AGU PUBLICATIONS

Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1002/2016MS000634

A stochastic scale-aware parameterization of shallow cumulus convection across the convective gray zone

Mirjana Sakradzija[1], Axel Seifert[2], and Anurag Dipankar[1]

[1]Max Planck Institute for Meteorology, Hamburg, Germany, [2]Hans-Ertel Centre for Weather Research, Deutscher Wetterdienst, Hamburg, Germany

Key Points:
- A scale-aware parameterization is proposed which includes stochastic sampling and convective memory
- Stochastic parameterizations can improve the physical aspects of modeled convection and clouds
- Interactions of parameterizations with model dynamics play a significant role in the gray zone

Correspondence to:
M. Sakradzija,
mirjana.sakradzija@mpimet.mpg.de

Citation:
Sakradzija, M., A. Seifert, and A. Dipankar (2016), A stochastic scale-aware parameterization of shallow cumulus convection across the convective gray zone, *J. Adv. Model. Earth Syst.*, *8*, doi:10.1002/2016MS000634.

Received 22 JAN 2016
Accepted 27 APR 2016
Accepted article online 2 MAY 2016

Abstract The parameterization of shallow cumuli across a range of model grid resolutions of kilometre-scales faces at least three major difficulties: (1) closure assumptions of conventional parameterization schemes are no longer valid, (2) stochastic fluctuations become substantial and increase with grid resolution, and (3) convective circulations that emerge on the model grids are under-resolved and grid-scale dependent. Here we develop a stochastic parameterization of shallow cumulus clouds to address the first two points, and we study how this stochastic parameterization interacts with the under-resolved convective circulations in a convective case over the ocean. We couple a stochastic model based on a canonical ensemble of shallow cumuli to the Eddy-Diffusivity Mass-Flux parameterization in the icosahedral nonhydrostatic (ICON) model. The moist-convective area fraction is perturbed by subsampling the distribution of subgrid convective states. These stochastic perturbations represent scale-dependent fluctuations around the quasi-equilibrium state of a shallow cumulus ensemble. The stochastic parameterization reproduces the average and higher order statistics of the shallow cumulus case adequately and converges to the reference statistics with increasing model resolution. The interaction of parameterizations with model dynamics, which is usually not considered when parameterizations are developed, causes a significant influence on convection in the gray zone. The stochastic parameterization interacts strongly with the model dynamics, which changes the regime and energetics of the convective flows compared to the deterministic simulations. As a result of this interaction, the emergence of convective circulations in combination with the stochastic parameterization can even be beneficial on the high-resolution model grids.

1. Introduction

Behavior and characteristics of numerically simulated convection and clouds are highly dependent on the horizontal resolution of atmospheric models. This dependence is pronounced when convection and clouds are not effectively resolved on the model grid, and is attributed to parameterization schemes and to nonlinearity of model dynamics [e.g., *Kiehl and Williamson*, 1991; *Giorgi and Marinucci*, 1996; *Williamson*, 1999; *Pope and Stratton*, 2002; *Bryan et al*., 2003; *Jung and Arakawa*, 2004]. This implies that information about the scales of model discretization should be inherently included in the parameterization of clouds and convection. The purpose of this paper is to recognize and uncover the main requirements for a scale-aware parameterization of shallow convective clouds. Rather than developing an universal parameterization for shallow convection, we develop a stochastic scale-aware parameterization based on a single convective case and we focus on the interactions of the stochastic parameterization with organized convective circulations that emerge due to nonlinearity of the model equations.

The main cause of nonconverging behavior of parameterization schemes with increasing horizontal resolution are the assumptions made in order to close the system of model equations. The main closure assumption in a conventional parameterization of convective clouds is the validity of quasi-equilibrium (QE), first introduced as a convective QE by *Arakawa and Schubert* [1974]. By the QE assumption, the subgrid cloud ensemble is uniquely related to and fully determined by the grid-scale variables, with the assumption that the convective environment is slowly changing so that convection has enough time to adjust to changes in the forcing. Assuming this, the average properities of the convective ensemble are used to estimate the response of the subgrid convective processes to an imposed large-scale forcing. The physical aspect of QE is a mechanism that determines the strength of the convective activity in the model grid column as a

response to destabilization of the atmosphere by some large-scale process. In the case of shallow convection, this is usually some form of a boundary layer equilibration mechanism [e.g., *Neggers et al*., 2006]. Another requirement for QE is a robust statistical sample of convective clouds within a model grid column. This statistical requirement is fulfilled for shallow convection if the horizontal resolution of the model is coarser than approximately 20–50 km.

In the range between the coarse grid scales where QE is valid and very high resolution scales of O(10–100 m), where the dynamics of shallow convection and clouds is considered as effectively resolved, convection is not adequately parameterized by any of the conventional deterministic approaches. These intermediate scales are commonly known as the convective and shallow-cloud gray zone. In the gray zone for a cloud ensemble, the model grid cell can hold only a subsample of the convective cloud ensemble, which invalidates the statistical aspect of the QE assumption [see *Plant and Craig*, 2008; *Dorrestijn et al*., 2013; *Sakradzija et al*., 2015]. Because of the limited cloud sample size, the fluctuations of subgrid convective states around the QE state increase as the grid resolution increases. Thus, the ensemble average response of subgrid convection is no longer representative of the actual subgrid state, nor does it correspond to the most probable outcome. As a direct consequence of the failure to meet the statistical requirement for QE, the cloud ensemble on the gray-zone model grids is also not subject to the physical mechanism that is defined to control and govern the convective response, thus the physical aspect of QE also breaks down.

Apart from the QE breakdown, another cause of divergence in modeled clouds and convection on the kilometre-scale resolutions is the nonlinear behavior of model dynamics. In an undisturbed, slowly changing convective environment on the kilometre-scale grids in both large-eddy simulations (LES) and in numerical weather prediction (NWP) models, the solution of the model dynamical equations leads to emergence of organized convective motions in form of convective cells or rolls [*Piotrowski et al*., 2009; *LeMone et al*., 2010; *Zhou et al*., 2014; *Ching et al*., 2014]. This type of organized motion is similar in mechanism to the Rayleigh-Bernard convection, and it emerges in the boundary layer flows heated from below when the buoyancy forces overcome the diffusive heat transport in the vertical direction. In the zone between effectively resolved convection and effectively subgrid convection, organized convective circulations are highly grid-dependent with spatial overturning scales larger than the natural convective scales [*Piotrowski et al*., 2009; *LeMone et al*., 2010; *Zhou et al*., 2014; *Ching et al*., 2014].

The under-resolved convective circulations, as a manifestation of the scale-dependent behavior of model dynamics, are undesirable for at least two reasons. First, in the absence of a convective parameterization, the overenergetic and under-resolved flow on the kilometre-scale grids can lead to biases in forecasts such as over-prediction of rainfall amount and delayed convective activity, which is already documented for the case of deep convection [e.g., *Bryan and Rotunno*, 2005; *Lean et al*., 2008; *Roberts and Lean*, 2008]. Second, the artificial truncation of convective processes and cloud ensembles at the model grid-resolution scale might cause an artificial interaction between model dynamics and a subgrid parameterization. The parameterization will react on the under-resolved convective flow, and influence it back through the convective tendencies. So it is expected that this interaction with the under-resolved flows might bring additional uncertainty and scale-dependence in parameterization of shallow clouds.

To address the two major model properties that degrade the model convergence with increasing resolution, i.e., the validity of statistical QE assumption and the effects of nonlinearity of model equations, we extend the eddy-diffusivity mass-flux (EDMF) scheme [*Neggers*, 2009] in the icosahedral nonhydrostatic (ICON) model [*Zängl et al*., 2015] by introducing the stochastic shallow cumulus ensemble model [*Sakradzija et al*., 2015] into the EDMF framework. The stochastic shallow cumulus ensemble provides a way to parameterize the variability of subgrid convective states around the QE convective state [*Plant and Craig*, 2008]. The distribution of convective states parameterized by the stochastic ensemble scheme automatically adapts to the model resolution, being narrow and Gaussian-like for the coarse grids, and changing the shape toward a heavy-tailed distribution on the kilometre-scale. In the new stochastic EDMF scheme, the requirements for the statistical QE are fulfilled by applying the deterministic EDMF closure across a large area around each model grid cell, while the subgrid convective states are subsampled from the cumulus ensemble defined in a compound stochastic process. In this way, the macroscopic state of the cloud ensemble is constrained based on the physical principles of the model closure, and the uncertain microscopic states of the cloud subensembles within model grid cells are governed by a probabilistic law.

We choose to investigate here a shallow convective case over the trade-wind region, the Rain In Cumulus over the Ocean (RICO) case, where the cloud field can be randomly distributed over wide areas but also convection and clouds can develop into organized systems, where convective roll circulations oriented along the wind direction are very common [*LeMone and Pennell*, 1976; *Rauber et al*., 2007]. The organization of convection and clouds into rolls, cloud streets and mesoscale arcs is also documented in LES studies of RICO [*Seifert and Heus*, 2013; *Seifert et al*., 2015]. So clouds are prone to organize in the RICO case, and roll organization is a physical phenomenon observed in both nature and LES. Here we use the LES of the RICO case as a reference case, against which the developed stochastic parameterization and its interactions with the model dynamics are tested.

In the following we first make an overview of stochastic convection and scale-aware methods and emphasize the new understanding and contribution of our work. A description of the ICON model and the convective case study is provided in section 2. Section 3 contains the description of the stochastic cloud ensemble parameterization that is coupled to the EDMF scheme in ICON. The unified EDMF parameterization of turbulence, convection and clouds and the main EDMF closure assumptions relevant for our study are also described in section 3. The results of a set of numerical simulations at different resolutions using both the deterministic and the stochastic model are given in section 4. Section 5 provides summary and conclusions.

1.1. A Short Overview of Stochastic and Scale-Aware Methods

A stochastic parameterization within the EDMF framework has already proven to improve the model skill of a global weather prediction model and to reproduce various convective cases in the study of *Sušelj et al*. [2014]. Their stochastic EDMF parameterization uses a Monte Carlo sampling method to estimate the properties of a predefined number of convective plumes in the model grid cell and a stochastic representation of the lateral entrainment rate [*Sušelj et al*., 2013]. To achieve scale-awareness in the stochastic sampling of a cloud ensemble, it is necessary to represent fluctuations in the number of cloud elements and their properties. Hence, in this paper we take a different route to stochastic parameterization and use an approach similar to the statistical ensemble framework of *Plant and Craig* [2008]. By this framework, cloud number and cloud properties fluctuate, but these fluctuations are constrained by a statistical-physical principle.

A physically based constraint imposed on the statistical cloud ensemble and convective fluctuations is important mainly because stochastic fluctuations alter the energetics of model flow dynamics (see section 4.4). An example of a scale-aware sampling that is conditioned on the resolved-scale state is a statistical approach of *Dorrestijn et al*. [2013]. By this approach, pairs of turbulent fluxes of heat and moisture are precomputed, and a stochastic parameterization of convective transport is developed based on the LES statistics. Turbulent fluxes in a grid column are sampled using a conditional Markov chain, where the subgrid state is conditioned on the resolved-scale state of the same model column. One of the main differences between our approach and this local and purely statistical method, is that in our approach a constraint is based on statistical physics and is defined for a system's macrostate. Macrostate of a cloud ensemble is the ensemble average state that is considered as a predictable state defined for a larger region around the grid-column and it satisfies the requirements for a QE closure.

The scale-awareness of shallow convection is frequently studied by partitioning the LES coarse-grained turbulent fluxes between subgrid and resolved scales and by blending the transition from fully resolved to completely subgrid processes across different resolutions [e.g., *Honnert et al*., 2011; *Dorrestijn et al*., 2013]. An implication of the scale-dependence of flow dynamics for the modeling of convection is that blending or subgrid-resolved partitioning in a parameterization employed in a nonhydrostatic model will not behave as expected from LES. The subgrid spectrum of convective motions, which is estimated from LES, is disparate from the resolved spectrum of model dynamics, which furthermore changes with resolution. This is why we refer to convection as under-resolved instead of partly resolved. A requirement for scale-awareness is thus a stochastic parameterization that has a capacity to alter the model flow dynamics to approach the coarse-grained LES statistics, and to join together the subgrid and resolved convective spectra. In this paper, we show how the two spectra can be matched more closely by the interactions of stochastic parameterization with flow dynamics, however, the effective resolution of the numerical model will always be a remaining limitation.

A scale-aware parameterization is predominantly identified as a formulation of subgrid processes that includes one or more scale-aware parameters. One of the most essential parameters in such formulations is

the area fraction covered by clouds a_c, which is assumed to be negligibly small $a_c \ll 1$ in a classic mass flux scheme [e.g., *Arakawa and Schubert*, 1974]. In order to make the parameterization scale-aware, the condition $a_c \ll 1$ has to be relaxed and a scale-aware formulation for a_c has to be developed for the high resolution grids [e.g., *Arakawa et al*., 2011; *Grell and Freitas*, 2014]. Our approach retains the assumption of a small convective area only on the large-scales at which QE is assumed valid. Then, instead of a single scale-aware parameter, we model a scale-aware distribution of the convective mass flux and convective area fraction.

The motivation for a stochastic parameterization of convection and clouds covers a much broader range of topics than what is addressed in this paper. The need for stochastic perturbations comes from the fact that the small-scale high-frequency variability of convection has a potential to influence the larger spatial and longer time scales, because of the multiscale and continuous atmospheric flow spectrum. So the stochastic convective perturbations can grow and propagate to the synoptic scales and represent the upscale growth and propagation of the model error, which is already documented for the deep convective case [*Teixeira and Reynolds*, 2008; *Selz and Craig*, 2015]. Thereby, the stochastic parameterization of convection represents the uncertainty of model physics in the ensemble prediction systems by inflating the ensemble spread [e.g., *Palmer*, 2001; *Teixeira and Reynolds*, 2008].

Most studies on stochastic convection were done for deep convection, but similar results can be expected for shallow convection. The unresolved deep convective variability has a substantial impact on the ability of numerical models to represent the low-frequency variability modes of the atmosphere in the tropics on intraseasonal and longer timescales [*Lin and Neelin*, 2000, 2002]. There are as well indications that convectively generated spatial variability in atmospheric thermodynamic fields can influence the timing of the diurnal cycle of deep convection over land [*Stirling and Petch*, 2004]. Thus a stochastic parameterization of convection might be beneficial also in improving the short-timescale atmospheric variability in models. Another example is the advantage of the stochastic multicloud convective parameterization [*Khouider et al*., 2010] over a deterministic scheme in improving the variability of tropical convection and the intermittent structure of synoptic and mesoscale convective systems and convectively coupled waves [*Frenkel et al*., 2012]. Related to this topic, the subgrid stochastic cellular automata (CA) scheme that parameterizes the lateral organization of convection and convective memory can mimic the organization by gravity waves originating from deep convection [*Bengtsson et al*., 2011]. This overview might not even be complete, because the field is new and developing rapidly.

Indications that a stochastic parameterization of deep convection can increase the forecasting skill of a model [*Bright and Mullen*, 2002; *Berner et al*., 2015] and reduce the systematic model biases [*Berner et al*., 2012] provide us some guidance for shallow cloud parameterization and are promising. We show in this paper that the stochastic parameterization of shallow convection, beside the possibility to induce a statistical impact in the ensemble modeling systems and improve the measure of uncertainty, also provides a tool to improve the distribution of the convective properties and to alter the regime and energetics of convective flow dynamics.

2. Description of the Numerical Model and the Case Study

A scale-aware approach to parameterization of shallow convective clouds is partly motivated by the new-generation global high-resolution NWP and climate models that offer the possibility of local grid refinement or two-way nesting [e.g., *Tomita and Satoh*, 2004; *Skamarock et al*., 2012; *Zängl et al*., 2015]. The ICON model, built on the icosahedral horizontal mesh of points with local grid refinement [*Zängl et al*., 2015], is one example of such a model that would require parameterizations that can adapt to the model resolution without the possibility of additional tuning of the scheme parameters. In this study, we use the fully compressible nonhydrostatic ICON model equations, which are solved for the following prognostic variables: the horizontal velocity component normal to the triangle edges v_n, the vertical wind component w, density ρ, virtual potential temperature θ_v, and the specific masses and number densities of tracers. For the purpose of this study, the tracers include water vapor (q_v), liquid water (q_l), rain(q_r), snow (q_s), ice (q_i), and total water variance ($\sigma^2_{q_t}$). The model domain used in our study is doubly-periodic with cyclic boundary conditions on a pseudo 2D torus triangular grid similar to the one used in *Dipankar et al*. [2015]. In this configuration, the horizontal resolution is uniform and the vertical coordinate is height-based. For more information about the numerical methods we refer the reader to *Dipankar et al*. [2015].

ICON comes with the three physics parameterization packages: the package suitable for climate modeling adopted from the ECHAM model [*Stevens et al*., 2013], the numerical weather prediction physics package (NWP) developed from the COSMO model (Consortium for Small-scale Modeling) [*Doms et al*., 2011] and the large eddy simulation (LES) physics [*Dipankar et al*., 2015]. The physics package applied in our study is a modified and extended numerical weather prediction package (NWP). Turbulence, convection and shallow clouds are parameterized using the unified EDMF parameterization scheme. Shallow clouds are parameterized in EDMF by means of a bimodal probability density function, in which the two normal modes represent the diffusive and convective contributions to cloud cover [*Lewellen and Yoh*, 1993; *Neggers*, 2009]. The parameterizations of radiation, grid-scale condensation and precipitation, as well as deep convection are excluded from the model configuration.

We focus on the undisturbed period of the RICO shallow convective case over the ocean. A detailed description of the setup of the RICO case is provided in *van Zanten et al*. [2011]. The effects of radiation, large-scale advection and subsidence on the convective case are applied through the prescribed constant large-scale forcing tendencies. The surface conditions for the RICO case are the constant sea surface temperature, the constant surface pressure, and a surface layer bulk aerodynamic parameterization for the surface turbulent fluxes [see *van Zanten et al*., 2011]. The UCLA-LES (University of California, Los Angeles - Large Eddy Simulation) RICO-140 case, which is the case with suppressed precipitation by doubling the cloud droplet number density to $N_c = 140$ cm^{-3}, is used as a reference simulation (similar to *Sakradzija et al*. [2015]). Precipitation is supressed because we choose to focus on the nonprecipitating behavior of the convective case. For the same reason, the parameterization of precipitation formulated within the moist updraft budget equations is not included in the EDMF configuration. Domain size of the reference LES simulation is 51.2 km, and the horizontal resolution is 25 m.

3. Unified Parameterization of the Moist Convective Boundary Layer in ICON

In the EDMF scheme, the turbulent flux of a conserved quantity ϕ, such as liquid water potential temperature θ_l and total water mixing ratio q_t, $\phi=\{\theta_l, q_t\}$, is decomposed into the diffusive and convective transport terms [*Siebesma et al*., 2007; *Neggers et al*., 2009]:

$$\overline{w'\phi'} = -K\frac{\partial\bar{\phi}}{\partial z} + \sum_{i=1}^{2} M_{ui}(\phi_{ui}-\bar{\phi}) \tag{1}$$

where K denotes the eddy-diffusivity, $\partial\bar{\phi}/\partial z$ is a local vertical gradient of a conserved quantity, M_{ui} is the convective updraft kinematic mass flux, and $\phi_{ui}-\bar{\phi}$ is the excess of the updraft property ϕ_{ui} with respect to the environment. The first term on the right hand side represents local turbulent mixing and is parameterized by the eddy-diffusivity approach of *Troen and Mahrt* [1986] and *Holtslag and Boville* [1993]. The second term represents nonlocal transport by strong organized updrafts via the dual-updraft mass flux approach [*Neggers et al*., 2009]. The mass flux term is split into the bulk dry ($i = 1$) and bulk moist ($i = 2$) updraft terms to represent all subcloud layer updrafts and positively buoyant shallow clouds that form on top of the moist updrafts. The kinematic mass flux is defined as the product of the updraft area fraction a_{ui} and the updraft vertical velocity w_{ui}:

$$M_{ui} \equiv a_{ui}w_{ui}. \tag{2}$$

The focus of our study is on the moist ($i = 2$) bulk updraft mass flux term of the EDMF parameterization.

Vertical profiles of thermodynamic and momentum tendencies are estimated in EDMF by applying an updraft plume model to both dry and moist updrafts. The vertical transport below and within the cloud layer is modeled using the same equations. The updraft budget equations for a conserved variable ϕ_{ui} and vertical velocity w_{ui} used in EDMF are formulated as in *Siebesma et al*. [2007]; *Neggers et al*. [2009], and references therein:

$$\frac{\partial\phi_{ui}}{\partial z} = -\epsilon_{ui}(\phi_{ui}-\bar{\phi}), \tag{3}$$

$$\frac{1}{2}(1-2\mu)\frac{\partial w_{ui}^2}{\partial z} = -b\epsilon_{ui}w_{ui}^2 + B_{ui}, \tag{4}$$

where

$$B_{ui} = \frac{g}{\bar{\theta}_v}(\theta_{v,ui} - \bar{\theta}_v). \quad (5)$$

ϵ_{ui} are the fractional entrainment rates of the dry and moist updrafts. The fractional entrainment rate for vertical velocity is assumed proportional to the fractional entrainment rate for ϕ with the factor of proportionality $b = 0.5$, i.e., $\epsilon_{wi} = b\epsilon_{ui}$, and $\mu = 0.15$ [*Siebesma et al*., 2007].

Equations (3) and (4) are solved starting at the lowest model level, and the integration terminates at the height where $w_{ui}^2 = 0$. The vertical profile of mass flux M_{ui} is parameterized by the buoyancy sorting of environmental and updraft air mixtures (see Appendix A).

The closure of the EDMF equations requires estimation of the updraft area fractions a_{ui} (see section 3.1), lateral entrainment rates ϵ_{ui} and estimation of the vertical structure of the updraft mass flux $M_{ui}(z)$ [*Neggers et al*., 2009]. The original formulation of EDMF is modified in the following, by introducing stochastic perturbations into the moist-updraft area fraction a_{u2}, and by reconstructing the vertical structure of the convective mass flux $M_{u2}(z)$ using a new multiupdraft method.

3.1. Perturbed Moist Updraft Area Fraction

A typical cumulus cloud fraction over the trade wind region is only a few percent, and the cloud-base height is always kept close to the mixed layer top as a result of the feedback mechanism between the mixed layer humidity, the convective mass flux and the mixed layer depth [*Betts*, 1976; *Neggers et al*., 2006; *van Stratum et al*., 2014]. Destabilization of the mixed layer by moistening results in the increase of the convective mass-flux out of the mixed layer and reduction of the mixed layer depth. After this phase, the system will adjust back to the previous state by decreasing the vertical mass transport and by entraining dry air at the top of the mixed layer, and thus the mixed layer depth will recover to its balanced state. This regulation mechanism is recognized as a change in depth of a weak inversion layer between the top of the mixed layer and cloud base, δ_{tr}, also called the transition layer depth, which reacts on the destabilizing effects within the boundary layer [*Betts*, 1976; *Bretherton et al*., 2004; *Neggers et al*., 2006]. Thus, the fraction of the updrafts a_{u2} that rise out of the mixed layer and form buoyant clouds is parameterized in EDMF as the ratio of the transition layer depth δ_{tr} and the subcloud mixed layer depth h:

$$a_{u2} = \left(\frac{\delta_{tr}}{h}\right)\frac{1}{2p+1} \quad (6)$$

where the parameter $p = 2.2$ is estimated from LES [*Neggers et al*., 2007]. The dry updraft area is then calculated as $a_{u1} = \mathcal{A}^{up} - a_{u2}$, where the total updraft area fraction $\mathcal{A}^{up} = 0.1$ is set as a model constant. The convective mass flux M_{ui} is then calculated following equation (2).

One of the main requirements of conventional deterministic mass-flux schemes is that the cloud area fraction has to be much smaller than the total grid cell area fraction. This requirement is not fulfilled in the gray zone where the model grid scale approaches the scale of individual clouds. That is why the definition of a_{u2} (equation (6)), as the QE closure assumption of EDMF, is valid only in the case where a model grid cell contains the cumulus ensemble with a robust statistical representation of clouds of all sizes. In this study we assume that equation (6) can be applied to a grid cell with size larger than ~50 km. For the model grids finer than ~ 50 km, fluctuations around the equilibrium value of a_{u2} increase with resolution and become substantial at the kilometre-scales. To represent these fluctuations, the stochastic cloud ensemble model of *Sakradzija et al*. [2015] is coupled to EDMF after the a_{u2} closure calculation (equation (6)). The closure assumption for the moist updraft area fraction (equation (6)) is assumed to be valid for the large area A around a model grid cell (Figure 1), and the outcomes of the subgrid moist updraft area fraction are subsampled from the QE cloud ensemble defined for the area A. This subsampling results in the perturbed moist updraft area fraction a_{u2}^p (see section 3.1.1), which is then passed to the EDMF parameterization to compute the cloud properties and subgrid tendencies at the actual model grid resolution. Thus, by the stochastic formulation of EDMF, the gray-zone dependence of the QE closure assumption is avoided, and equation (6) is used in the same form as in the deterministic EDMF scheme, but now applied on the scale where the statistical QE is considered valid.

In the stochastic EDMF scheme, the total updraft area fraction $\mathcal{A}^{up*}$ is grid-dependent, because the perturbed moist updraft area fraction a_{u2}^p can take values between zero and one. So if a grid column is likely to

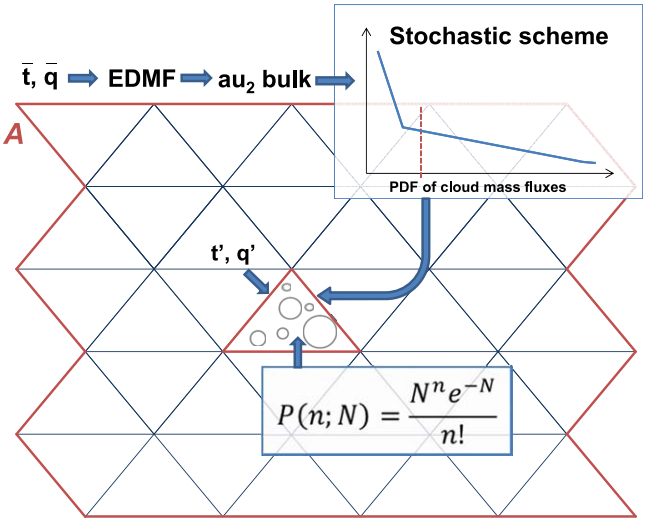

Figure 1. A schematic diagram of the stochastic shallow cumulus ensemble on the ICON model grid. The subgrid updraft area fraction is estimated as a random sum of the individual updraft area fractions $a^p_{u2}=\sum_{i=1}^{n} a_{ci}$ within each model grid cell. The number of clouds n that occupy the model grid cell is randomly sampled from the Poisson distribution. Here N represent the ensemble average of the total number of clouds in the cell. The EDMF closure is applied over the large-scale area around the grid cell (red outer rim), while the EDMF scheme updates the local grid cell thermodynamic properties t and q based on the random subsample of the cloud ensemble.

hold a cloud, the total convective area fraction in the stochastic formulation $\mathcal{A}^{up*}$ is allowed to grow according to the area fraction occupied by clouds

$$\mathcal{A}^{up*}=MAX(\mathcal{A}^{up},\ a^p_{u2}), \tag{7}$$

$$a^{u1}=\mathcal{A}^{up*}-a^p_{u2}. \tag{8}$$

The dry updraft area fraction a_{u1} is still not parameterized as scale-aware in the current model formulation, because the focus of this study is on the moist updraft and cloud parameterization component of EDMF.

3.1.1. Stochastic Shallow Cloud Ensemble

The stochastic model is a stationary canonical cloud ensemble framework that operates on a two-dimensional horizontal domain and keeps the memory of individual cloud lifecycles [*Sakradzija et al.*, 2015]. Coupling of the stochastic model to EDMF introduces a spectral cloud ensemble and a local memory component to the otherwise fully diagnostic and bulk EDMF cloud scheme. By the stochastic model formulation, the individual clouds are independent elements that reside on the model grid during their lifetimes. The convective response in every grid cell is calculated by summing the effects of the individual clouds that are assigned to a given cell, which then represents a random cloud subensemble. In that way, dependence on the grid resolution is included intricately in the model formulation, by counting the number of clouds that fall into the grid cell area. In the case where a single cloud area is larger than the model grid cell area, the cloud is allowed to spread across and cover the neighbouring grid cells. This latter feature is a step toward a parameterization in the gray zone for an individual cloud, where the communication between the neighbouring grid cells has to be implemented instead of a classic localized grid-column approach.

The stochastic model generates a compound distribution of the subgrid moist updraft area fraction a^p_{u2} [see *Sakradzija et al.*, 2015]. This compound distribution is the distribution of the random sum of individual cloud areas a_{ci} within the model grid cell $a^p_{u2}=\sum_{i=1}^{n} a_{ci}$. A random sample of clouds from the compound distribution is generated using the inverse transform method in two steps. First, a number of clouds n in the given grid cell is sampled from the Poisson distribution

$$p(n)=\frac{(G\Delta t)^n e^{-G\Delta t}}{n!},\quad n=0,1,2,\ldots \tag{9}$$

where G is the cloud production rate per model time step Δt and per model grid cell area. And second, the cloud area a_c of each sampled cloud is generated from the Weibull distribution

$$p(a_c)=\frac{k}{\lambda}\left(\frac{a_c}{\lambda}\right)^{k-1} e^{-(a_c/\lambda)^k} \tag{10}$$

where $\lambda>0$ refers to the distribution scale, and parameter $k>0$ is the distribution shape (see also Figure 1). The Weibull distribution used in this model configuration is a single-mode distribution that represents the active cumulus clouds. Passive clouds are parameterized by the diffusive transport and the corresponding cloud parameterization of EDMF. The Weibull distribution for active clouds is used to define the fluctuations of the moist updraft area fraction a_{u2}, and also mass flux M_{u2} through equation (2), around the QE convective state area fraction $\langle a_{u2}\rangle$. We assume that the ensemble average area fraction $\langle a_{u2}\rangle$ is equivalent to the bulk

updraft area fraction a_{u2} that results from the EDMF closure equation (6). As shown in *Sakradzija et al*. [2015], the resulting distribution of the perturbed a_{u2}^{p} is scale-aware, and its variance and skewness change with the model resolution because the number of initiated cloud elements in a grid cell depends on the grid cell area through the parameter G. Using this approach, a_{u2} is perturbed in a physically based manner, having the large-scale ensemble average of area fraction $\langle a_{u2} \rangle$ (equivalent to a_{u2} from the equation (6)) and average cloud lifetime $\langle \tau \rangle$ as physical constraints on the system following the closure equations:

$$\langle \tau \rangle = \alpha \lambda^{\beta} \Gamma \left(1 + \frac{\beta}{k} \right), \tag{11}$$

$$\langle a_{u2} \rangle = \frac{G\alpha}{w_*} \lambda^{1+\beta} \Gamma \left(1 + \frac{1}{k} + \frac{\beta}{k} \right), \tag{12}$$

as formulated in *Sakradzija et al*. [2015]. α and β are the coefficients of the cloud lifetime relation $\tau = \alpha (a_c)^{\beta}$ estimated from LES, and $\langle \tau \rangle$ is the constant average cloud lifetime (Table 1). The shape parameter of the Weibull distribution is set to constant $k = 0.7$ (Table 1). w_* is the domain average convective velocity scale calculated by the EDMF surface layer scaling formulation [*Siebesma et al*., 2007]. Thus, the two unknown parameters are λ and G and the system of equations (11) and (12) is closed.

3.1.2. Stochastic EDMF Flowchart

A time step of the stochastic EDMF parameterization starts with averaging of the prognostic variables that are input to EDMF as shown on the flow chart, Figure 2. The averaging is performed over the large area around each grid cell in ICON (Figure 1), which is currently the full simulation domain (see section 4). This is justified by the uniformity of the LES RICO-140 cloud field [*Sakradzija et al*., 2015, Figure 1]. The EDMF scheme is called twice (Figure 2). The first call of EDMF calculates the large-scale closure for the moist updraft area fraction a_{u2} and the convective velocity w_* averaged over the large-scale area A (Figure 1). After the closure point, the bulk moist updraft area fraction a_{u2} and convective velocity w_* are passed to the stochastic model closure equations (equations (11) and (12)). The stochastic model is run for the current time step to perturb a_{u2} by producing a distribution of subgrid convective area fractions a_{u2}^{p} that is dependent on the model resolution. The perturbed area fraction a_{u2}^{p} is then provided as an input to the second EDMF call, initialized with the nonaveraged prognostic fields. In the second call, the EDMF scheme continues further from the a_{u2} closure point to compute the moisture and heat budgets and estimate the vertical structure of the cloud layer for each grid cell in the domain following equations (3) and (4).

3.2. Vertical Structure of the Cloud Layer

When the stochastic model is implemented in the 3-D ICON model, the vertical structure of cloud layer also has to be reformulated. So in this section, we propose a preliminary formulation of the parameterization of the cloud layer vertical structure in the stochastic EDMF framework.

The rising plume budget equations, equations (3) and (4), are applied to both bulk updrafts in the EDMF scheme. *Neggers et al*. [2009] argue that such an approach is feasible only if the updraft lateral entrainment rate is dependent on the state of the updraft. As in *Neggers et al*. [2002] the dependence on the state of the updraft is achieved by parameterizing the updraft lateral entrainment rate as inversely proportional to the updraft velocity:

Table 1. , Parameters of the Stochastic Cumulus Ensemble Model: the Shape Parameter of the Weibull Distribution k, α and β Parameters of the Cloud Lifetime Relation, and Ensemble Average Cloud Lifetime $\langle \tau \rangle$

Parameter	Value	Unit
Domain size	410^2	km^2
k	0.7	
α	0.33	s/m^2
β	0.72	
$\langle \tau \rangle$	720	s

$$\epsilon_{ui} = \frac{1}{\tau_\epsilon w_{ui}} \tag{13}$$

In the deterministic EDMF formulation, the entrainment turnover timescale τ_ϵ is constant and is estimated from LES as $\tau_\epsilon = 400$ s [*Neggers et al*., 2002]. By this entrainment formulation, an updraft that is rising faster through the model grid layer will have less time to mix with the environment and thus will be less diluted, and vice versa. The same updraft entrainment rate is applied to the moist updraft in the subcloud and cloud layers.

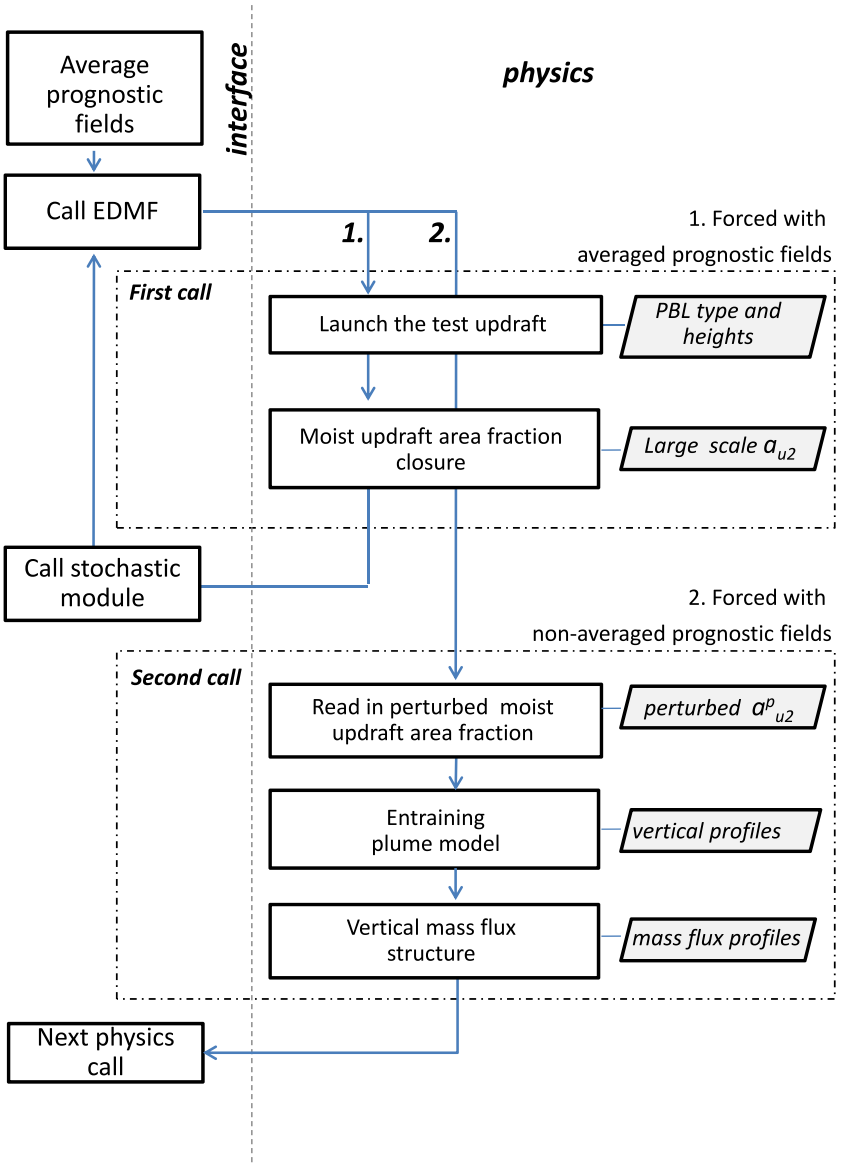

Figure 2. The flow chart of the stochastic EDMF scheme in ICON during a single model physics time step. The ICON model consists of the model dynamics level where the prognostic equations are discretized and solved (left side) and the model physics level where the contributions of the subgrid processes are calculated by using different parameterization schemes, which are mainly diagnostic (right side). The two model components are coupled at the interface level, at which the exchange of model variables between the model dynamics and physics takes place.

The updraft budget equations, equations (3) and (4), can in principle be applied to a bulk convective updraft within a bulk convection scheme or to a single updraft in a multiupdraft approach. On the gray-zone scales, where the convective ensemble is subsampled, the parameters in the bulk updraft budget equations are unknown. The model parameters that are defined specifically for the case of a robust statistical sample of an updraft ensemble are: the entrainment rate for the conserved quantities ϵ, the entrainment rate for the updraft vertical velocity ϵ_w, and the proportionality parameter μ in the updraft velocity equation (equation (4)). All these parameters are dependent on the spatial scales of the model and are causing the scale-dependence of the cloud layer vertical structure. We do not provide here the solution to scale-dependence of all these parameters, and the formulation of the original scheme is retained in the stochastic model formulation. The only difference is in the entrainment time scale, which is tuned to $\tau_\epsilon = 320$ s for all resolutions in both the deterministic and the stochastic model configurations. As will be shown in section 4,

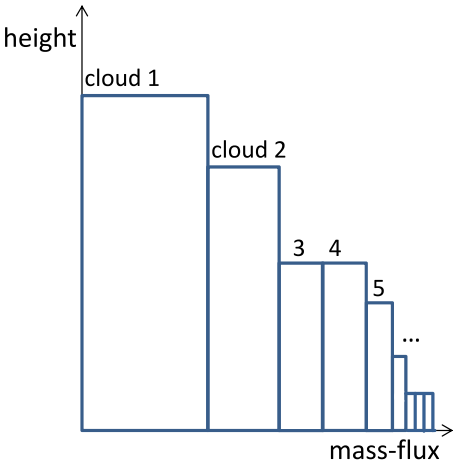

Figure 3. A sketch of the subgrid mass flux vertical profiles. Individual clouds have different heights, which are calculated as dependent on the cloud area $h_i \propto a_{ci}, i=1,n$. It is assumed that the mass flux of each cloud is constant with height. Depending on the random number of clouds and their areas within a grid column, the total cloud mass flux profile will have various shapes.

the original EDMF entrainment formulation and the tuned parameter τ_ϵ still cause some scale-dependence of the cloud layer properties.

In the parameterization at hand, the vertical mass-flux profile is the key quantity that influences scale-dependence of the cloud-layer vertical structure. Thus, it is necessary to reformulate the estimation of the vertical mass-flux profile in the stochastic convective scheme. Instead of using the buoyancy sorting mechanism to estimate the mass flux vertical profiles, which is the approach used in the deterministic EDMF (see Appendix A), we collect the individual cloud area fractions from the stochastic cumulus ensemble and construct the vertical structure of a random sum of cloud area fractions within each model grid column. It is assumed that individual clouds occupy areas constant with height (Figure 3), while the height of a cloud depends on the cloud area following the relation $h_i = c_1 (a_{ci}/a_0)^{0.25}$, $c_1 = 75$ m, $a_0 = 1$ m^2 estimated from LES and additionally tuned to match the cloud layer depth determined in EDMF. So the random sums of subgrid cloud areas show the height dependent structure (Figure 3). Finally, the vertical profiles of the cloud mass flux $M_{ui}(z)$ are calculated using equation (2). This approach can be seen as a preliminary but appropriate parameterization of the vertical mass flux structure that is consistent with the stochastic cumulus ensemble framework.

4. Numerical Simulations

A single column model (SCM) configuration is commonly used to test the newly developed parameterizations. However, this approach might only be appropriate in cases where the subgrid processes are parameterized assuming no direct communication or correlation between the neighbouring model columns, and only in those cases where model dynamics is not interfering directly with the processes that are being parameterized. Because of the high model resolutions used in this study, convection develops on the grid scale, so the standard SCM setup does not provide the complete picture of the convective processes at different scales and their interactions.

The LES of RICO case is forced by the homogeneous and constant large-scale tendencies in a doubly periodic domain, so the entire model domain can be considered equivalent to a SCM. Without the effects of model dynamics, in a SCM-like but spatially extended setup, the deterministic EDMF parameterization would provide identical outcome for all independent grid columns. In the case of the stochastic parameterization, the entire model domain can be considered equivalent to a SCM setup for testing the parameterization capability to predict the macroscopic state of the convective system, while the stochastic model is applied across this domain for the probabilistic modeling of the microstates of the system.

In this section, we perform deterministic and stochastic model simulations using the ICON model on the double-periodic domain of around 410^2 km^2. The deterministic parameterization is tested only in model configuration with full dynamics. The stochastic parameterization is tested first in a SCM-like setup where interactions with model dynamics are excluded. Second, the stochastic model is allowed to interact with the model dynamics. This provides a way of testing and quantifying the effects of grid-scale convective motions and their interaction with the stochastic parameterization of convection.

All experiments are conducted multiple times using a range of horizontal resolutions from around 1 to 50 km (Table 2). In all experiments the vertical grid spacing is stretched with a minimum layer depth of 100 m near the surface and with a total of 50 levels up to 10 km height. Near the top boundary of the modeling domain, the sponge layer is set for wave damping starting at the height of 7 km. The ICON simulations are compared to the UCLA-LES simulation of the RICO-140 case to assess their performance.

4.1. Vertical Structure of the Convective Boundary Layer

The domain-averaged vertical profiles of the RICO boundary layer compare well with the LES for all three EDMF configurations, the deterministic EDMF, the stochastic EDMF parameterization in a SCM-like setup and the stochastic EDMF parameterization in the ICON model including model dynamics (Figure 4). The vertical structure of temperature and moisture, and the wind profiles in the boundary layer are equally well modeled on all tested resolutions and there is almost no resolution dependence of the simulations. The most prominent resolution dependence is present in the vertical profiles of wind components in the stochastic simulations using the model configuration with the grid-scale dynamics. The mean wind profiles are very similar to the LES profiles but are degraded for the highest resolutions (Figure 4). The reason for this comes from the interaction of the stochastic parameterization with model dynamics, which makes the existing EDMF formulation of the entrainment rate inadequate for the vertical momentum transport in the boundary layer on high model resolutions.

Cloud fraction and liquid water content are overestimated by the deterministic EDMF scheme (Figure 5), however, this is a known limitation of a bulk approach, and it is evident in other models as well (see *Neggers* [2009, Figure 14], which results from the case without updraft precipitation). This result confirms that the bulk approach is not sufficient for parameterizing cloudiness in transient cases such as RICO [see also *Neggers*, 2015], but can nevertheless capture the boundary layer structure (Figure 4).

The deterministic EDMF scheme shows only a slight dependence on the model resolution in the cloud layer, even though it is not formulated as scale-aware. A reason for weak scale dependence lies in the homogeneous and constant forcing of the RICO case. The EDMF closure formulations of the vertical mass flux structure and the entrainment rate are impaired on the kilometre-scale resolution (see Appendix A, Figure A1). However, this closure formulation, inadequate for high resolutions, affects the boundary layer structure only if inhomogeneities are present in the modeling domain. In the ICON model, spatial inhomogeneities are present and are caused by convective circulations that emerge on the model grids of resolutions finer than 10 km (see section 4.4). These inhomogeneities influence the parameterization of the cloud layer vertical structure and decrease the vertical profiles of cloud mass flux on the high resolution model grids (Figure 5, bottom). As a result of the slightly reduced mass flux in the high resolution simulations, the vertical structure of the cloud fraction and cloud liquid water also show a reduction on all vertical levels (Figure 5).

In the stochastic simulations, the vertical profiles of cloud fraction a_c and cloud liquid water content q_c demonstrate a decrease in cloudiness with increasing resolution while approaching the shape of the LES RICO profile (Figure 5). In a similar way as in the deterministic simulations, this reduction in a_c and q_c results from the changes in the mass flux profile with resolution (Figure 5, bottom). The formulation of the mass flux profile in the stochastic EDMF scheme is scale adaptive because it is based on the random cloud ensemble subsampling (see section 3.2). On the coarse model grids, the stochastic scheme acts similar to a bulk scheme limited by the moist-plume bulk approach of EDMF. At higher resolutions, the stochastic scheme approaches the behavior of a spectral scheme by exploiting the stochastic spectral cloud ensemble more efficiently. So the liquid water content and cloud fraction profiles change from the bottom-heavy in the coarse resolution simulations to the top-heavy profiles on the high resolutions. It also has to be noted that the structure of the cloud layer on different grids is affected by the entrainment formulation, which is still not developed as scale-aware. So the inadequate formulation of entrainment might also be slightly responsible for the resolution dependence of the mass flux profiles and subsequently of the a_c and q_c.

The stochastic simulations that include model dynamics converge faster toward the LES profiles with the increase of model resolution than the SCM-like stochastic simulations. Here the under-resolved circulations apparently have a positive impact and their effects on the mean structure of the cloud layer are not negligible (Figure 5). These changes in the profiles come mostly from the transition of the stochastic parameterization behavior from a bulk-like to a spectral-like scheme (see the next section). In difference to the deterministic simulations where the grid-scale convective circulations impacts the parameterization only on the scales smaller than 10 km, the effects of the convective

Table 2. Model Horizontal Resolutions and Time Steps

Triangle Edge (km)	Rectangular Edge (km)	Number of Points	Time Step (s)
2.43	1.6	168 × 196	10
4.86	3.2	84 × 98	20
9.72	6.4	42 × 48	30
19.45	12.8	20 × 24	50
38.90	25.6	10 × 12	80
77.81	51.2	8 × 8	100

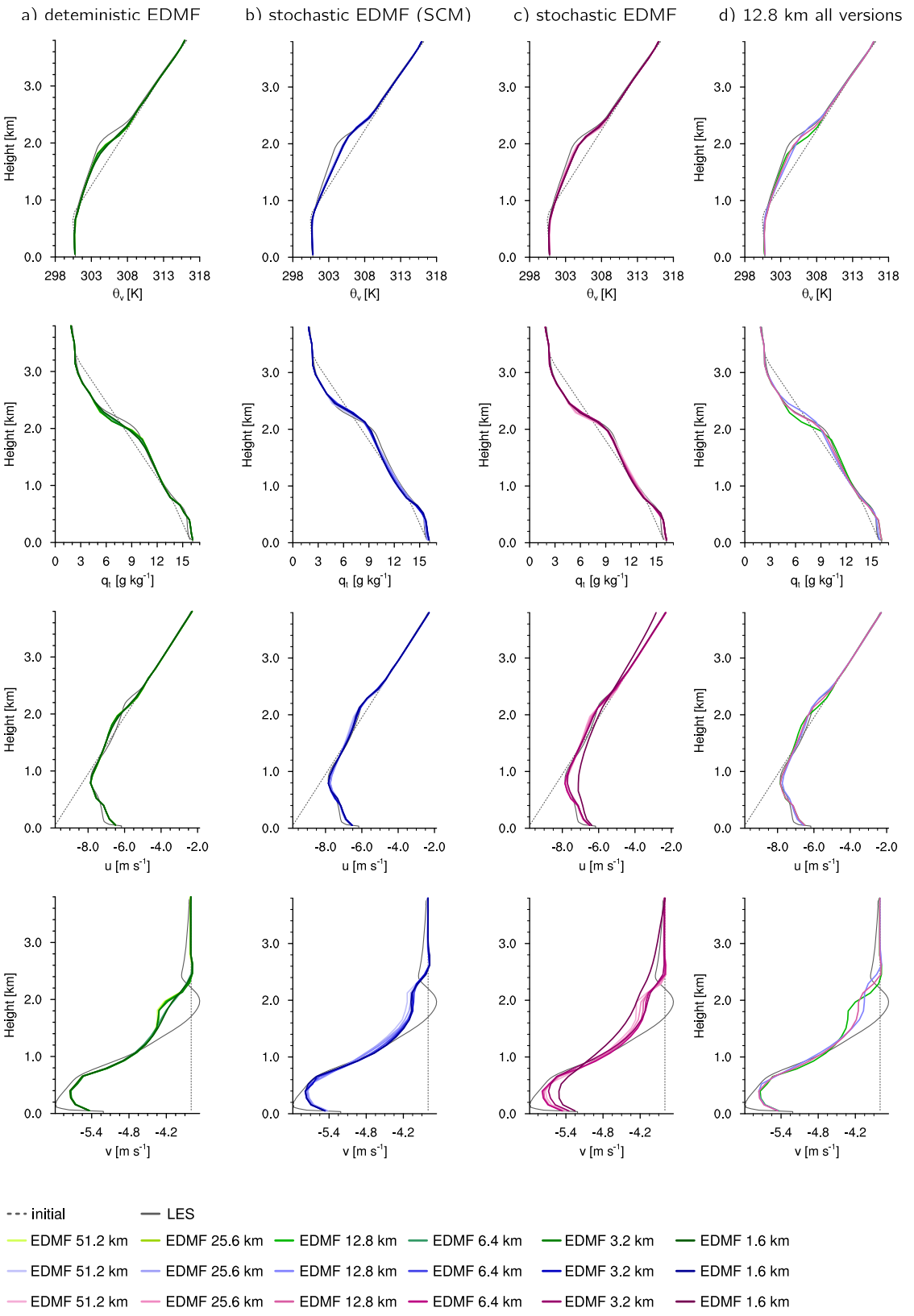

Figure 4. Boundary layer vertical profiles of the nonprecipitating RICO case averaged over the time period from 20 to 24 h. The vertical profiles of virtual potential temperature θ_v, total water mixing ratio q_t, zonal wind component u, and meridional wind component v are plotted for the three model configurations: the deterministic EDMF scheme (green), the stochastic EDMF scheme in a SCM (blue), and the stochastic EDMF scheme in the setup that includes model dynamics (red). The LES profiles are plotted in thin black lines.

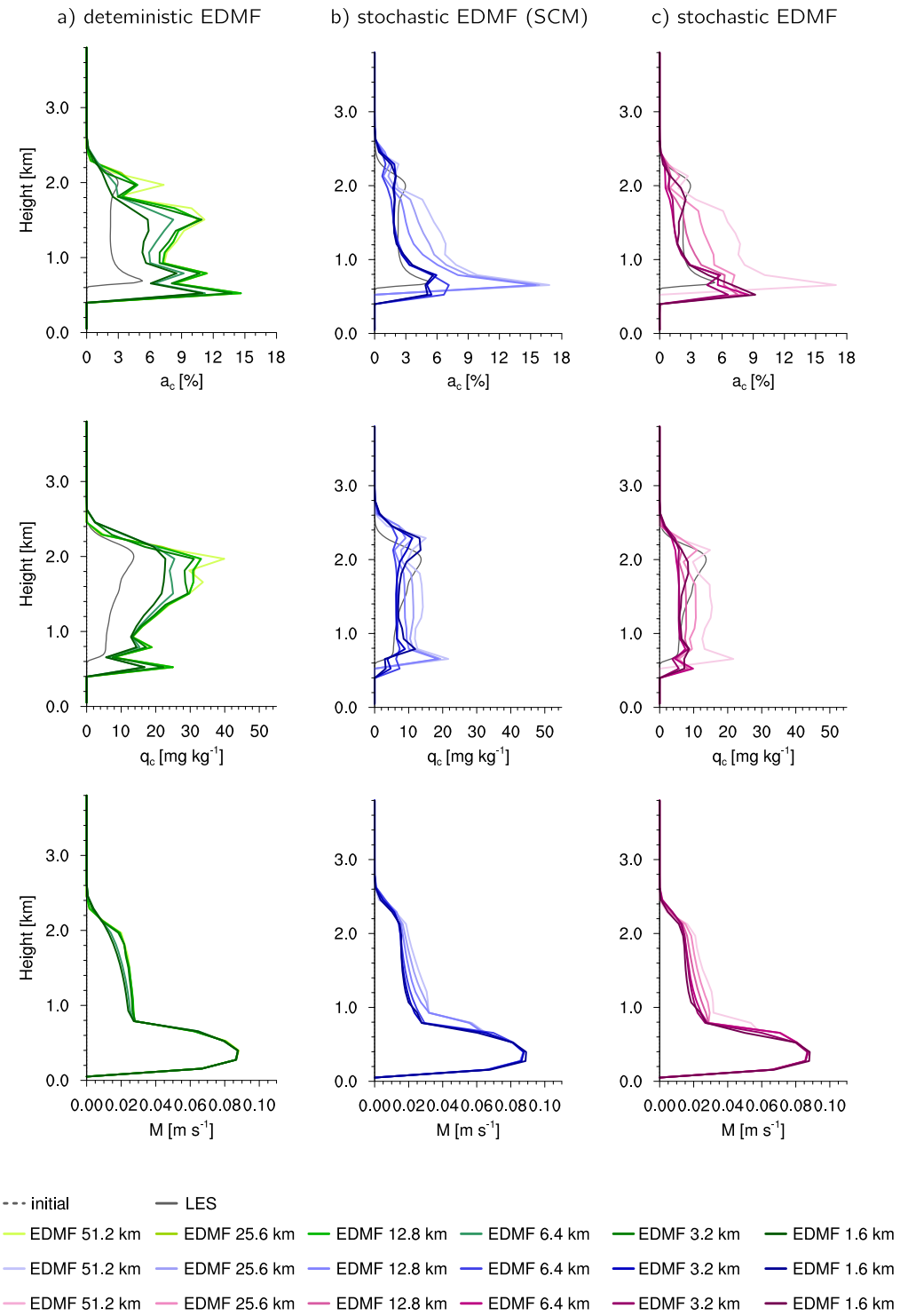

Figure 5. Boundary layer vertical profiles of the nonprecipitating RICO case averaged over the time period from 20 to 24 h. The vertical profiles of cloud fraction a_c, cloud liquid water content q_c, and updraft mass flux M are plotted to show the resolution dependence of the three model configurations: the deterministic EDMF scheme, the stochastic EDMF scheme in a SCM, and the stochastic EDMF scheme in the setup that includes model dynamics. The LES profiles are plotted in thin black lines. We do not compare here the updraft mass flux to the LES results.

circulations that develop on the model grid act on all grid resolutions tested in the stochastic simulations (Figure 6). Cloud fraction is mostly affected near the cloud base, where the absolute difference can be up to 9%. Cloud liquid water content shows a change of up to 10–12 mg kg^{-1} at the cloud base and cloud top levels.

Turbulent fluxes also change by the influence of model dynamics by around 20–50 Wm^{-2} in absolute value, and this change in the turbulent fluxes reflects in a change in virtual potential temperature by 0.2–0.6 K and in total water content by up to 1.2 g kg^{-1}. The largest changes in the wind profiles are on the 1.6 km resolution with difference in magnitudes from 0.3 to 0.6 m s^{-1}.

4.2. Time Evolution

The time series of cloud cover and cloud liquid water path are plotted on Figure 7. It is evident that the deterministic EDMF scheme overestimates cloudiness on all model grids. Besides this, the behavior of the deterministic EDMF scheme is intermittent and the time series show abrupt jumps at times (Figure 7). This behavior can be avoided by including the generation of precipitation from convecting clouds in EDMF, however, because liquid water is overestimated in RICO simulations, precipitation amount will as well be overestimated. As already stated in section 2, we choose to study a nonprecipitating convective case instead. Furthermore, the solution of the model equations is distinct across different resolutions and this difference is the result of diverging model dynamics and its interaction with the convective parameterization.

The stochastic EDMF captures the RICO time series better than the deterministic scheme and clearly smooths the abrupt peaks in the time series of cloud cover and liquid water path on all model resolutions (Figure 7). The stochastic scheme converges to the LES solution more effectively with the increase of model resolution, because it interacts more actively with the model grid scale flow (see section 4.4). On the coarse resolution grids (12.8 - 25.6 km) the stochastic scheme performs in a similar manner as the bulk EDMF deterministic scheme where the cloudiness is overestimated. In the model configuration with full dynamics, the model switches between the states of overestimated to slightly underestimated cloud cover. On the intermediate scales, around 6.4 km resolution, the two regimes interchange in both simulations using the stochastic model configuration. And finally, on the high resolution grids (1.3 - 3.2 km), the stochastic spectral cloud model becomes more effective and the convection stays in the regime with the correct solution during the whole simulated period. Very similarly to the behavior of the vertical cloud structure (section 4.1), the high resolution simulations show the full benefit of using the spectral stochastic cloud model for the time evolution of the cloud field.

4.3. Convective Variability Across Scales

Based on the coarse-graining study of the LES RICO case [*Sakradzija et al*., 2015], the statistical QE can be considered valid on grid resolutions coarser than approximately 20 km. It should be emphasized that this conclusion holds in a slowly changing convective environment and in the case where the effects of nonlinearities in model dynamics are negligible. In this section, we show how the fluctuations around QE develop in ICON in the deterministic and in the stochastic RICO simulations forced by the constant and spatially homogeneous large-scale forcing tendencies.

Variability of the subgrid convective states in ICON increases with model resolution (Figure 8). The histograms resulting from simulations using the deterministic EDMF scheme are narrow and very similar across the grid resolutions coarser than 10 km (not shown here). In the deterministic simulations, on the model grids with resolution finer than 10 km, the variability of the subgrid cloud fraction is overestimated. The shape and spread of the histograms of cloud fraction differ significantly from the histograms of the coarse-grained LES. The shape of the histograms is irregular and it shows two peaks associated with the updraft and downdraft regions of the convective rolls. Hence, the spatial variability of cloud fraction found in the deterministic simulations results solely from the under-resolved convectively induced circulations that develop on the model grid and control the behavior of the convective parameterization (see section 4.4). The second mode of the histograms forms too long tails and, in some cases, reversed skewness compared to the coarse-grained LES. The spread of the histograms partly also originates in the time variations of the mean cloud fraction (section 4.2), because samples of cloud fraction are collected from snapshots taken every 15 min during 4 h of the deterministic simulations.

The histograms of cloud fraction resulting from the stochastic SCM-like simulations resemble the LES histograms well and show an increase in variance and skewness with resolution (Figure 8). Compared to deterministic simulations, the stochastic scheme improves the histograms and their scale-awareness. However, the variance of cloud fraction is overestimated in the stochastic simulations, especially on the highest resolutions where the peak at the zero value of cloud fraction is too high compared to the peak value of the LES

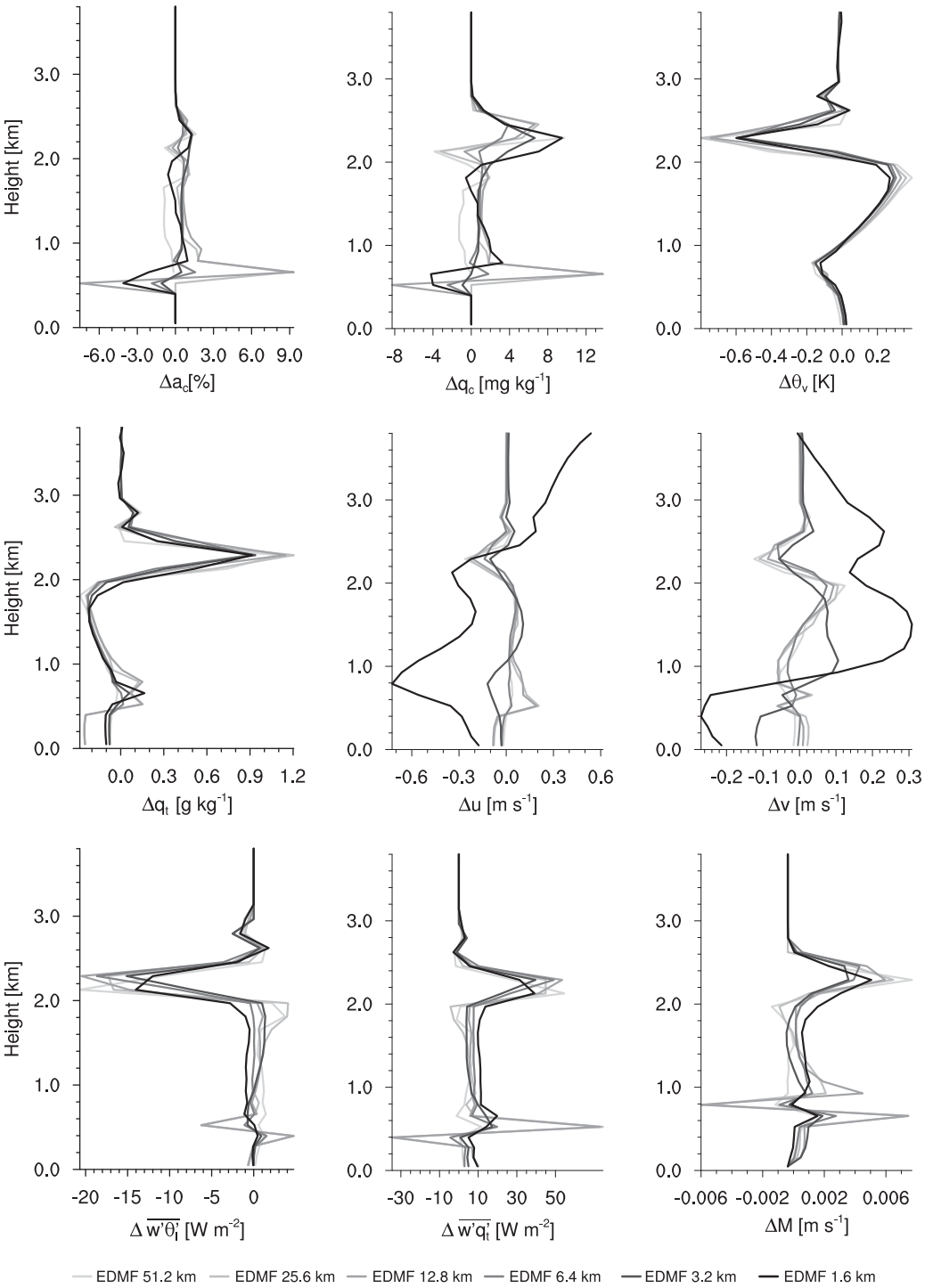

Figure 6. Absolute difference in the boundary layer vertical profiles between the stochastic RICO case modeled in the SCM ICON and the ICON configuration with the full dynamics. The figures show the difference in the profiles of cloud fraction a_c, cloud liquid water content q_c, virtual potential temperature θ_v, total water mixing ratio q_t, zonal wind component u, meridional wind component v, sensible heat flux $\overline{w'\theta_l'}$, latent heat flux $\overline{w'q_t'}$, and updraft mass flux M.

coarse-grained histograms (Figures 8a and 8b). The reason for this overestimation partly lies in the estimation of the boundary layer type in EDMF, which is a switch that turns moist shallow convection on or off based on some criterion. Hence, the stochastic model does not operate across all grid cells but only those

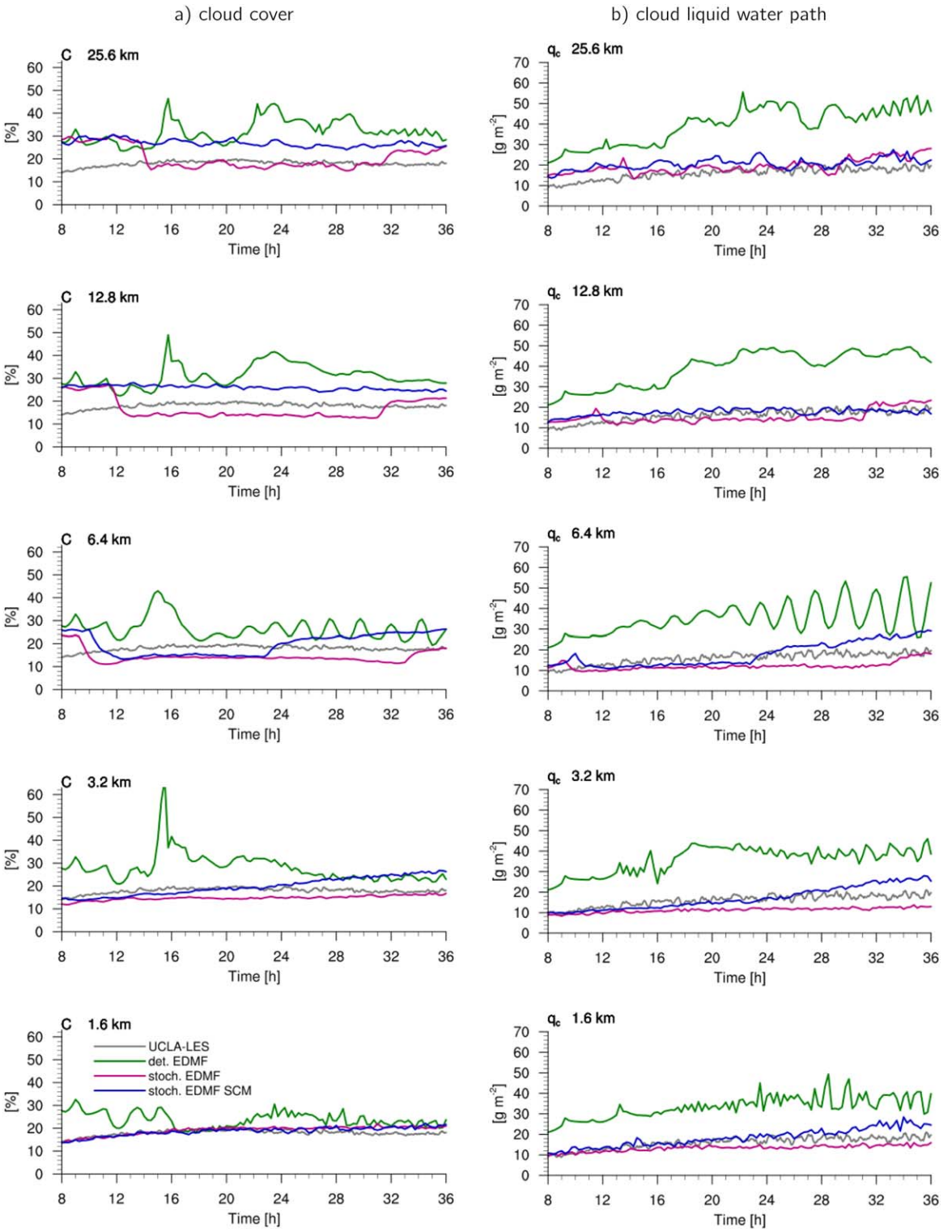

Figure 7. The time series of cloud cover and cloud liquid water path in the RICO case, showing the deterministic EDMF (green), the two stochastic EDMF simulations (blue and purple), and the LES RICO-140 time series (gray). The LES time series are computed by averaging the fields of the original 25 m resolution over the domain and are the same on all graphs corresponding to different ICON resolutions. The SCM stochastic EDMF simulations are plotted in blue color, while the simulation using the full model dynamics are plotted in purple colour.

identified as moist convective cells by EDMF. Another reason for overestimation of variance might be in separation of the cloud field into convective and diffusive modes, while the stochastic scheme acts only on the convective mode of the cloud ensemble.

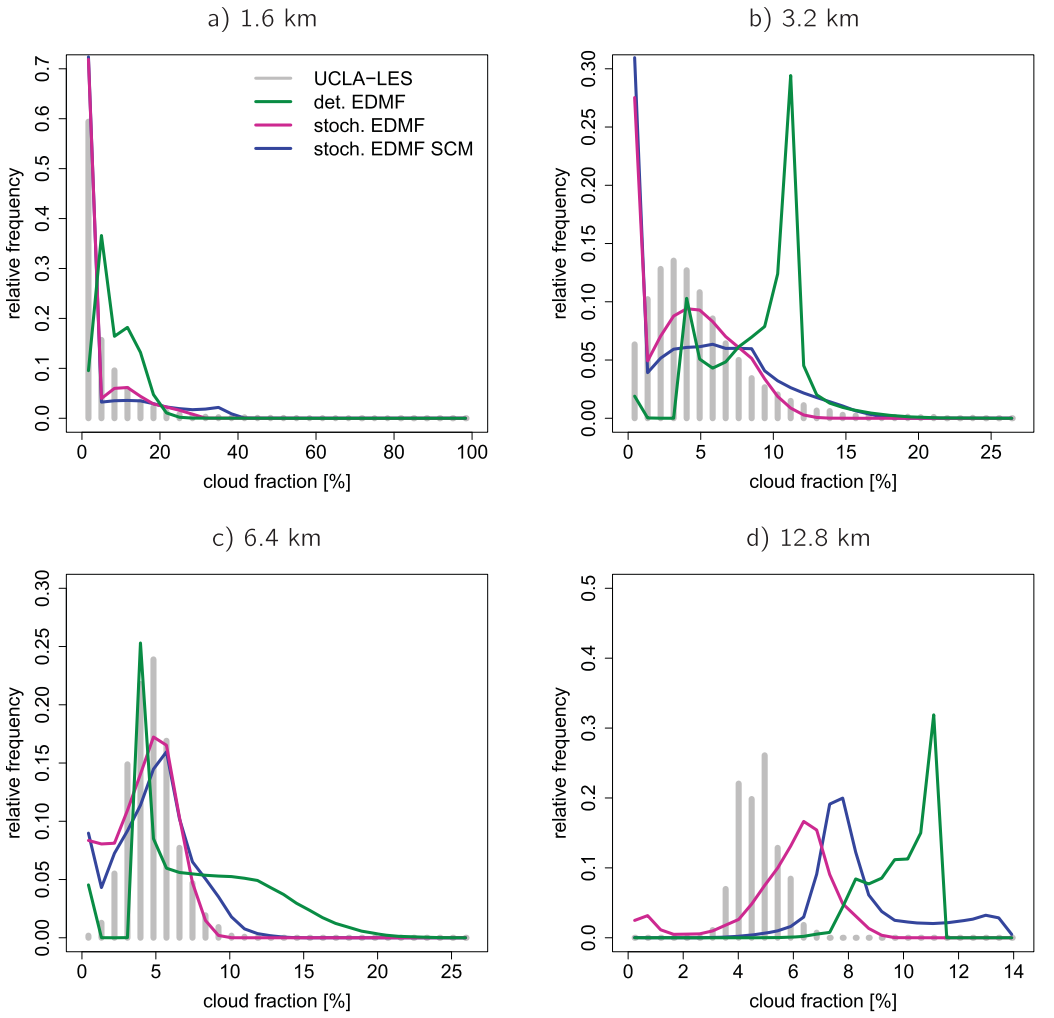

Figure 8. Histograms of the cloud fraction above the cloud base across different grid resolutions. The histograms computed from the coarse-grained cloud fraction of LES RICO-140 are used as a reference case. The histograms resulting from the deterministic EDMF (green), the stochastic EDMF in the SCM setup (blue) and the stochastic EDMF in the ICON setup that includes model dynamics (purple) are compared to the reference histograms.

The histograms resulting from the stochastic simulations in the model configuration with active grid-scale dynamics are improved further and match the LES histograms better than in the SCM-like configuration (Figure 8). This additional improvement comes from the grid-scale flow and emergence of convective circulations that impose a spatial preference in EDMF for stronger convection in the updraft regions, and also because stochastic parameterization improves the grid-scale convective flow regime (see section 4.4). These results imply that the SCM configuration of numerical models is not appropriate nor sufficient for testing convective parameterizations, because the interaction of parameterization with model dynamics can have significant implications for the test outcome. On top of this, the LES coarse-grained diagnostics would be sufficient for training parameterizations only if the effects of model dynamics are negligible, which is not the case in the RICO cloud field on a wide range of model grid resolutions.

4.4. Grid-Scale Dependent Secondary Circulations

As demonstrated in the previous sections, the emergence of grid-scale circulations has a strong effect on the simulation outcomes. Here we describe the circulations that emerge in the convective gray zone, and we emphasize the difference in the convective flow regimes caused by the choice of the deterministic or stochastic convective parameterizations.

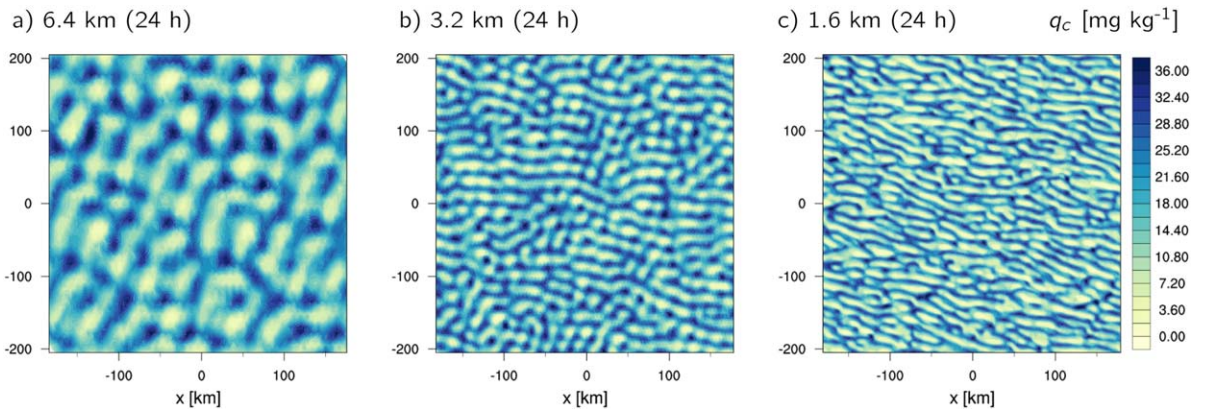

Figure 9. Horizontal snapshots of the cloud liquid water content in the RICO case at a level above the cloud base. The simulations are carried out using the deterministic EDMF scheme configuration. Snapshots are taken after 24 h of the three simulations on the grids of different horizontal resolutions.

4.4.1. Convective Regimes in the Deterministic Simulations

On the model grids coarser than 10 km, the deterministic EDMF parameterization produces horizontally homogeneous thermodynamic fields of the RICO-140 case as a result of the homogeneous surface conditions and of the constant and homogeneous large-scale forcing tendencies (not shown here). In contrast to the coarse grids, in the gray zone, spatial inhomogeneities develop as a result of the secondary convective circulations that self-organize into convective rolls (Figure 9) [see also *LeMone et al*. 2010; *Ching et al*., 2014]. Organization of convective motions on the model grid is similar in mechanism to the organization of convection in nature. In a convective boundary layer in nature, convective and wind shear instabilities induce perturbation waves that organize convection into roll patterns and dictate the direction and orientation of the rolls along the wind direction, see e.g., *LeMone* [1973] or *Brown* [1980]. However, in the numerical simulations, artificial truncation of the scales of motion and the homogeneous large scale forcing affect the spatial scales and evolution of the organized convective motions.

On the LES model grids, convection and clouds are effectively resolved and convective organization can take place in a thermally and mechanically driven convective flow under the constant large-scale forcing, such as in RICO. The LES of the RICO-140 case simulates the emergence and development of convective rolls that are well defined after a day of simulation time [e.g., *Sakradzija et al*., 2015, Figure 1]. These rolls have spatial scales that are comparable to the scales of natural convection and are controlled by the CBL depth. The wavelength of the natural convective rolls is around 2–4 times the CBL depth [*Brown*, 1980]. As the LES model resolution coarsens, the model grid scale becomes comparable to the dominant convective scale and convection is no longer effectively resolved, so the roll organization of convection develops spurious properties [*Piotrowski et al*., 2009; *Zhou et al*., 2014].

A similar behavior occurs in the NWP models, where the amplitudes and spatial scales of the convective organization depend on the model grid resolution because of the artificial truncation of the flow at the grid scale (Figure 9) [*LeMone et al*., 2010; *Ching et al*., 2014], while the homogeneous large-scale forcing allows for the undisturbed competition between the perturbation modes [e.g., *Drazin*, 2002]. The undisturbed conditions allow for development of multiple stationary states with different convective patterns, which develop in a sequence and remain stable for limited time periods. The convective system first develops into convective rolls along the wind direction until another dominant mode takes over and influences the roll orientation, so the rolls become positioned almost perpendicular to the mean flow after several hours from their emergence (Figure 9c). The convecting system does not remain steady for a long time and it eventually develops into a very regular stripe pattern with a larger wavelength, if the simulation is run long enough (not shown here). This is a typical behavior of nonlinear flows with the low Reynolds numbers far from the onset of convection [e.g., *Drazin*, 2002]. Hence, these under-resolved convective regimes in the gray zone are unrealistic and their emergence and further development should be suppressed. A suggestion how to suppress the under-resolved convective rolls is proposed in *Ching et al*. [2014]. In the following, we show that the stochastic parameterization, developed to adequately represent shallow cumulus clouds in the gray zone, also suppresses the artificial roll organization. Furthermore, instead of removing

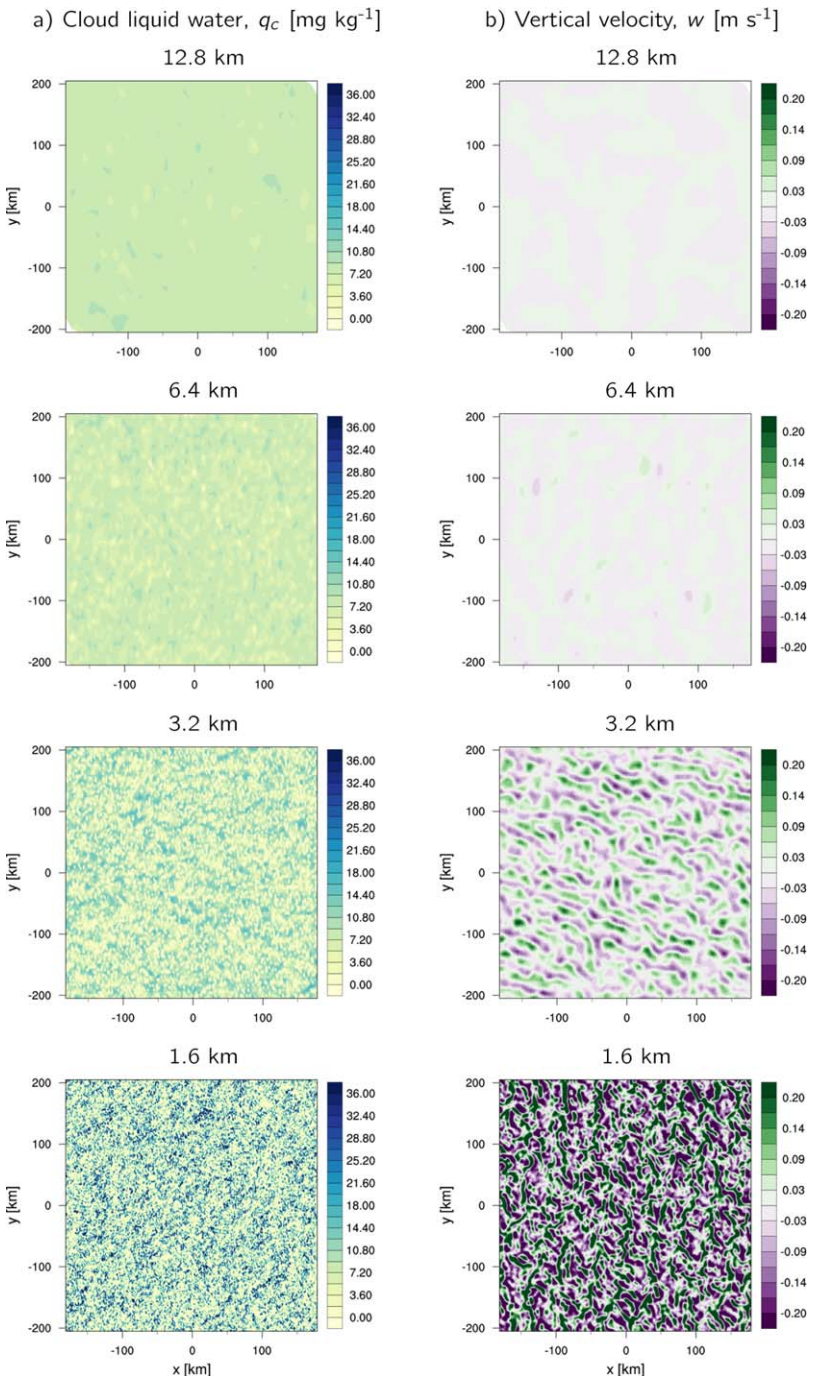

Figure 10. Horizontal snapshots of the cloud liquid water content and vertical velocity in the RICO case simulated using the stochastic EDMF scheme. The snapshots are taken at a level above the cloud base after 24 h of the simulations. The simulation of RICO is repeated using different horizontal resolutions, from around (top) 12.8 km to (bottom) 1.6 km.

convective circulations completely, the stochastic scheme has a potential to improve their representation. This might be important for modeling of deep convection and mesoscale convective processes on the kilometre-scale grids.

4.4.2. How Does the Stochastic Parameterization Affect the Convective Flow?

The horizontal cross sections of liquid water content (Figure 10a) and vertical velocity (Figure 10b) are taken at a level above the cloud-base after 24 h of simulation time using the stochastic EDMF. The variability of the subgrid liquid water content increases with model resolution, as predicted by the coarse-grained LES.

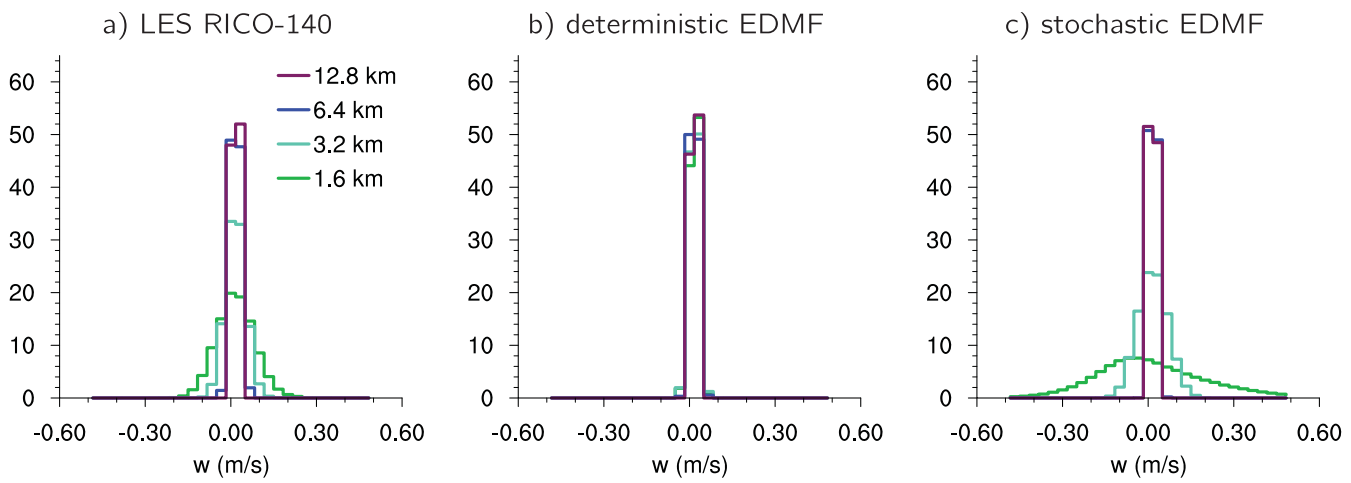

Figure 11. Relative frequency histograms of vertical velocity from the coarse-grained LES, the deterministic EDMF ICON, and the stochastic EDMF ICON simulation. LES histograms are computed based on the coarse-grained vertical velocity to mimic the different ICON horizontal resolutions. Time frame from 20 to 24 h is taken as the time period during which all three simulations develop convective circulations on the grid scale.

This variability is introduced through subsampling of the stochastic cloud ensemble, but is also influenced by the convective circulations. The emergence of convective circulations is supported by the stochastic fluctuations that introduce inhomogeneities into the convective field, so the circulations emerge immediately from the simulation start. The structure of circulations is distinct from the deterministic simulations (Figure 9) and the spurious organization of convection into roll structures is now suppressed. The circulations become significantly stronger with the increase in model resolution (Figure 10b), which is a direct consequence of the increase in fluctuation amplitudes.

Stronger circulations on higher resolutions are an expected behaviour based on the coarse-grained LES statistics. Such resolution dependence is reflected in the increasing variance of the subgrid vertical velocity as the model grid resolution becomes finer, which can be estimated by the spread of the histograms on Figure 11. The stochastic simulations produce convective circulations with the vertical velocity variance (Figure 11c) that shows an increase with resolution similar to the coarse-grained LES vertical velocity (Figure 11a). However, the variance of vertical velocity in the stochastic simulations is overestimated compared to the coarse-grained LES histograms. A better match could be achieved by tuning one of the parameters in the stochastic model, the distribution shape parameter k or the average cloud lifetime $\langle\tau\rangle$, in order to compensate for the additional variance that inevitably arises from other components of the EDMF parameterization. Even though the stochastic model is not tuned here to match the LES variance exactly, the stochastic simulations represent the LES statistics much better compared to the deterministic simulations. In the deterministic simulations, the convective circulations are by an order of magnitude too weak and the variance of the vertical velocity is too low compared to LES (Figure 11b).

In the gray-zone regime of the simulated shallow convective case in ICON, the convective kinetic energy of the flow is shifted toward larger scales compared to the coarse grained and effectively resolved LES cases (Figure 12). The energy-carrying scales in the ICON simulations with grid resolution of 1.6 km are around 10 km in the wavelength, while the coarse grained LES shows the spatial scales of around 3 km for the same grid resolution. The effectively resolved convection in LES on the grid of 25 m resolution develops eddies that are the most energetic at the scales between 10^2 m and 10^3 m. Both deterministic and stochastic simulations are in this respect equally limited by the grid scale and the effective resolution of ICON. The input of energy is at the scales permitted by the model resolution, so the spatial scales of convective circulations are larger than expected from the coarse-grained LES (Figure 12b, left).

The convective energy of the flow is highest in the LES simulation with 25 m resolution, while the coarse-grained LES shows by two orders of magnitude lower energy (Figure 12b). The flow in the deterministic ICON simulation is by two orders of magnitude less energetic than the reference coarse-grained LES (Figure 12a,

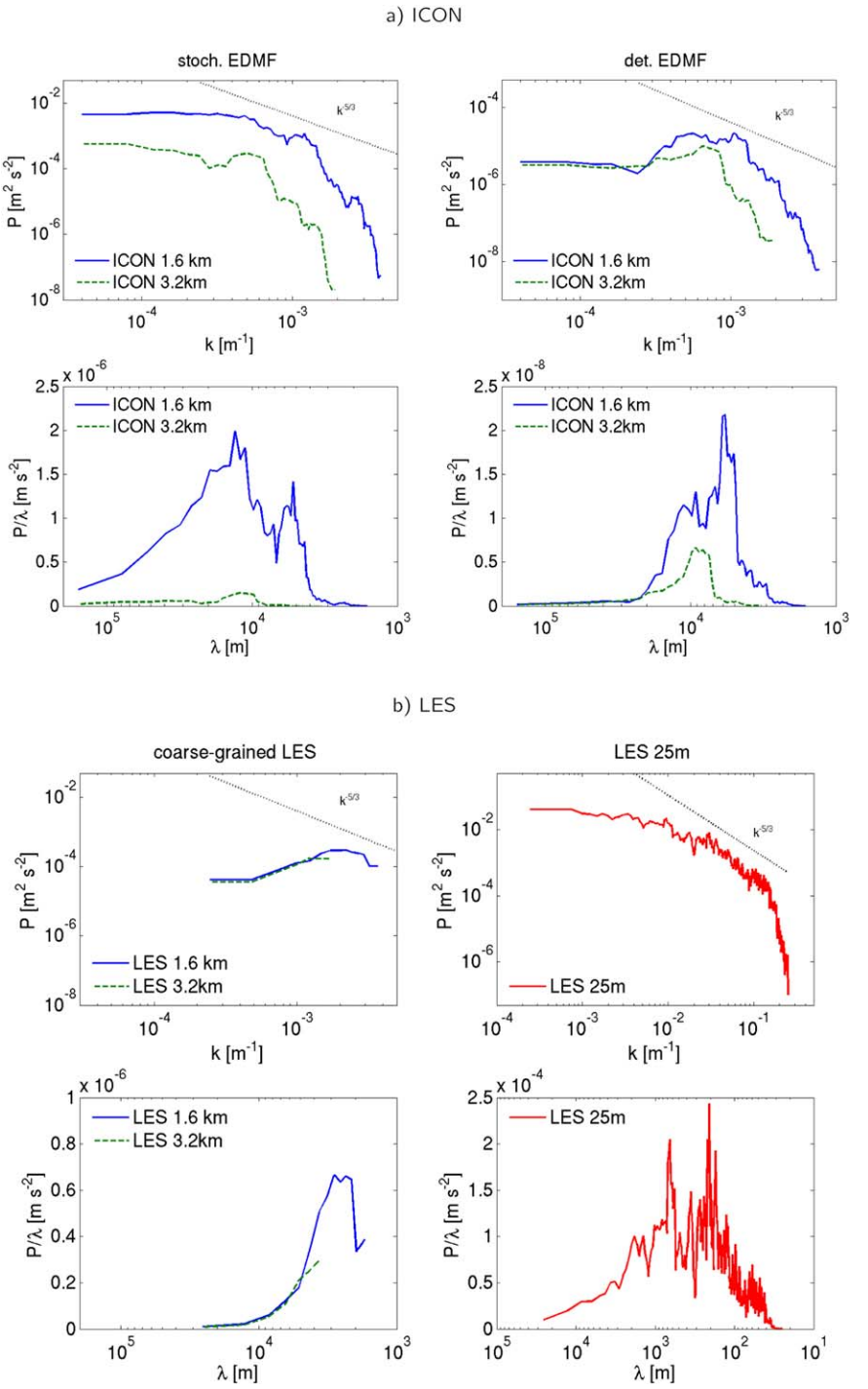

Figure 12. Vertical velocity power spectra along *y* direction from the simulations using (a) (left) the stochastic EDMF in ICON, and (right) the deterministic EDMF in the ICON model, and (b) (left) the coarse-grained LES statistics, and (right) the 25 m resolution LES. Horizontal model resolutions used in ICON and for the LES coarse-graining are 1.6 km (blue) and 3.2 km (green). Top rows represent the turbulent spectra of vertical velocity P as a function of wavenumber k. The bottom rows represent the vertical velocity spectra normalized by the wavelength λ. Inertial-subrange slope of $-5/3$ is plotted as a dashed black line.

right), while the convective energy in the stochastic simulations (Figure 12a, left) matches the amplitude of the peak of the coarse-grained LES case closely. The results here show that the stochastic fluctuations introduced in the cloud parameterization are necessary to reproduce the convective energetics more correctly compared to the deterministic simulations.

In both stochastic and deterministic ICON simulations, the turbulent fluxes are almost completely subgrid, with only a negligible contribution from the resolved scales (not shown here). This result confirms that the distribution of turbulent kinetic energy across spatial scales (turbulent energy cascade) is not resolved on the kilometre-scale grids in the ICON model (Figure 12, top rows). The stochastic parameterization improves the power spectrum of vertical velocity on the kilometre-scale ICON model grids compared to the deterministic simulations. However, the vertical velocity power spectrum does not show the inertial subrange $-5/3$ slope, which is on the other hand clearly visible on the LES vertical velocity spectrum (Figure 12b, right).

5. Summary and Conclusions

In this paper, a stochastic and scale-aware parameterization of the RICO shallow cumulus cloud field has been developed and tested across a range of model grid resolutions in the ICON model. The EDMF scheme is employed in ICON to represent the subgrid effects of turbulence, convection and clouds of the RICO case. The scale-awareness of the cloud parameterization is achieved by coupling the stochastic shallow cumulus ensemble model developed in *Sakradzija et al*. [2015] into the EDMF framework. The developed stochastic parameterization provides a way to apply the statistical QE assumption at the large scales where it is still considered valid, while the subgrid clouds are parameterized on the grid-cell scale by subsampling the stochastic cloud ensemble. By this stochastic approach, the distribution of fluctuations of cloud properties around the equilibrium state is modeled based on the statistical ensemble theory combined with the empirical knowledge from LES. Furthermore, the shallow cloud statistical ensemble is constrained in the mean based on the physical principle that controls the development of the moist convective boundary layer. Scale-dependence of the three ICON model configurations, the deterministic EDMF, the stochastic EDMF in a SCM configuration and the stochastic EDMF with the full model dynamics, have been tested for six different horizontal resolutions in the range from 1 to 50 km. The LES of the RICO case and its coarse-grained statistics have been used as the reference case.

In principle, this stochastic and scale-aware parameterization based on the statistical ensemble theory and LES provides a path to a correct representation of the convective distribution across a range of model horizontal resolutions. However, in practice, the interaction of the stochastic parameterization of shallow cumulus clouds with the nonlinear flow permitted on the model grid becomes another important factor that influences the performance of the scheme and its scale-awareness. This factor is usually not taken into account when parameterizations are developed. Hence, we have also quantified the effects of the interaction between the parameterization and the grid-scale flow and showed that these effects are not negligible in the convective gray zone. In this particular convective case and in combination with EDMF, the interaction of the stochastic scheme with model dynamics further improves the cloud statistics compared to the SCM model setup.

The deterministic EDMF scheme shows some dependence on model resolution because it was not formulated as scale-aware originally. In the simulations tested in this study, the dependence of the domain average quantities on resolution is prominent mainly in the cloud layer as a reduction of cloudiness in the high resolution simulations. The difference across resolutions arises mostly as a result of the diverging convective flows that develop on the model grid and that are in different regimes across different resolutions. This divergence of the grid-scale flow introduces inhomogeneities in the cloud field, which then cause differences in behavior of the parameterization of the cloud layer vertical structure across resolutions.

The stochastic EDMF scheme developed in this study produces the correct average RICO boundary layer structure and its time evolution. With the increase in model resolution, the cloud layer vertical structure and its time evolution are converging toward the LES cloud layer structure. This transition of the stochastic parameterization to a more correct behavior on the higher resolution grids comes from a better use of information that the stochastic cloud ensemble can provide. On the coarse grids, the stochastic parameterization is approaching the performance of the bulk EDMF scheme, while on the higher resolutions the stochastic subsampling of the cloud ensemble results in more variability because of its spectral cloud representation. The result of such behavior is a convergence toward the LES vertical profiles with resolution,

which is a less bottom-heavy and more top-heavy cloud layer. A similar effect can be seen in the time series of cloud properties in the stochastic flow regime, where the cloudiness shows two distinct regimes - a bulk-like regime on the coarse grids and a spectral-like regime on the fine grids. Hence, the stochastic subsampling of the cloud ensemble shows a great potential in improving the average statistics of the RICO case on high-resolution grids. A similar improvement could be achieved also on the coarse model grids by employing a multiupdraft approach [e.g., *Neggers*, 2015] and combining it with the stochastic cloud ensemble model.

The stochastic perturbations interact strongly with the dynamical flow, and as a consequence, a new flow regime develops distinct from the deterministic flow regimes. The stochastic fluctuations dissolve the artificial roll organization of the modeled flow, and instead of regular patterns, they support the development of the spatially irregular circulations. The strength of these circulations and the energy they carry increase with model resolution. While the magnitude of the energy that the circulations carry on the kilometre-scale resolution approaches the magnitude of the LES convective circulations closer than in the deterministic simulations, the spatial scales of these circulations are still too large and are determined by the model resolution scale. Furthermore, the circulations in the stochastic simulations are stronger and the vertical velocity distribution shows a higher variance than the distribution of the effectively resolved flow (coarse-grained LES). In spite of this, the agreement of the stochastic simulations with LES is significantly better compared to the deterministic simulations.

The overly-energetic convective flows modeled at the scales that are not naturally convective scales, affect the parameterization of variability of convection and clouds, and its scale-dependence. In deterministic simulations, convection and clouds are strongly organized into spurious convective rolls and this organization inflates the variance and changes the shape of the distribution of cloud properties. This distribution is dissimilar from the distribution expected from the coarse-grained LES statistics, and can have two peaks associated with the updraft and downdraft regions of convective rolls. In the stochastic simulations, the distribution of cloud fraction is scale-aware and the distribution shape resembles the LES coarse-grained distribution across the range of model horizontal resolutions. When the effects of model dynamics are included into the model configuration, the spurious roll organization is dissolved by the effects of the stochastic scheme, which improves the variability of the cloud field compared to the deterministic and stochastic SCM simulations. So the interaction of the stochastic scheme and its host convective parameterization with model dynamics can have a positive impact on the distribution of cloud properties.

These results also reveal that the limits of validity of the QE assumption are set by one more factor beside the cloud sample size in the model grid column and the memory set by the large-scale forcing. This additional factor is introduced by the nonlinearity of model equations and it manifests through the self-organized convective circulations that are out of equilibrium, and that are, furthermore, under-resolved on the kilometre-scale grids.

To apply this approach to other shallow cumulus cases, further knowledge is required about the processes that control the probability distribution of individual shallow cloud areas $p(a_c)$, mass fluxes $p(m)$, and lifetimes τ. These local properties of convective clouds depend on the convective environment, but might not necessarily be controlled by the large-scale forcing. Thus, the parameterization developed here should not be considered as a universal method for shallow cumulus representation and further development is required. The framework used in this study also has a great potential in unification of shallow and deep convection parameterization. In such a unified parameterization, the distribution of the mass flux of individual deep convective clouds would be represented as a third mode constituting the probability distribution functional form $p(m)$, next to the passive and active shallow cloud distribution modes. However, in order to develop such a parameterization in the EDMF framework, a new closure assumption to estimate the total mass flux M in the cloud ensemble would be necessary.

In this paper we have demonstrated that the stochastic parameterization has a potential to improve the representation of shallow clouds and convection in the convective gray zone to a great extent, in the average properties as well as in the higher moments. The stochastic parameterization interacts with the convective flow that emerges on the model grids, and changes the flow energetics and improves the power spectrum of vertical velocity in the boundary layer. These interactions alter the behavior of parameterization

and show a benefit for the model convergence with resolution toward LES statistics. The parameterization scheme developed and studied in this paper should be seen as an example that sets the basic requirements for a scale-aware representation of shallow clouds, which include stochastic cloud sampling, convective memory and a spatially nonlocal approach to parameterization.

Appendix A: Scale-Dependence of the EDMF Buoyancy-Sorting Closure Assumption

The closure for the mass flux vertical structure in EDMF is an implicit version of a buoyancy sorting scheme [see e.g., *Kain and Fritsch*, 1990]. In such schemes, the turbulent mixing between clouds and their environment is represented by the mixtures of cloudy and environmental air that can entrain or detrain the air mass into and from the cloud near its periphery. In EDMF, the dependence of the mass flux profiles on the updraft properties and on the environmental conditions is introduced through the moist zero buoyancy deficit $q_t^x - \overline{q_t}$, where the critical total water mixing ratio q_t^x corresponds to the critical fraction of environmental air at which the mixture is neutrally buoyant. In terms of this formulation, if the mean total water mixing ratio of a model grid column $\overline{q_t}$ is far from the critical total water mixing ratio q_t^x, only a small fraction of the cumulus updraft will be positively buoyant and able to form clouds. And the other way around, if the distance between $\overline{q_t}$ and q_t^x is small, a large fraction of the updraft will become cloudy.

The normalized moist zero buoyancy deficit Q_c is defined as:

$$Q_c = \frac{q_t^x - \overline{q_t}}{\sigma_{q_t}} \tag{A1}$$

where σ_{q_t} is the standard deviation of the total water mixing ratio. The normalized moist zero buoyancy deficit Q_c is correlated to the cloud fraction a_c, which is expressed as a correlation function of the vertical gradients of a_c and Q_c [*Neggers et al*., 2009]:

$$\frac{1}{a_c}\frac{\partial a_c}{\partial z'} = C_a \frac{1}{Q_c}\frac{\partial Q_c}{\partial z'} \tag{A2}$$

The correlation coefficient C_a is found to be equal to -1.8 based on the LES studies of several different cases, over the ocean and over the land [*Neggers et al*., 2009]. z' is the height above the cloud base normalized by the cloud layer depth.

As an example of the scale-dependence of the cloud layer vertical structure parameterization, we test the parameterization of the vertical mass flux profile in terms of the gradient scaling relation (equation (A2)) in a coarse-graining study of the RICO-140 LES case. The original resolution of the LES simulation is 25 m, thus to represent the range of grid resolutions of a NWP model, the LES field is divided into areas of different sizes and cloud properties are averaged over these areas. The gradients of the normalized moist zero buoyancy deficit Q_c and cloud area fraction a_c are estimated from LES, and are plotted on the scatter plots for the different coarse-graining resolutions (Figure A1). The scaling relation equation (A2) is plotted in its original form without fitting it to the LES scatter points (red line on Figure A1). This scaling is originally introduced for the coarse model grids and it fits very well for the resolution coarser than 20 km. The cumulus ensemble is under-sampled on the gray-zone scales, and the spread around the employed scaling relation (equation (A2)) becomes wider with the increase of resolution. There is no unique vertical gradient of the cloud area fraction a_c for each gradient of the normalized moist zero buoyancy deficit Q_c.

Based on these results, we conclude that on the model grids with a horizontal resolution coarser than 20 km the scaling relation equation (A2) is valid because the estimated gradients corresponding to the RICO-140 case fall close to the parameterization line. On the grids with a horizontal resolution finer than 20 km, the scatter of points that represent the pairs of vertical gradients of Q_c and a_c is wide and the scaling relation is no longer valid. Thus, the closure scaling equation (A2) is impaired on the high resolution grids because of the uncertainty in the moist zero buoyancy deficit Q_c on high resolutions, and of course, also because of a high variability in the subgrid cloud fraction a_c on the high resolution grids.

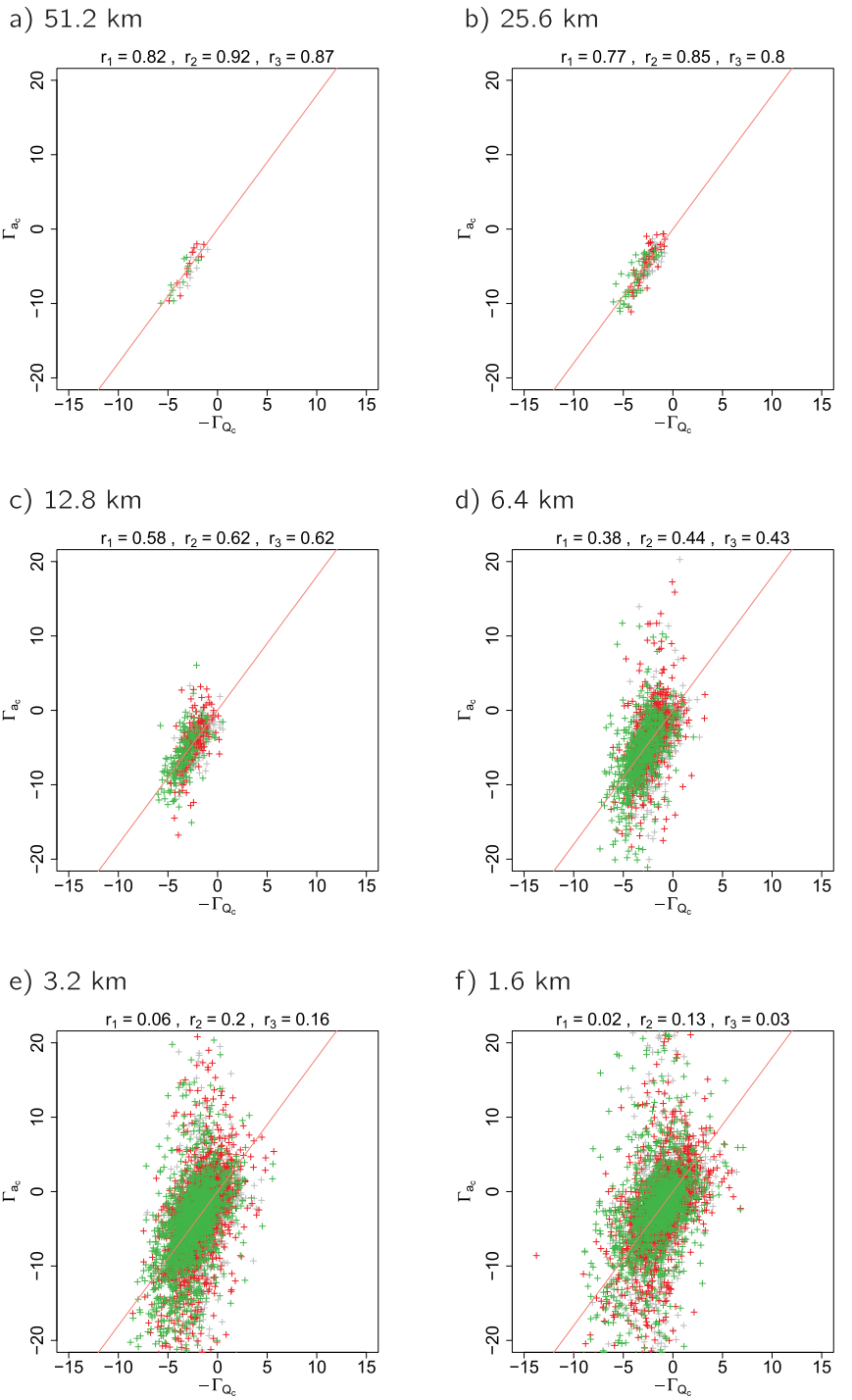

Figure A1. Resolution dependence of the original EDMF closure assumption given by equation (A2) (red line) tested for the RICO case. Plots are showing the pairs of the vertical gradients of the moist updraft area fraction and zero-buoyancy deficit of q_t above the cloud base, within the time frames starting at 6 (gray), 12 (red) and 18 h (green). Results are coarse-grained across the range of horizontal model resolutions from (top-left) 51.2 km to (bottom-right) 1.6 km. Correlation coefficients r_1, r_2, and r_3, are calculated for all three time frames (above the plots).

Acknowledgments
For constructive and fruitful discussions about convective parameterizations we thank Pier Siebesma, Robert Plant, and George Craig, and especially Martin Köhler and Roel Neggers. The authors are also thankful for the technical support provided by Leonidas Linardakis and Daniel Reinert among other ICON model developers. The manuscript is greatly improved owing to suggestions and comments provided by Cathy Hohenegger, Alberto De Lozar, Gualtiero Badin, and two anonymous reviewers to whom we also express our thanks. The work described here was initially a part of the PhD research of the corresponding author [*Sakradzija*, 2015], and it is presented here in a changed and further developed form and content. Primary data and scripts used in the analysis and other supplementary information that may be useful in reproducing the author's work are archived by the Max Planck Institute for Meteorology and can be obtained by contacting publications@mpimet.mpg.de. This research was carried out as part of the Hans-Ertel Centre for Weather Research. This research network of Universities, Research Institutes and the Deutscher Wetterdienst is funded by the BMVI (Federal Ministry of Transport and Digital Infrastructure).

References

Arakawa, A., and W. H. Schubert (1974), Interaction of a cumulus cloud ensemble with the large-scale environment, Part I, *J. Atmos. Sci.*, *31*, 674–701.

Arakawa, A., J.-H. Jung, and C.-M. Wu (2011), Toward unification of the multiscale modeling of the atmosphere, *Atmos. Chem. Phys.*, *11*, 3731–3742.

Bengtsson, L., H. Körnich, E. Källén, and G. Svensson (2011), Large-scale dynamical response to subgrid-scale organization provided by cellular automata, *J. Atmos. Sci.*, *68*, 3132–3144.

Berner, J., T. Jung, and T. N. Palmer (2012), Systematic model error: The impact of increased horizontal resolution versus improved stochastic and deterministic parameterizations, *J. Clim.*, *25*, 4946–4962.

Berner, J., K. R. Fossell, S.-Y. Ha, J. P. Hacker, and C. Snyder (2015), Increasing the skill of probabilistic forecasts: Understanding performance improvements from model-error representations, *Mon. Weather Rev.*, *143*, 1295–1320.

Betts, A. K. (1976), Modeling subcloud layer structure and interaction with a shallow cumulus layer, *J. Atmos. Sci.*, *33*, 2363–2382.

Bretherton, C. S., J. R. McCaa, and H. Grenier (2004), A new parameterization for shallow cumulus convection and its application to marine subtropical cloud-topped boundary layers. Part I: Description and 1D results, *Mon. Weather Rev.*, *132*, 864–882.

Bright, D. R., and S. L. Mullen (2002), Short-range ensemble forecasts of precipitation during the southwest monsoon, *Weather Forecasting*, *17*, 1080–1100.

Brown, R. A. (1980), Longitudinal instabilities and secondary flows in the planetary boundary layer: A review, *Rev. Geophys.*, *18*(3), 683–697, doi:10.1029/RG018i003p00683.

Bryan, G. H., and R. Rotunno (2005), Statistical convergence in simulated moist absolutely unstable layers, in *11th Conference on Mesoscale Processes*, Am. Meteorol. Soc., Albuquerque, N. M., 1M.6, 5 pp. [Available at http://n2t.net/ark:/85065/d7cn732w.]

Bryan, G. H., J. C. Wyngaard, and J. M. Fritsch (2003), Resolution requirements for the simulation of deep moist convection, *Mon. Weather Rev.*, *131*, 2394–2416.

Ching, J., R. Rotunno, M. LeMone, A. Martilli, B. Kosovic, P. A. Jimenez, and J. Dudhia (2014), Convectively induced secondary circulations in fine-grid mesoscale numerical weather prediction models, *Mon. Weather Rev.*, *142*, 3284–3302.

Dipankar, A., B. Stevens, R. Heinze, C. Moseley, G. Zängl, M. Giorgetta, and S. Brdar (2015), Large eddy simulation using the general circulation model ICON, *J. Adv. Model. Earth Syst.*, *7*, 963–986, *doi*:10.1002/2015MS000431.

Doms, G., et al. (2011), *A Description of the Nonhydrostatic Regional COSMO Model, Part II: Physical Parameterization*, Deutscher Wetterdienst, Offenbach, Germany. [Available at http://www.cosmo-model.org, last accessed 9 May 2014.]

Dorrestijn, J., D. T. Crommelin, A. P. Siebesma, and H. J. J. Jonker (2013), Stochastic parameterization of shallow cumulus convection estimated from high-resolution model data, *Theor. Comput. Fluid Dyn.*, *27*, 133–148.

Drazin, P. G. (2002), *Introduction to Hydrodynamic Stability*, Cambridge Univ. Press, University of Cambridge, Cambridge, U. K.

Frenkel, Y., A. J. Majda, and B. Khouider (2012), Using the stochastic multicloud model to improve tropical convective parameterization: A paradigm example, *J. Atmos. Sci.*, *69*, 1080–1105.

Giorgi, F., and M. R. Marinucci (1996), A investigation of the sensitivity of simulated precipitation to model resolution and its implications for climate studies, *Mon. Weather Rev.*, *124*, 148–166.

Grell, G. A., and S. R. Freitas (2014), A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling, *Atmos. Chem. Phys.*, *14*, 5233–5250.

Holtslag, A. A. M., and B. A. Boville (1993), Local versus nonlocal boundary-layer diffusion in a global climate model, *J. Clim.*, *6*, 1825–1842.

Honnert, R., V. Masson, and F. Couvreux (2011), A diagnostic for evaluating the representation of turbulence in atmospheric models at the kilometric scale, *J. Atmos. Sci.*, *68*(12), 3112–3131, doi:10.1175/JAS-D-11-061.1.

Jung, J.-H., and A. Arakawa (2004), The resolution dependence of model physics: illustrations from nonhydrostatic model experiments, *J. Atmos. Sci.*, *61*, 88–102.

Kain, J. S., and J. M. Fritsch (1990), A one-dimensional entraining/detraining plume model and its application in convective parameterization, *J. Atmos. Sci.*, *47*, 2784–2802.

Khouider, B., J. Biello, and A. Majda (2010), A stochastic multicloud model for tropical convection, *Commun. Math. Sci.*, *8*, 187–216.

Kiehl, J. T., and D. L. Williamson (1991), Dependence of cloud amount on horizontal resolution in the national center for atmospheric research community climate model, *J. Geophys. Res.*, *96*(D6), 10,955–10,980.

Lean, H. W., P. A. Clark, M. Dixon, N. M. Roberts, A. Fitch, R. Forbes, and C. Halliwell (2008), Characteristics of high-resolution versions of the Met Office unified model for forecasting convection over the United Kingdom, *Mon. Weather Rev.*, *136*, 3408–3424.

LeMone, M. A. (1973), The structure and dynamics of horizontal roll vortices in the planetary boundary layer, *J. Atmos. Sci.*, *30*, 1077–1091.

LeMone, M. A., and W. T. Pennell (1976), The relationship of trade wind cumulus distribution to subcloud layer fluxes and structure, *Mon. Weather Rev.*, *104*, 524–539.

LeMone, M. A., F. Chen, M. Tewari, J. Dudhia, B. Geerts, Q. Miao, R. L. Coulter, and R. L. Grossman (2010), Simulating the IHOP_2002 fair-weather CBL with the WRF-ARW-NOAH modeling system. Part II: Structures from a few kilometers to 100 km across, *Mon. Weather Rev.*, *138*, 745–764.

Lewellen, W. S., and S. Yoh (1993), Binormal model of ensemble partial cloudiness, *J. Atmos. Sci.*, *50*(9), 1228–1237.

Lin, J. W.-B., and J. D. Neelin (2000), Influence of a stochastic moist convective parameterization on tropical climate variability, *Geophys. Res. Lett.*, *27*(22), 3691–3694.

Lin, J. W.-B., and J. D. Neelin (2002), Considerations for stochastic convective parameterization, *J. Atmos. Sci.*, *59*, 959–975.

Neggers, R. A. J. (2009), A dual mass flux framework for boundary layer convection. Part II: Clouds, *J. Atmos. Sci.*, *66*, 1489–1506.

Neggers, R. A. J. (2015), Exploring bin–macrophysics models for moist convective transport and clouds, *J. Adv. Model. Earth Syst.*, *7*, 2079–2104, doi:10.1002/2015MS000502.

Neggers, R. A. J., A. P. Siebesma, and H. J. J. Jonker (2002), A multiparcel model for shallow cumulus convection, *J. Atmos. Sci.*, *59*, 1655–1668.

Neggers, R. A. J., B. Stevens, and J. D. Neelin (2006), A simple equilibrium model for shallow-cumulus-topped mixed layers, *Theor. Comput. Fluid Dyn.*, *20*(5–6), 305–322, doi:10.1007/s00162-006-0030-1.

Neggers, R. A. J., B. Stevens, and J. D. Neelin (2007), Variance scaling in shallow-cumulus-topped mixed layers, *Q. J. R. Meteorol. Soc.*, *133*, 1629–1641, doi:10.1002/qj.105.

Neggers, R. A. J., M. Köhler, and A. C. M. Beljaars (2009), A dual mass flux framework for boundary layer convection. Part I: Transport, *J. Atmos. Sci.*, *66*, 1465–1487.

Palmer, T. N. (2001), A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models, *Q. J. R. Meteorol. Soc.*, *127*, 279–304, doi:10.1002/qj.49712757202.

Piotrowski, Z. P., P. K. Smolarkiewicz, S. P. Malinowski, and A. A. Wyszogrodzki (2009), On numerical realizability of thermal convection, *J. Comput. Phys.*, *228*, 6268–6290.

Plant, R. S., and G. C. Craig (2008), A stochastic parameterization for deep convection based on equilibrium statistics, *J. Atmos. Sci.*, *65*, 87–105.

Pope, V. D., and R. A. Stratton (2002), The processes governing horizontal resolution sensitivity in a climate model, *Clim. Dyn.*, *19*, 211–236, doi:10.1007/s00382-001-0222-8.

Rauber, R. M., et al. (2007), Rain in shallow cumulus over the ocean: the RICO campaign, *Bull. Am. Meteorol. Soc.*, *88*, 1912–1928, doi: 10.1175/BAMS-88-12-1912.

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

Sakradzija, M. (2015), A stochastic parameterization of shallow cumulus convection for high-resolution numerical weather prediction and climate models, PhD thesis, Reports on Earth System Science, Int. Max Planck Res. Sch. on Earth Syst. Modell., Max Planck Inst. for Meteorol., Hamburg, Germany, ISSN 1614–1199.

Sakradzija, M., A. Seifert, and T. Heus (2015), Fluctuations in a quasi-stationary shallow cumulus cloud ensemble, *Nonlinear Proc. Geophys.*, *22*, 65–85, doi:10.5194/npg-22-65-2015.

Seifert, A., and T. Heus (2013), Large-eddy simulation of organized precipitating trade wind cumulus clouds, *Atmos. Chem. Phys.*, *13*, 5631–5645, doi:10.5194/acp-13-5631-2013.

Seifert, A., T. Heus, R. Pincus, and B. Stevens (2015), Large-eddy simulation of the transient and near-equilibrium behavior of precipitating shallow convection, *J. Adv. Model. Earth Syst.*, *7*, 1918–1937, doi:10.1002/2015MS000489.

Selz, T., and G. C. Craig (2015), Simulation of upscale error growth with a stochastic convection scheme, *Geophys. Res. Lett.*, *42*, 3056–3062, doi:10.1002/2015GL063525.

Siebesma, A. P., P. M. M. Soares, and J. Teixeira (2007), A combined eddy-diffusivity mass-flux approach for the convective boundary layer, *J. Atmos. Sci.*, *64*, 1230–1248.

Skamarock, W. C., J. B. Klemp, M. G. Duda, L. D. Fowler, S.-H. Park, and T. D. Ringler (2012), A multiscale nonhydrostatic atmospheric model using centroidal Voronoi tesselations and C-grid staggering, *Mon. Weather Rev.*, *140*, 3090–3105.

Stevens, B., et al. (2013), Atmospheric component of the MPI-M earth system model: ECHAM6, *J. Adv. Model. Earth Syst.*, *5*, 146–172, doi: 10.1002/jame.20015.

Stirling, A. J., and J. C. Petch (2004), The impacts of spatial variability on the development of convection, *Q. J. R. Meteorol. Soc.*, *130*, 3189–3206.

Sušelj, K., J. Teixeira, and D. Chung (2013), A unified model for moist convective boundary layers based on a stochastic eddy-diffusivity/mass-flux parameterization, *J. Atmos. Sci.*, *70*, 1929–1953.

Sušelj, K., T. F. Hogan, and J. Teixeira (2014), Implementation of a stochastic eddy-diffusivity/mass-flux parameterization into the navy global environmental model, *Weather Forecasting*, *29*, 1374–1390.

Teixeira, J., and C. A. Reynolds (2008), Stochastic nature of physical parameterizations in ensemble prediction: A stochastic convection approach, *Mon. Weather Rev.*, *136*, 483–496.

Tomita, H., and M. Satoh (2004), A new dynamical framework of nonhydrostatic global model using the icosahedral grid, *Fluid Dyn. Res.*, *34*, 357–400.

Troen, I. B., and L. Mahrt (1986), A simple model of the atmospheric boundary layer; sensitivity to surface evaporation, *Boundary Layer Meteorol.*, *37*, 129–148, doi:10.1007/BF00122760.

van Stratum, B. J. H., J. Vilá–Guerau de Arellano, C. C. van Heerwaarden, and H. G. Ouwersloot (2014), Subcloud-layer feedbacks driven by the mass flux of shallow cumulus convection over land, *J. Atmos. Sci.*, *71*, 881–895.

van Zanten, M. C., et al. (2011), Controls on precipitation and cloudiness in simulations of trade-wind cumulus as observed during RICO, *J. Adv. Model. Earth Syst.*, *3*, M06001, doi:10.1029/2011MS000056.

Williamson, D. L. (1999), Convergence of atmospheric simulations with increasing horizontal resolution and fixed forcing scales, *Tellus, Ser. A*, *51*, 663–673.

Zängl, G., D. Reinert, P. Rípodas, and M. Baldauf (2015), The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core, *Q. J. R. Meteorol. Soc.*, *141*, 563–579, doi:10.1002/qj.2378.

Zhou, B., J. S. Simon, and F. K. Chow (2014), The convective boundary layer in the terra incognita, *J. Atmos. Sci.*, *71*, 2545–2563.