Physically Based Stochastic Perturbations (PSP) in the Boundary Layer to Represent Uncertainty in Convective Initiation

KRISTIN KOBER AND GEORGE C. CRAIG
Ludwig-Maximilians-Universität München, Munich, Germany

(Manuscript received 27 May 2015, in final form 21 March 2016)

ABSTRACT

Stochastic perturbations allow for the representation of small-scale variability due to unresolved physical processes. However, the properties of this variability depend on model resolution and weather regime. A physically based method is presented for introducing stochastic perturbations into kilometer-scale atmospheric models that explicitly account for these dependencies. The amplitude of the perturbations is based on information obtained from the model’s subgrid turbulence parameterization, while the spatial and temporal correlations are based on physical length and time scales of the turbulent motions. The stochastic perturbations lead to triggering of additional convective cells and improved precipitation amounts in simulations of two days with weak synoptic forcing of convection but different amounts of precipitation. The perturbations had little impact in a third case study, where precipitation was mainly associated with a cold front. In contrast, an unphysical version of the scheme with constant perturbation amplitude performed poorly since there was no perturbation amplitude that would give improved amounts of precipitation during the day without generating spurious convection at other times.

1. Introduction

Forecasting convective initiation more than a few hours in advance is an ongoing challenge in atmospheric research. The exact timing and location will probably not be forecast by numerical weather prediction (NWP) models in the near future, but forecasts of the probability of precipitation can show useful skill.

Probabilistic forecasts aim to represent uncertainty that results from several sources of varying importance. The intrinsic uncertainty of a chaotic system like the atmosphere contributes to the overall uncertainty, as do inaccuracies in the numerical integration of the equations of motion, including initial and boundary condition uncertainty and model error. Model error describes not only the numerical error resulting from the computational need to truncate at a specific resolution in the integration of the equations but also the inability to exactly represent the effects of physical processes inherent on scales below the grid size (subgrid) (see, e.g., Lorenz 2006). Multiple integrations of NWP models (ensembles) can be used to provide probabilistic information but can be set up in different ways, depending on the represented sources of uncertainty. Recent studies have shown that in different weather regimes, different sources of uncertainty dominate: in cases of strong large-scale forcing, initial and boundary conditions uncertainty contributes more to the overall uncertainty, whereas in weak large-scale forcing, model error is more important (Groenemeijer and Craig 2012; Keil et al. 2014; Kober et al. 2015).

Model error is addressed with two primary approaches: the so-called model diversity approach combines different forecast models or different parameterizations (for a specific process), whereas the random sampling approach (or stochastic physics) aims to represent subgrid-scale variability that has the potential to grow to larger scales (Eckel and Mass 2005). Within stochastic physics, several methods of varying complexity exist to include perturbations in parameterizations. They range from perturbing input fields before they enter a parameterization (Lin and Neelin 2003; Bright and Mullen 2002), through perturbing tunable parameters within a parameterization (Bright and
Mullen 2002), perturbing the parameterized tendencies (Buizza et al. 1999; Teixeira and Reynolds 2008), adding additional terms to equations to consider upscale transport (Shutts 2005, 2015), and a theory-based entirely stochastic formulation of a parameterization (Plant and Craig 2008). The application of stochastic parameterizations has shown improved skill (Lin and Neelin 2003) and increased spread in ensemble prediction systems (Buizza et al. 1999; Berner et al. 2012; Christensen et al. 2015).

Most of the existing approaches to stochastic physics aim to maximize ensemble spread and do not distinguish between the varying contributions of specific physical processes (Buizza et al. 1999). But depending on the meteorological situation, different physical processes contribute differently to the overall variability. Teixeira and Reynolds (2008) have shown the potential of considering a specific process in detail by applying perturbations to the tendencies that are scaled with a variance calculated in the convection parameterization. This gives a physically reasonable spatial variability of the perturbations based on model-intrinsic information.

A second factor determining which physical processes require a stochastic representation is the resolution of the model relative to the physical scales of variability. For example, in a convective boundary layer, the variability is dominated by eddies with sizes corresponding to the depth of the boundary layer, of order 1 km. Global atmospheric models with grid lengths of tens of kilometers are unlikely to require a stochastic representation of this variability, since each grid box is influenced by the accumulated effects of many eddies. A deterministic parameterization that represents this average effect should be accurate. In a kilometer-scale weather prediction model, on the other hand, the boundary layer eddies should lead to significant variability on the grid scale, which may require a stochastic representation, since the eddies themselves are not resolved. This problem is exacerbated by numerical diffusion, which can strongly damp what grid-scale variability the model has. In certain weather situations, such as the diurnal cycle of convection, this missing variability could lead to significant forecast errors: for example, the systematic tendency for convection to be weak and late in convection-permitting NWP models (Trentmann et al. 2009; Kühnlein et al. 2014).

In this paper, we will consider how stochastic perturbations can be introduced in a kilometer-scale model in a physically consistent way. While the initial focus will be on the convective boundary layer, we will introduce a framework that will be extended in future work to include other processes relevant to the initiation of convection. In general, the initiation of convection requires certain atmospheric conditions. In synoptic situations without large-scale forcing, convection can develop if instability, measurable by the convective available potential energy (CAPE), as well as local triggers to overcome a possible inversion above the boundary layer, measurable by the convective inhibition (CIN), are available together with moisture in specific heights (Done et al. 2012). The distribution of moisture and temperature in the atmosphere determines the amount of CAPE and is influenced by several processes, especially in the atmospheric boundary layer. Soil type, land use, orography, and elevation are important because they define properties of the interaction between the land surface and the atmosphere. Several processes contribute to create variability in a convective boundary layer. The heating of surfaces with different heat conduction properties by the sun causes turbulent eddies to evolve. Orography modifies the boundary layer by lifting air parcels, differential heating along slopes causes thermals to rise, and the lateral displacement of flows can lead to convergence. Deep convective cells modify the stability in the boundary layer by downdrafts associated with precipitation and its evaporative cooling effect (so-called cold pools). Circulations with mesoscale organization (e.g., mountain–valley winds, land–sea breezes) influence the boundary layer structure.

The scheme introduced here will use information from the model’s subgrid turbulence parameterization to determine the amplitude of variability that should appear on the smallest resolved scales. The sensitivity to key parameters, including the choice of variables to be perturbed and the length and times scales of the perturbations, will be examined. The impact of the perturbation scheme on precipitation forecasts for three case studies will then be examined. Primary focus will be on a case with weak synoptic forcing and a strong diurnal cycle of precipitation, since as noted above, unperturbed models can have systematic errors in such situations. Three scientific questions will be addressed:

1) How can information from physical processes in the numerical model be used to create stochastic variability that adapts consistently to different weather situations?
2) Is the variability predicted by these physical processes large enough to have a significant impact but not so large as to degrade the simulations?
3) Do the physically based stochastic perturbations respond appropriately to changing weather situations, and could similar results be obtained by a simpler scheme that does not use physical information from the model?

The analysis here will mainly consider the bias in area-averaged precipitation. Future work will assess the contribution of the stochastic perturbations to ensemble spread in comparison to other sources of forecast uncertainty.
The paper is structured as follows: In section 2, the physically based stochastic perturbations will be introduced as well as the mesoscale weather prediction model used in this study and the verifying observational data. In section 3, the method will be applied in several case studies representing two different weather regimes to assess the structure of the perturbations, their impact on the precipitation fields, and the sensitivities in parameter settings of the perturbations scheme. Additionally, forecast quality of other variables will be assessed with additional measures. Finally, the results will be discussed (section 4), and summarizing conclusions will be drawn (section 5).

2. Data and methods

In this study, we introduce a method to consider the variability of a specific physical process in the atmospheric boundary layer relevant for convective initiation: namely, boundary layer turbulence driven by surface heating. This section explains the approach and describes the NWP model in which the perturbations are implemented. Note that the perturbation method is applicable in any atmospheric model that allows for calculation of the relevant physical process information. The observational data used to evaluate the forecasts and the selected case studies in which the parameterization is tested will be introduced briefly as well as the analysis strategy for the suggested method.

a. Physically based stochastic perturbations in the boundary layer

We propose a concept of process-based model error representation in terms of a stochastic parameterization that considers the missing variability due to a specific physical process relevant in the initiation of convection. This is realized by adding a term to the tendencies $\frac{\partial \Phi}{\partial t}$ of a resolved variable $\Phi$ that is a product of a random number field $b_{\Phi}$ representing the respective physical processes. Different seeds for the random number field $\eta$ give different realizations of the small-scale variability to provide an ensemble of simulations that samples the possible outcomes from this source of uncertainty.

In the initial implementation, we consider the variability due to turbulence driven by surface heating by adding stochastic perturbations to the resolved fields of temperature $T$, vertical velocity $w$, and moisture $q$. The physical information $b_{\Phi}$ is obtained from the turbulence parameterization in the form of variances of the variables so that the perturbations are formulated as

$$\left( \frac{\partial \Phi}{\partial t} \right)_{stoch} = \frac{\partial \Phi}{\partial t} + \alpha_{sh} \eta_{sh} \langle \Phi^2 \rangle^{1/2},$$

with $\Phi$ being the resolved variables $T, w,$ and $q$, and $\langle \Phi^2 \rangle$ being the variances predicted by the turbulence scheme. The Gaussian random number field $\eta_{sh}$ and the scaling parameter $\alpha_{sh}$ must have certain properties and are specified as follows.

The quantity $\eta_{sh}$ is a random number field with Gaussian distribution (Figs. 1a,c) and additionally contains spatial and temporal correlations. The temporal correlation is incorporated by creating a new random number field every 10 min, representing the average eddy turnover time in a well-mixed convective boundary layer. Spatially, $\eta_{sh}$ is constructed to be vertically coherent since eddies in the convective boundary layer typically extend through the depth of the layer, meaning that every level is perturbed with the same random number field. Horizontally, the random number field is convolved with a two-dimensional Gaussian function with half-width ($\sigma$) of 2.5 grid points, resulting in a correlation representing the effective model resolution of $5dx$ (Bierdel et al. 2012; Verrelle et al. 2015). This significantly reduces the numerical damping that would remove most of the perturbation amplitude in a few time steps if the perturbations were applied at the actual model resolution. After normalization, the mean of the random number field is 0, and the standard deviation is 1 (Figs. 1b,d). Ensemble members are created through different seeds of $\eta_{sh}$. Figure 1 shows the variability with two different realizations (one in Figs. 1a and 1b and another in Figs. 1c and 1d).

The parameter $\alpha_{sh}$ scales the amplitude of the perturbations and is chosen such that it represents the contribution to convective variability of all eddies within a region corresponding to the effective model resolution in one integration time step:

$$\alpha_{sh} = \frac{\ell_{e}}{\sqrt{5dx}} \frac{1}{dt},$$

where $\ell_{e}$ is the asymptotic mixing length describing the average size of an eddy assumed in the turbulence parameterization. In the Consortium for Small-Scale Modeling (COSMO) model version 5.01 used in this study, this value is 150 m. The horizontal grid length of the model $dx$ is $0.025^\circ$ (approximately 2800 m) in this study, $5dx$ is the effective horizontal resolution of the COSMO model (Bierdel et al. 2012), and $dt$ is the
temporal resolution of the model, here 25 s. Finally, we allow an overall (dimensionless) factor $a_{sh,F}$, expected to be of order unity, that will be fixed independent of weather situation, model resolution, and variable to be perturbed. In theory, this parameter should be exactly one, but in practice the perturbation amplitudes will be affected by numerous implementation details such as numerical diffusion, and a slightly different value may be optimal. Note, however, that if best results were obtained with a value substantially different than one (say 10 or 100), this would be an indication that the physical processes considered here are not the most relevant ones. In the present work, $a_{sh,F} = 2$ for all $\Phi$. Sensitivities to the settings in $h_{sh}$ and $a_{sh}$ are tested in the following section.

The resulting perturbations vary in time and space because the physical information ($b_{l\Phi}$) is updated by the turbulence parameterization every time step. This update allows the perturbations to be scaled to the respective weather situation. In situations when fluxes due to surface heating are small (e.g., at night in weak synoptic forcing), the turbulence scheme predicts small variances, and the perturbations added to the tendencies are negligibly small. The perturbation fields vary in space and time through their physical relation to the buoyant heat flux and through the properties of the random number field.

b. Weather forecast model COSMO

This study is conducted within the COSMO model (Baldauf et al. 2011) version 5.01. COSMO solves the fully compressible equations of motion on an Arakawa-C grid that is chosen to have a horizontal resolution of $0.025^\circ$ (approximately 2800 m) and 50 vertical levels by a terrain following coordinate system (Lorenz grid staggering). Model forecasts are computed over 24 h with a time step of 25 s. The domain over which the model is integrated is smaller than the operational version run at the Deutscher Wetterdienst (DWD) and is centered over Germany (Fig. 1). Initial and boundary conditions

---

**Fig. 1.** Random number field $h_{sh}$: (a),(c) without correlation and (b),(d) after convolution with Gaussian functions and scaling for two different realizations: one in (a) and (b) and another in (c) and (d).
are interpolated from COSMO simulations with 0.625° (approximately 7000 m) resolution over the same 24 h driven by global forecasts of the European Centre for Medium-Range Forecasts (ECMWF) Integrated Forecast System (IFS) and updated hourly. The resolution of the COSMO simulations allows deep convection to be computed explicitly, while shallow convection is parameterized based on a simplified version of the Tiedtke (1989) parameterization. The cloud microphysics parameterization for cloud formation and decay is of Lin type (Lin et al. 1983), with a one-moment scheme for the autoconversion, accretion, and self-collection. The lower boundary condition is calculated with the multilayer version of the soil model TERRA (Schraff et al. 2016). Turbulence is parameterized by a 1.5-order closure (Buzzi et al. 2011), corresponding to level 2.5 of the Mellor and Yamada notation (Mellor and Yamada 1982) with horizontal wind components, accumulated from the observations with 5-min frequency and projected onto the COSMO model grid (Stephan et al. 2008). Not all of the model domain is covered by radar data (in part because not all radars are available at all times). Therefore, the evaluation domain is restricted to the areas with available radar data, and hence the number of evaluated model grid points can vary with time. Areas where data are not available are indicated by gray shading in the respective figures.

Forecasts of 2-m temperature and 10-m wind speed are compared to hourly observations at specific locations (SYNOP data). For each available SYNOP observation, the closest model grid point is chosen for the comparison. The number of observations varies with time and is in the range of 700 over the model domain.

d. Case studies

The performance of the stochastic perturbation scheme will be examined in simulations of three case studies with different meteorological conditions. Initial testing and parameter sensitivities will be explored in 24-h forecasts starting from 0000 UTC 1 July 2009. This case was highlighted in several studies by Kober et al. (2014), Keil et al. (2014), and Kober et al. (2015) as an example of convection that followed a classical diurnal cycle evolution under weak synoptic forcing. Examples of the observed precipitation field and time evolution of area-integrated precipitation will be shown in the next section for comparison with the simulation results. Keil et al. (2014) showed that simulations of such situations are most sensitive to perturbations in model parameterizations, and hence this is a case where the stochastic method is likely to have a significant effect.

To verify that the scheme can adapt to different meteorological situations, we then consider two additional case studies. First, another case with weak synoptic forcing is examined (27 June 2010), but this time where the precipitation was weak. In this case, there is the possibility that a poorly configured scheme could lead to over-prediction of precipitation. Last, a strongly forced case (24 June 2012) is examined, where the precipitation is convective, but associated with a summertime cold front. In such situations, the results of Kober et al. (2015) lead us to expect little positive impact from the stochastic perturbations, but we hope no negative effects.

e. Analysis strategy

This study introduces a method to represent uncertainty in convective initiation aiming for improved precipitation forecasts. The analysis of the case studies focuses on the impact of the process-specific stochastic perturbations scheme on properties of precipitation fields like the evolution of the overall precipitation amount with domain-averaged values and the distribution over...
several precipitation thresholds in comparison to radar observations. The physically based perturbations are compared to unphysical perturbations, where the physical information is eliminated by setting \( \text{Then}^2 \) in Eq. (2) to 1, resulting in the loss of the horizontal, vertical, and weather-situation-adaptive scaling of the perturbations. As an indicator for the influence of the perturbations on the energy and instability, time series of domain-averaged values of CAPE and CIN are compared for both perturbation types. To ensure that the perturbations do not negatively influence as well perturbed variables and others, the forecast quality of temperature at 2 m and the wind speed in 10 m is evaluated against surface station observations with the evolution of the root-mean-square error (RMSE; Wilks 2011).

To test the statistical significance of our findings in comparison to the unperturbed reference simulation, we applied bootstrapping on the ensembles resulting from physically based and unphysical stochastic perturbations (Wilks 2011). Within this procedure, we sample (with \( N = 1000 \)) from the 10 ensemble members at each grid point, assuming spatially uncorrelated fields. On the resulting fields, we apply the domain-averaging or RMSE statistics. We mark our results as different with statistical significance in comparison to the reference simulation with an asterisk if the statistic of the reference is outside of the 95% confidence interval of the distributions derived from the physically based and unphysical samples. If the assumption of no spatial correlation is not satisfied, the degree of significance will be overestimated, so, as an alternative measure, the 10 individual ensemble members are also plotted together with the mean, which provides an indication of significance under the assumption of perfect spatial correlation.

3. Results

The stochastic perturbations are introduced in simulations of a specific case study where strong precipitation occurred under weak large-scale synoptic forcing. First, properties of the perturbation fields are presented in terms of their components and their variability. The impact of the perturbations on the precipitation forecasts is investigated in terms of time series of the domain-averaged precipitation and stability measures, the distribution over several precipitation thresholds and sensitivities to different parameter settings are explored. After this, the stochastic perturbations are applied in the two other weather situations to determine if the basic concept is applicable and the parameter settings are robust so that no additional tuning is necessary. The impact of the physically based perturbations is compared to that of perturbations with constant amplitude to assess the importance of adapting the perturbations to different weather conditions. Finally, for all three case studies, forecast quality is evaluated for the reference (unperturbed), the physically based (phys) and unphysical (unphys) perturbations forecasts of 2-m temperature and 10-m wind speed.

a. Structure of the physically based perturbations

The 1 July 2009 case is chosen as the reference case study with which we will demonstrate the effect of the perturbations and sensitivities to the parameter settings. This day is characterized by intense and short-lived small-scale convective cells in weak large-scale synoptic forcing, as the radar observation at 1200 UTC shows (Fig. 2a). The forecast with the operational version of the COSMO-DE model (which in the following will be referred to as the
reference or unperturbed simulation; DE stands for Deutschland and specifies the operational version run at DWD) shows that most of the observed precipitation features are missed, although the large-scale structure, with no precipitation in the northwestern part of the domain and small local cells, is captured (Fig. 2b). This weather situation is ideal to test if additional process-related variability can improve the convection forecast in terms of the timing, the intensity, and the amount of precipitation.

In this subsection, we will examine how the spatial distribution of diagnosed turbulent variances influences the perturbation amplitudes. First, the horizontal structure of the perturbations will be discussed, then the vertical structure, and finally the contribution of the stochastic scheme to the total tendencies of temperature, humidity, and vertical velocity will be shown.

Figure 3 shows an example of the horizontal structure of the variance components and the final perturbations at a specific model level. The random number field \( \eta_{bh} \) (Fig. 1d) is the same for all variables and at all levels and is updated every 10 min. Figures 3a, 3c, and 3e are snapshots of the variances \( \Phi_i^2 \) as calculated in the turbulence parameterization, and Figs. 3b, 3d, and 3f show the resulting perturbation fields that will be added to the respective tendency fields, all at 1200 UTC and at level 45 (approximately 300 m). All model fields show boundary effects due to the update with boundary data at each full hour. The temperature variance (Fig. 3a) shows larger variability over mountainous areas, especially the Alps and the Black Forest. In areas where the model already simulated precipitation (see Fig. 2b), the temperature variance has higher values (e.g., over central or northeastern Germany). This is also reflected in the resulting perturbation field (Fig. 3b), where these areas show a high small-scale variability. Areas with low variances (e.g., the western part of central Germany and over the North Sea) result in low amplitudes of the perturbations. The variance of vertical velocity \( (w^2) \) (Fig. 3c) is more homogeneous, and areas associated with precipitation show very small values, whereas some areas in the southern Alps and over the Netherlands have high variances. This projects on the resulting perturbations field (Fig. 3d). Despite this, the structure of the random number field (Fig. 1d) can be seen in the perturbations clearly. The variance of moisture \( (q^2) \) (Fig. 3e) in general has very small amplitudes with very localized maxima (e.g., over Switzerland). This is also seen in the resulting perturbation field (Fig. 3f), where the amplitude is small in comparison to the other two variables and the structure of the random number field \( \eta_{bh} \) is less visible. In general, the structure of the physical information in the variances is retained in the perturbation fields and can be identified clearly in comparison with the original random number field.

While Figs. 3b, 3d, and 3f showed the horizontal variability of the variances \( \Phi_i^2 \) at a specific level within the boundary layer, Fig. 4 displays their vertical variability with box-and-whisker plots, analyzed from 1145 to 1215 UTC in 5-min intervals. The median and the variability (in terms of the 25th and 75th percentiles as boxes and the 5th and 95th percentiles as whiskers) of the temperature variance (Fig. 4a) is largest in the lowest level and decreases rapidly through the depth of the boundary layer. Above level 40 (approximately 900 m), values are very small, although in the highest 15 levels (above approximately 10 km) the medians increase somewhat. The reasons for this are unclear and may be associated with the model numerics. There is no evidence of any impact on these nonzero values in the uppermost troposphere on the simulations. The median and the variability of the variance of the vertical velocity (Fig. 4b) also clearly show the signature of the boundary layer with positive values between level 40 (approximately 900 m) and the ground. The moisture variance (Fig. 4c) is generally much smaller, but the structure is similar: the convective boundary layer can be identified by nonzero values in the lowest model levels. The evolution of boundary layer depth over the day can be seen in the vertical structure of the variances as well (not shown). Thus, while in principle the turbulence parameterization acts everywhere in the model domain, in practice the effects are confined to the boundary layer. This vertical structure of the variances and their diurnal cycle is projected on the perturbation fields, resulting in decreasing influence of the perturbations with height and during nighttime.

The impact of the perturbations on the model tendency fields \( \partial \Phi_i / \partial t \) is illustrated in Fig. 5 with snapshots at 1200 UTC on level 45 (approximately 300 m). The tendencies are displayed for all three perturbed variables \( (T, w, q) \) of the unperturbed reference COSMO-DE simulation (Figs. 5a,d,g), the perturbed COSMO-DE simulation (Figs. 5b,e,h), and, for clarity, the difference between the two (Figs. 5c,f,i). It should be noted that, although time and level agree, the tendencies of the perturbed COSMO-DE simulation cannot be interpreted as the sum of Figs. 3b, 3d, and 3f and Figs. 5a, 5d, and 5g since perturbations are added continuously over the entire model simulation, and hence Figs. 5b, 5e, and 5h show the integrated effect of the perturbations.

The unperturbed temperature tendencies (Fig. 5a) are a smooth, mostly positive, field. The isolated negative values are associated with precipitation (cf. with Fig. 2b). Some higher positive values occur in the Alps and close to the coast, whereas areas covered with water have values close to zero. The tendencies of the perturbed COSMO-DE simulation (Fig. 5b) generally have the same properties, but as a result of more precipitation cells, there are
more isolated negative values, and more variability in the small positive values (e.g., northern Germany or Netherlands) is found. This is also seen in the difference between the two fields (Fig. 5c). The difference field shows the nonlinear effect of the perturbations: extrema are not only shifted in position, but their number and amplitude changes. The perturbation structure as seen in Fig. 3c can hardly be identified in the full tendency field.

The tendencies of the vertical velocity of the unperturbed reference simulation (Fig. 5d) are small in general. In areas
around precipitating cells, negative values are found, and the signature of mountains can be seen with high intensity and high variability. The tendency field of the perturbed simulation is dominated by the perturbations field (Fig. 5e), but the effects of precipitation cells and orography can still be identified. The comparison between the reference and the perturbed simulation (Fig. 5f) shows large differences in particular because of the structure of the perturbations.

The moisture tendencies of the reference simulation are very weak (Fig. 5g). In the northwestern part of the domain and in the mountains some negative values can be found. Some very small and isolated positive tendencies occur in mountainous regions like the Alps or the Bavarian Forest. In the perturbed simulation (Fig. 5h), the structure of the perturbations shows up as low-amplitude values as well as variability in the northwestern domain and the Alps, consistent with the unperturbed simulation. This is seen in the difference fields as well (Fig. 5i). Except for the temperature variances, the difference between the unperturbed and the perturbed tendencies is obvious, and the structure of the perturbations field can be identified clearly.

b. Impact on precipitation and sensitivity to parameters

The impact of the perturbation tendencies on the spatial structure of the precipitation forecast is displayed in Fig. 6. The perturbed forecast is still characterized by small-scale and intense precipitation cells over the model domain, but no cells over northwestern Germany. The comparison with Fig. 2 shows that more convective cells and hence more precipitation is simulated than in the reference simulation, for example, over eastern Germany, although it is still less than observed, for example, over southern Germany.

The temporal development of the domain-averaged precipitation in COSMO-DE simulations and radar observations will be used to provide a quantitative measure of the impact of the perturbations and to assess the sensitivities to parameter settings in the perturbation scheme \[ \Phi, \alpha_{sh}, \eta_{sh}, \text{and} \langle \Phi^2 \rangle \text{in Eq. (2)} \]. Note that model grid points are only considered in the calculation of the average if radar data are available at the respective time (indicated by the white areas within the black boxes in Figs. 2 and 6). To provide some context for judging the importance of changes in the various parameters, Fig. 7 and Fig. 8 also show time series for the unperturbed reference COSMO-DE simulation, one realization of the perturbed simulations using final parameter settings, and the radar observation. The comparison with the radar observation (light gray) shows that the reference simulation (black) substantially underestimates the growth of precipitation, reaching a value of 0.05 mm h\(^{-1}\) about 2 h later than observed and reaching a maximum intensity...
that is just over a third of the observed peak. The precipitation in the perturbed simulation (dark gray) increases much more rapidly, reaching 0.05 mm h\(^{-1}\) an hour later than observed, with the maximum rain rate underestimated by only about 20%. However, all the simulations show precipitation decreasing about 2 h earlier than in the observations, leading to strong underestimates in the second half of the day.

Figure 7a shows the sensitivity of the precipitation forecast to the choice of the perturbed variables $\Phi$. The simulations where the variables $T$, $w$, and $q$ are perturbed separately (red) show that their effects on precipitation add up in the full-perturbations scheme (where all three are perturbed simultaneously). Perturbations of the vertical velocity $w$ have the smallest effect on precipitation, whereas the moisture $q$ perturbations have the largest. The diurnal cycle of precipitation is not modified by this parameter setting, as it is found in all simulations. The sensitivity to the scaling of the amplitude of the perturbations is displayed in Fig. 7b, where $\alpha_{sh,\Phi}$ is varied to 1 and 3, in addition to the final value of 2. With increasing $\alpha_{sh,\Phi}$, the overall amount of precipitation increases, and
the onset of convection is earlier, with small nonzero values of precipitation appearing at 0900 UTC. The impact of the correlation length and time scales of the random number field $\eta$ is investigated in Fig. 7c. Doubling the length of the horizontal correlations (cf. Fig. 1a) does not affect the onset time of convection, but the maximum is higher and precipitation lasts longer. Halving the persistence time of the random number fields to 5 min results in a lower maximum precipitation but still improves the forecast in comparison to the unperturbed reference. The variability due to different realizations of the random number fields is shown in Fig. 7d with 10 ensemble members and an ensemble mean. The variability among the realizations is smaller than the variability due to parameter settings in Figs. 7a, 7b, and 7c. Ensemble member 1 is the perturbed member displayed in the previous panels. The properties of the mean of the 10 ensemble members agree with the findings for this member, and similar results were found for ensemble sizes ranging from 5 to 20 (not shown).

Figure 8 shows results of simulations with a simplified version of the perturbation scheme, where the physical information was eliminated by setting $\Phi^2$ to 1, resulting in the loss of the horizontal, vertical, and weather-situation-adaptive scaling of the perturbations. In Fig. 8a, the unphysical perturbations of each variable are first scaled separately and then again for all three variables together, choosing a scaling factor $\alpha_{sh,\Phi}$ that...
approximately reproduces the observed maximum precipitation rate. Perturbations of temperature $T$ and vertical velocity $w$ can be tuned such that the general structure of daytime-driven convection is maintained, but those of moisture $q$ and of all three variables simultaneously result in the loss of the diurnal cycle. Here, the diurnal cycle structure can be recreated if the moisture perturbations are scaled to be small. In contrast to the physical perturbations, which are all scaled with the same $\alpha_{sh,\theta} = 2$, the tuning of the constant perturbation amplitude is more difficult and the shutdown of precipitation in the night hours is only possible when the overall amount is underestimated. The best values for the perturbations’ amplitudes are indicated by the orange lines (solid for the ensemble mean and dashed for member 1), and these values will be used in the following sections for the other case studies. The comparison of the physically based and the unphysical ensembles and their means (Fig. 8b) confirms the findings for a single member described above. With these parameter settings, both ensembles are characterized by a similar variability. At all times that the forecasts have precipitation, the difference in rate from the reference simulation is statistically significant.

The sensitivity tests show an advantage of the physical perturbations in comparison to the unphysical perturbations since there is no single perturbation amplitude that is appropriate throughout the diurnal cycle. Using physically based variances means that once an appropriate value of the tuning parameter $\alpha_{sh}$ has been identified, it can be used throughout the diurnal cycle.

Comparing the precipitation snapshots of the perturbed and unperturbed forecasts in Figs. 2b and 6 gives the impression that the increased precipitation in the perturbed model forecasts comes in the form of additional convective cells, rather than by changing their size or intensity. This impression is confirmed by examining the frequency distribution of different rain rates in the simulations (Fig. 9). The relative frequency of high and low rain rates is similar for all the simulations but stronger in every category for the physically perturbed simulations and even stronger with the unphysical perturbations. It is worth noting, however, that none of the simulations reproduce the observed distribution, which has more frequent occurrences of relatively low rain rates (up to $2.0 \text{ mm h}^{-1}$) relative to the higher rates. As a result, the simulations with physical perturbations have approximately the correct frequency of the highest rain rates while substantially underpredicting the low values, while the simulations with unphysical perturbations are better for lower rain rates but overestimate the occurrence of high values.

The examination of CAPE and CIN allows some insight into how the perturbations lead to changes in the precipitation rates. Time series of domain-averaged CAPE and CIN, based on a mean boundary layer parcel, are shown in Fig. 10 for a representative ensemble member with physical perturbations and one with unphysical perturbations. If a lack of variability in the boundary layer makes triggering of convective cells more difficult, then CAPE would be expected to accumulate. Consistent with this interpretation, CAPE is smaller in the perturbed simulations throughout the period where precipitation occurs. However, this does not explain the origin of the spurious convection early in the simulation with unphysical perturbations, since CAPE is no different from the other simulations. Instead, the large unphysical perturbations at this time lead to a reduction in CIN. Apparently this is sufficient to allow convective updrafts to
form, although the occurrence is rare enough that the domain-averaged CAPE is not noticeably affected.

c. Application to other case studies

The perturbation scheme will now be applied to two other weather situations to test its adaptability to changes in synoptic forcing and the robustness of the parameter settings. We compared 10 realizations of the physical perturbations and their mean to the reference forecast and to 10 realizations of the unphysical perturbations and their mean in two case studies. These are chosen to be different to the case study in the previous section: another weak forcing case but now with a small number of convective cells and small total precipitation and a strong forcing case with a coherent frontal system.

The 27 June 2010 case is a weak synoptic forcing weather situation with small amounts of precipitation as a result of a strong inversion layer. The radar observes rain only between 1300 and 1700 UTC with a maximum at 1500 UTC (Fig. 11), resulting from isolated short-lived cells. The evaluation of the evolution of the domain-averaged precipitation reveals that the reference COSMO-DE run produces very little precipitation. The physically perturbed ensemble has precipitation starting and peaking about an hour earlier than in observations, with an underestimate of the peak amplitude, but is able to reproduce the decay of precipitation around sundown. The mean of the forecasts with unphysical perturbations also improves the initiation and amount of precipitation in the early afternoon in comparison to the reference forecast but misses the termination and repeatedly triggers cells in the nighttime when no precipitation was observed. As a result, the diurnal cycle is obscured when unphysical perturbations are included with an amplitude chosen to give good results for the first case study. As for the previous case study, the frequency distribution of rain rates shows that both physical and unphysical perturbations lead to distributions that are too flat (i.e., not enough occurrence of weak rain rates in comparison to strong) (Fig. 12). The overestimation of total precipitation with the unphysical perturbations is

![Figure 9](image9.png)  
**FIG. 9.** Comparison of COSMO-DE simulations with radar observation. Distribution over precipitation thresholds on 1 Jul 2009 with radar observation (light gray), reference COSMO-DE simulation (black), mean over 10 members of the physical perturbations scheme (dark gray), and unphysical perturbations (orange). Values for the 0.0 mm h\(^{-1}\) threshold are divided by 100, and actual values are written above each bar for this threshold.

![Figure 10](image10.png)  
**FIG. 10.** Comparison of COSMO-DE simulations. Temporal evolution of domain-averaged (top) CAPE and (bottom) CIN on 1 Jul 2009 with reference COSMO-DE simulation (black), one realization of the physical perturbations scheme (dark gray), and one realization of the unphysical perturbations (orange).
associated with an excess of large values, while the underestimate with the physical perturbations comes from a lack of low-threshold occurrences. While the unphysical perturbation amplitude chosen for the 1 July case is clearly too large for this case study, the physical perturbation amplitude has adapted to this weather situation without retuning.

The third case study, 24 June 2012, is characterized by strong synoptic forcing and a cold front passing the domain over the course of the day. In terms of the domain-averaged precipitation, all model simulations underestimate the overall amount of precipitation. Both types of perturbations improve on the reference COSMO-DE simulation by increasing the maximum precipitation rate, with a larger impact from the unphysical perturbations. However, most of the low bias in comparison to radar observations remains (Fig. 13). The radar shows a faster decrease after the maximum at 1600 UTC. The rain-rate frequency distributions show that the stochastic perturbations have less influence in this case (Fig. 14). As in the previous cases, the increased precipitation in the forecasts with unphysical perturbations is distributed over all categories. Snapshots of the hourly precipitation at 1600 UTC (Fig. 15) show that the underestimate in all of the simulations is associated with the frontal band being too narrow. This error is unaffected by the stochastic perturbations. The physical perturbations have very little impact on the front at all, whereas the unphysical perturbations (with amplitude chosen for the 1 July 2009 weak forcing case) produce more small-scale variability in the precipitation field than was observed, although the large-scale distribution is essentially unchanged. Only a single member of each ensemble is shown in Fig. 15, but the same characteristics were found in all 10 members of the perturbed ensembles.

d. Comparison with surface temperature and wind observations

In Fig. 16, the RMSE and ensemble spread are shown for 2-m temperature and the 10-m wind speed forecasts. The impact of the physical perturbations is almost always neutral, indicating the additional variability is within the range of the model error. The notable exception is the increased temperature error found in the afternoon for the 1 July 2009 case when the boundary layer is most disturbed (Fig. 16a). The very large perturbations produced by the unphysical scheme lead to increased temperature errors in all three cases, except during the nighttime hours, when temperature variability is strongly damped by stable stratification in the boundary layer (Figs. 16a,c,e). The reverse of this pattern is found for the wind errors associated with the unphysical perturbations, where larger errors occur during the night in all three cases (Figs. 16b,d,f). This is to be expected since stronger vertical mixing will damp variability in the boundary layer during the day in comparison to night.

4. Discussion

As described in the preceding sections, we have developed and tested a physically based method for
introducing stochastic perturbations associated with parameterized turbulence in a numerical model with kilometer-scale resolution. The perturbations are added to the tendencies of three model variables (temperature, moisture, and vertical velocity). They are computed as the product of a random number field, the flux information from the turbulence parameterization, and a constant scaling factor. For each of these components, the sensitivity of the results to important parameter choices was tested: the combination of perturbed variables (temperature, moisture, and vertical velocity), correlation scales in the random number field (temporal and spatial), and the amplitude of the scaling factor. These results were compared to those of the unperturbed model as well as simulations based on different realizations of an unphysical perturbation scheme with similar spatial and temporal structures but a constant mean amplitude that did not take information from the turbulence parameterization.

Since the introduction of perturbations was mainly confined to the boundary layer and therefore directly impacted the stability of those air parcels, it is not surprising that the additional variability led to increased convective precipitation where it had any effect at all. Frequency distributions of rain rates showed that precipitation of all intensities was increased when perturbations were added, suggesting that the impact of the perturbations was indeed to trigger more convective updrafts, rather than to change their properties significantly. The largest contribution comes from perturbations to the humidity field, consistent with previous studies that have shown spatial variability in CAPE to be dominated by moisture variations (Muller et al. 2009), but the contributions of $T$ and $w$ are additive.
A single parameter multiplying the amplitude of all stochastic perturbations must be tuned to give optimal results. A value of 2 was found to be appropriate, indicating that the variances derived from the turbulence parameterization are of a sufficient order of magnitude to affect the model behavior. More importantly, the same parameter value was found to give good results throughout the diurnal cycle in a case of strong precipitation but weak synoptic forcing and in two other case studies with different meteorological conditions. This is in contrast to results obtained from the simplified version of the method that does not use physical information from the turbulence scheme. In this case, the amplitude had to be retuned for the different cases, and even then no tuning could be found that would reproduce the range of variation of precipitation through the diurnal cycle. The tuning procedure in this study is very pragmatic, but in future work we will explore the potential of more sophisticated parameter estimation methods, usually applied in data assimilation (e.g., Hacker and Snyder 2005; Ruiz and Pulido 2015).

While the physically based stochastic perturbations introduced here provide a systematic improvement in the diurnal cycle of convective precipitation, it depends upon the fluxes computed by the existing turbulence scheme in the model and will not improve model biases that are not associated with inadequate variability on small scales. Similarly, other deficiencies in the model, for example in the surface flux or cloud microphysics parameterizations, may contribute to the model biases seen here and will not be improved by stochastic perturbations. Furthermore, the initial implementation of the perturbations presented here considers only variability represented by classical turbulence theory. Other processes relevant to convective initiation, such as orography, would have to be considered separately. Finally, it is conceivable that small-scale
variability in other subgrid processes, such as cloud microphysics, land surface interactions, and soil moisture treatment, and their representation in numerical weather prediction models might be a significant source of uncertainty in aspects of kilometer-scale weather prediction other than convective initiation.

While it was found here that the physically based stochastic perturbations respond to different weather regimes more accurately than simple additive perturbations, some of this improvement may be obtainable with simple multiplicative schemes like stochastically perturbed physics tendencies (SPPT; Buizza et al. 1999) that randomly re-scale the parameterization tendencies. For example, one would expect such a scheme to capture the termination of convection that the constant perturbations missed since the multiplicative perturbations would go to zero as the mean fluxes decay. A systematic comparison between the physically based scheme and a more pragmatic stochastic perturbation method would be of great interest in future.

The potential of the suggested method was demonstrated in case studies representing different weather regimes. The robustness of the results will have to be
proven on the basis of a larger dataset. Additionally, the impact of the perturbations on the forecast skill of other model variables like temperature in comparison to surface observations has shown promising results but needs further investigation in future work as does the scale dependence aspect of the predictability and the probabilistic forecast skill.

5. Summary and conclusions

In this study, a method is proposed for introducing stochastic perturbations related to turbulent fluctuations into kilometer-scale numerical weather prediction models. This is expected to be especially relevant for initiation of convective clouds, since this depends on processes in the subcloud layer that occur on scales near the model resolution. The results of this study can be summarized in terms of the three questions posed in the introduction:

1) How can information from physical processes in the numerical model be used to create stochastic variability that adapts consistently to different weather situations?

   The scheme implemented here uses TKE and flux information from the model’s turbulence parameterization to compute the corresponding variances in temperature, vertical velocity, and moisture. Stochastic increments are then added to the model fields to include the resolved portion of this turbulent variability, as determined from the ratio of the model grid length to the mixing length scale used in the turbulence scheme. This provides an automatic adjustment for changes in model resolution, while the scaling of the perturbation amplitude to the turbulent variances makes the scheme self-adapting to changing weather situations.

2) Is the variability predicted by these physical processes large enough to have a significant impact but not so large as to degrade the simulations?

   For the 1 July 2009 weak forcing case, the representation of the diurnal cycle was improved. Precipitation in the physically perturbed model reached significant intensity about an hour earlier than the unperturbed run, although still about an hour later than observed, and the maximum value was substantially improved. The increase appeared to be associated with an increased number of convective cells. Since the turbulence in the boundary layer was less intense at night, the amplitude of the perturbations was smaller, and no convection was produced. The impact on the RMSE for surface wind and temperature was neutral, except during periods of strong surface heating, when temperature errors were increased. This would indicate that the perturbation amplitude was, for the most part, within the observed variability of the boundary layer. If the physical information from the turbulence parameterization was omitted, it was not possible to find a perturbation amplitude that would reproduce the afternoon maximum precipitation rate without triggering spurious convection during the night and introducing large errors in surface wind and temperature.

3) Do the physically based stochastic perturbations respond appropriately to changing weather situations, and could similar results be obtained by a simpler scheme that does not use physical information from the model?

Two additional case studies were examined using the parameter settings that were found to give best results for the first case. For the weak synoptic forcing case, the physical perturbation scheme again substantially improved the peak precipitation rate without triggering spurious convection. In contrast, the unphysical perturbations were clearly larger than optimal for this case and produced too much precipitation throughout the day. A third case was also considered, where the precipitation was associated with a cold front. In this case, the physical perturbations had very little impact on the simulations, whereas the unphysical perturbations were large enough to disrupt the structure of the frontal precipitation in an unrealistic way.

The promising initial results from the three case studies presented here suggest that a long-term trial of the method that samples the full range of weather regimes would be a desirable next step. Work is also in progress to implement perturbations associated with unresolved orographic features and to represent secondary convective initiation by cold pools.

Acknowledgments. We thank the German weather service (DWD), especially Klaus Stephan for kindly providing the radar observations and Ulrich Blahak and Matthias Raschendorfer for helping with COSMO model details. We gratefully acknowledge the German Research Foundation (DFG) for funding the research unit “Predictability and Dynamics of Weather Systems in the Atlantic-European Sector (PANDOWAE)” and the transregional collaborative research center SFB/TRR165 “Waves to Weather (W2W).” The comments of three anonymous reviewers improved this manuscript significantly.

REFERENCES
Baldauf, M., A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, 2011: Operational convective-scale numerical weather prediction with the COSMO model: Description


