to lower ETS values for the individual members of the EPS with ICPs, as their perturbed initial conditions deviate from the unperturbed COSMO-DE analysis which initialises all members of the EPS without ICPs. However, the ICP impact on ensemble variance and probabilistic forecast quality is clearly positive and consistent. In addition, the probabilistic forecasts by the two EPSs outperform the deterministic forecast of COSMO-DE at the same convection-permitting resolution, and this is seen more strongly in the first  $\sim 9$  forecast hours for the EPS with the ICPs.

An important aspect of the present study is that the ICPs have been studied in the 'full' EPS combined with BCPs and PYPs. In this setting, the decay of the ICPs is related to the increasing impact of the BCPs and PYPs with lead time. Specifically, the BCPs representing the uncertainty in the large-scale weather provided by the driving EPS generally outperform the other sources of uncertainty at forecast lead times  $> 12$  h. The PYPs representing model error uncertainty associated with tunable parameters in selected physics schemes tend to have a significant impact throughout the entire 21 h forecast lead time, with the maximum impact during the convectively most active time of the day. Moreover, the relative impact of the PYPs versus the ICPs and BCPs is significantly larger in weakly forced conditions with typically low predictability.

The observed similar impact of the downscaled ICPs in the different forcing regimes is in contrast to previous studies investigating convective-scale ICPs, which showed a larger and longer-lasting impact under weak forcing (e.g. Vié *et al.*, 2011). We suspect that the differing results obtained in the present study could therefore be associated with the large-scale structure of the downscaled ICPs. This would be in agreement with the recent results by Keil *et al.* (2013), investigating the influence of BCPs and PYPs in COSMO-DE-EPS under different weather regimes. One the one hand, their study confirmed that the PYPs, which act

mainly on small scales, are more important in weak than strong forcing conditions. On the other hand, their results showed that the impact of the large-scale BCPs have a similar impact in weak and strong forcing regimes, thus sharing the behaviour observed for the downscaled ICPs herein.

It is concluded that the ICPs based on a downscaling approach from a driving multi-model global EPS (via an intermediate-resolution LAM EPS) are useful in terms of short-range convection-permitting ensemble forecasts of warm season convective precipitation. Advantages are that the method of generating ICPs is practical, consistent with the forcing from lateral boundary conditions, provides ensemble variance from the initial time of the forecast, significantly improves the probabilistic forecasts of precipitation, and complements well the BCPs and PYPs which become more important at later forecast hours. In the configuration of the EPS with BCPs and PYPs, the ICPs work equally well under conditions of strongly and weakly forced convective precipitation.

Nevertheless, the implemented ICPs are not representative of the initial condition probability distribution around the convective-scale LAM analysis of COSMO-DE. Future research in the framework of the Hans-Ertel Centre for Weather Research at Ludwig-Maximilians-Universität Munich will investigate how the currently used ICPs based on the downscaling approach compare to ICPs as provided by means of a local ensemble transform Kalman filter employed in the convective-scale data assimilation within the LAM domain (Reich *et al.*, 2011). In addition, research is being performed towards the use of a stochastic boundary-layer parameterisation in the convection-permitting EPS to represent aspects of model error, in contrast to the varied parameter approach that is currently used. Stochastic physics perturbations in the EPS are seen as a means to improve ensemble variance, but also to decrease systematic errors – a prominent example being the wrong diurnal cycle representation of convective precipitation which has been shown to exist in current convection-permitting models under weak forcing conditions in regions of flat terrain.

#### Acknowledgements

We appreciate help with the radar data by K. Stephan of DWD. The present study was carried out in the framework of the Hans-Ertel Centre for Weather Research. This research network of Universities, Research Institutes and DWD is funded by the German BMVBS (Federal Ministry of Transport, Building and Urban Development). Useful comments from two reviewers are gratefully acknowledged.

#### References

- Anthes RA, Kuo YH, Baumhefner DP, Errico RM, Bettge TW. 1985. Predictability of mesoscale atmospheric motions. *Adv. Geophys.* **28**: 159–202.
- Baldauf M, Seifert A, Förstner J, Majewski D, Raschendorfer M, Reinhardt T. 2011. Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities. *Mon. Weather Rev.* **139**: 3887–3905.
- Bouttier F, Vié B, Nuissier O, Raynaud L. 2012. Impact of stochastic physics in a convection-permitting ensemble. *Mon. Weather Rev.* **140**: 3706–3721.
- Bowler NE, Mylne KR. 2009. Ensemble transform Kalman filter perturbations 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.
- Buizza R, Houtekamer PL, Toth Z, Pellerin G, Wei M, Zhu Y. 2005. A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Mon. Weather Rev.* **133**: 1076–1097.
- Caron J. 2013. Mismatching perturbations at the lateral boundaries in limited-area ensemble forecasting: A case study. *Mon. Weather Rev.* **141**: 356–374.
- Clark AJ, Gallus WA, Xue M, Kong F. 2009. A comparison of precipitation forecast skill between small convection-allowing and large convection-parameterizing ensembles. *Weather and Forecasting* **24**: 1121–1140.
- Craig GC, Keil C, Leuenberger D. 2012. Constraints on the impact of radar rainfall data assimilation on forecasts of cumulus convection. *Q. J. R. Meteorol. Soc.* **138**: 340–352.

- Done JM, Craig GC, Gray SL, Clark PA, Gray MEB. 2006. Mesoscale simulations of organized convection: Importance of convective equilibrium. *Q. J. R. Meteorol. Soc.* **132**: 737–756.
- Gebhardt C, Theis SE, Paulat M, Ben Bouallègue Z. 2011. Uncertainties in COSMO-DE precipitation forecasts introduced by model perturbations and variation of lateral boundaries. *Atmos. Res.* **100**: 168–177.
- Groenemeijer P, Craig GC. 2011. Ensemble forecasting with a stochastic convective parametrization based on equilibrium statistics. *Atmos. Chem. Phys. Discuss.* **11**: 30457–30485.
- Helmert K, Hassler B, Seltmann JEE. 2012. An operational tool to quality control 2D radar reflectivity data for assimilation in COSMO-DE. *Int. J. Remote Sens.* **33**: 3456–3471, doi: 10.1080/01431161.2011.592161.
- Hohenegger C, Schär C. 2007a. Atmospheric predictability at synoptic versus cloud-resolving scales. *Bull. Am. Meteorol. Soc.* **88**: 1783–1793.
- Hohenegger C, Schär C. 2007b. Predictability and error growth dynamics in cloud-resolving models. *J. Atmos. Sci.* **64**: 4467–4478.
- Hohenegger C, Walser A, Langhans W, Schär C. 2008. Cloud-resolving ensemble simulations of the August 2005 Alpine flood. *Q. J. R. Meteorol. Soc.* **134**: 889–904.
- Keil C, Craig GC. 2011. Regime-dependent forecast uncertainty of convective precipitation. *Meteorol. Z.* **20**: 145–151.
- Keil C, Heinlein F, Craig GC. 2013. The convective adjustment time-scale as indicator of predictability of convective precipitation. *Q. J. R. Meteorol. Soc.*, doi: 10.1002/qj.2143.
- Kong F, Droegemeier KK, Hickmon NL. 2006. Multiresolution ensemble forecasts of an observed tornadic thunderstorm system. Part I: Comparison of coarse- and fine-grid experiments. *Mon. Weather Rev.* **134**: 807–833.
- Leoncini G, Plant RS, Gray SL, Clark PA. 2010. Perturbation growth at the convective scale for CSIP IOP18. *Q. J. R. Meteorol. Soc.* **136**: 653–670.
- Marsigli C, Boccanera F, Montani A, Paccagnella T. 2005. The COSMO-LEPS mesoscale ensemble system: Validation of the methodology and verification. *Nonlinear Processes Geophys.* **12**: 527–536.
- Migliorini S, Dixon M, Bannister R, Ballard S. 2011. Ensemble prediction for nowcasting with a convection-permitting model—I: Description of the system and the impact of radar-derived surface precipitation rates. *Tellus Ser. A* **63**: 468–496.
- Molini L, Parodi A, Rebora N, Craig GC. 2011a. Classifying severe rainfall events over Italy by hydrometeorological and dynamical criteria. *Q. J. R. Meteorol. Soc.* **137**: 148–154.
- Montani A, Cesari D, Marsigli C, Paccagnella T. 2011b. Seven years of activity in the field of mesoscale ensemble forecasting by the COSMO-LEPS system: Main achievements and open challenges. *Tellus Ser. A* **63**: 605–624.
- Peralta C, Ben Bouallègue Z, Theis SE, Gebhardt C, Buchhold M. 2012. Accounting for initial condition uncertainties in COSMO-DE-EPS. *J. Geophys. Res.* **117**: 1–13, doi: 10.1029/2011JD016581.
- Reich H, Rhodin A, Schraff C. 2011. LETKF for the nonhydrostatic regional model COSMO-DE. *COSMO Newsl.* **11**: 27–31.
- Saito K, Hara M, Kunii M, Seko H, Yamaguchi M. 2011. Comparison of initial perturbation methods for the mesoscale ensemble prediction system of the Meteorological Research Institute for the WWRP Beijing 2008 Olympics Research and Development Project (B08RDP). *Tellus Ser. A* **63**: 445–467.
- Saito K, Seko H, Kunii M, Miyoshi T. 2012. Effect of lateral boundary perturbations on the breeding method and the local ensemble transform Kalman filter for mesoscale ensemble prediction. *Tellus A* **64**: 11594, doi: 10.3402/tellusa.v64i0.11594.
- Schättler U, Doms G, Schraff C. 2009. *A Description of the Nonhydrostatic COSMO-Model. Part VII: User's Guide*. Deutscher Wetterdienst: Offenbach, Germany. <http://www.cosmo-model.org/content/model/documentation/core/cosmoUserGuide.pdf> (accessed 13 March 2012).
- Stensrud DJ, Fritsch JM. 1994a. Mesoscale convective systems in weakly forced large-scale environments. Part II: Generation of a mesoscale initial condition. *Mon. Weather Rev.* **122**: 2068–2083.
- Stensrud DJ, Fritsch JM. 1994b. Mesoscale convective systems in weakly forced large-scale environments. Part III: Numerical simulations and implications for operational forecasting. *Mon. Weather Rev.* **122**: 2084–2104.
- 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.
- Stephan K, Klink S, Schraff C. 2008. Assimilation of radar-derived rain rates into the convective-scale model COSMO-DE at DWD. *Q. J. R. Meteorol. Soc.* **134**: 1315–1326.
- Vié B, Nuissier O, Ducrocq V. 2011. Cloud-resolving ensemble simulations of Mediterranean heavy precipitating events: Uncertainty on initial conditions and lateral boundary conditions. *Mon. Weather Rev.* **139**: 403–423.
- Walser A, Lüthi D, Schär C. 2004. Predictability of precipitation in a cloud-resolving model. *Mon. Weather Rev.* **132**: 560–577.
- Wang Y, Bellus M, Wittmann C, Steinheimer M, Weidle F, Kann A, Ivatek-Şahdan 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.
- Warner TT, Peterson RA, Treadon RE. 1997. A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction. *Bull. Am. Meteorol. Soc.* **78**: 2599–2617.
- Wilks DS. 2006. *Statistical Methods in the Atmospheric Sciences: An Introduction*. Elsevier: San Diego, CA.
- Zhang F, Snyder C, Rotunno R. 2003. Effects of moist convection on mesoscale predictability. *J. Atmos. Sci.* **60**: 1173–1185.
- Zimmer M, Craig GC, Keil C, Wernli H. 2011. Classification of precipitation events with a convective response timescale and their forecasting characteristics. *Geophys. Res. Lett.* **38**: 5802, doi: 10.1029/2010GL046199.