Stochastic Differential Equations for the Variability of Atmospheric Convection Fluctuating Around the Equilibrium

E. Machulskaya1 and A. Seifert1

1Deutscher Wetterdienst, Offenbach-am-Main, Germany

Abstract Most convection parameterization schemes used within current atmospheric models make a convective equilibrium assumption, which breaks at the resolutions currently used by many numerical weather prediction models. To account for fluctuations of the cloud base mass flux about its equilibrium value, stochastic convection schemes have been developed for both deep (Plant & Craig, 2008, https://doi.org/10.1175/2007JAS2263.1) and shallow (Sakradzija et al., 2015, https://doi.org/10.5194/npg-22-65-2015) convection. Due to the need to explicitly track individual clouds in each grid box, these schemes can be computationally expensive. Motivated by the above considerations, the present study demonstrates how the machinery of the above schemes can be reduced to the solution of two ordinary stochastic differential equations for the cloud number and the total cloud base mass flux. The properties of the resulting stochastic processes for the cloud number and the total mass flux are not exactly equivalent to those of the original schemes but recover them to a very good approximation.

1. Introduction

Most convection parameterization schemes used within current atmospheric models utilize the convective equilibrium assumption which states that the convective model response always matches the convective forcing. Hence, this assumption allows to uniquely relate the grid box tendencies of momentum, temperature, and moisture due to convection to the local and instantaneous values of the large-scale model fields. However, the peculiarity of the convective response is that it is not continuous in space, time, and intensity. It is materialized by individual localized entities—convective plumes and their visible manifestation—convective clouds. Hence, this response is discrete in intensity, may be delayed in time as clouds need time to appear (and to disappear), and is not homogeneous in space at small scales due to randomly placed clouds even for a homogeneous large-scale forcing. Therefore, the convective equilibrium assumption can only be valid if a sufficiently large area is considered, where the number of clouds in a grid box is large enough to represent an ensemble of clouds to assure an equilibrium response. This requirement may be taken as fulfilled at a relatively coarse spatial model resolution where the mesh size has an order of 50 km and more. At the present time, however, many numerical weather prediction (NWP) models operate with essentially higher mesh sizes of order of 10 km, where the equilibrium assumption is no longer valid.

Instead, the observed convective response at this scale starts to be affected by the finiteness of the cloud number (Plant & Craig, 2008). This means that the total mass flux transported by a finite number of clouds in the grid box of this scale may fluctuate around its equilibrium value, approaching this value only on average over some temporal interval and/or spatial domain, if the large-scale conditions in the corresponding temporal and spatial windows may be taken as stationary.

This problem may be regarded as a part of a more general problem of the lost variability which emerges in truncated models, where the interaction of model variables beyond some truncation scale is not resolved explicitly (see, e.g., Hasselmann, 1976). In this case, one and the same macrostate of resolved variables may be consistent with a variety of different microstates of unresolved variables, which, however, may differently affect the tendencies of resolved variables. This natural variability in resolved tendencies is then lost if only one tendency per macrostate is chosen for a resolved variable which is the case in conventional deterministic NWP. The problem of the accounting for the variety of other possible resolved tendencies, commonly referred to as model error, has been receiving much attention in the last two decades. The model error description, that is, the strategy of generating model perturbations, has become one of the crucial points in...
the design of the ensemble prediction systems (EPS), which have already become a necessary component of the weather prediction in the last decade. Due to the highly random nature of the convection phenomenon at small scales, the probabilistic description of the convection is one of the main candidates among the descriptions of the uncertainties in other physical processes to be included into an ensemble prediction system (ECMWF, 2011, 2016). Generally speaking, the phenomenon of convection may be regarded as random in different aspects; the aspect of the variability of the convective response due to its quantization and random plumes displacement discernible at small scales, outlined in the previous paragraph, is one of them, and this is the aspect the present study deals with.

To account for the fluctuations of the total mass flux in a grid box at the cloud base for high model resolutions, two stochastic schemes, to the best authors’ knowledge, were developed, one for deep convection (Craig & Cohen, 2006; Cohen & Craig, 2006; hereafter CC06), (Plant & Craig, 2008, hereafter PC08) and one for shallow convection (Sakradzija et al., 2015, 2016, hereafter SSH15). Specific details of the schemes are provided in the next section; here the general idea behind them is given that is needed to clarify the motivation of the present study. The main idea of both schemes is to regard the equilibrium mass flux at the cloud base computed by a deterministic convection scheme as a mean estimate to which the actual convective response tends only on average. The actual mass flux at the cloud base in each grid box is determined as the sum of the mass fluxes of single clouds which are randomly generated in each grid box with a rate consistent with the large-scale homogeneous forcing. To each generated cloud, a random mass flux is ascribed which is sampled from a probability distribution function (PDF), different for different schemes. Here an important observation from the cloud-resolving and large-eddy simulations is used that the mean mass flux per cloud does not depend on the strength of the convective forcing but that it is the number of clouds that responds to the convective forcing magnitude (Cohen & Craig, 2006, SSH15). This suggests using the constant in time PDF of the mass fluxes of generated clouds for changing environmental conditions, where only the number of generated clouds is adjusted to the strength of the convective forcing. To each generated cloud its lifetime is ascribed (see the next section for the details of its choice). After a cloud is generated, it is put into the list of living clouds which is maintained for each grid box. At each model step it is checked whether the cloud has to disappear according to its lifetime, after which the cloud is taken out from the list. The properties of clouds do not change during its life cycle in the PC08 scheme and in the basic version of the SSH15 scheme. The notion of the lifetime is used there solely to reasonably represent the temporal correlation of cloud base mass flux rather than to directly simulate the life cycle behavior of each cloud.

These stochastic convection schemes represent a natural way to introduce the unresolved variability outlined above by random finite sampling from the cumulus cloud ensemble. The schemes have been shown to reproduce the correct variability of precipitation for deep and shallow convection across a range of scales (Keane & Plant, 2012, SSH15) and are the candidates for using in the operational NWP practice. The problem that can hinder the operational use of these schemes, however, is that the described machinery related to the tracking of each single cloud may be quite expensive as to the required computational time. The question addressed in the present study concerns the possibility to represent the schemes developed in PC08 and SSH15 in effect by means of such a system of ordinary stochastic differential equations (SDE) for the cloud number $N$ and the total mass flux $M$ that would be capable to recover the statistical behavior of $N$ and $M$ of the corresponding full schemes. The solution of those ordinary equations would be much more computationally inexpensive and could be used in atmospheric models in place of the schemes of PC08 and SSH15. In what follows, we will refer to such equation systems as the coarse-grained solutions, as opposed to the respective fine-grained full schemes. It may be pointed out at once that no solution is found which would reproduce both schemes exactly (it is shown why for the PC08 scheme it seems to be even impossible); however, the proposed solutions recover the original schemes to a good approximation which is well within the uncertainty of the formulations of the original schemes.

It should be noted that the original implementation of the PC08 scheme is tightly coupled to the host deterministic Kain-Fritsch convection parameterization scheme (Kain & Fritsch, 1990; Kain, 2004). Within the framework of this scheme it is possible to use the information on the cloud base mass flux of individual clouds to compute the individual cloud properties such as the entrainment/detrainment intensity and cloud top height. In this way, not only the total cloud base mass flux is perturbed as compared to the mass flux from the deterministic convection scheme, which alone would mean the respective rescaling of the vertical profiles of the temperature and moisture tendencies, but the vertical structure of the tendencies is also affected. This issue remains, however, beyond the scope of the present paper where only the question of the
fluctuating cloud base mass flux is addressed. Such approach corresponds with the idea of the detachment of the stochastic treatment of the cloud base mass flux from the Kain-Fritsch parameterization and its application to any other bulk convection parameterization scheme operating with the notion of the cloud base mass flux. Along these lines, the stochastic scheme for shallow convection of SSH15 was developed from the very beginning for the coupling to a bulk deterministic convection parameterization. In what follows, as the PC08 scheme, we will refer to the stochastic component dealing with the cloud base mass flux only.

The paper is organized as follows. In the next section, the stochastic schemes developed in PC08 and SSH15 are briefly described. After that, the systems of SDEs emulating the behavior of the stochastic schemes for deep and shallow convection are proposed. Although the methodology of the derivation of those equations is the same in both cases, the solutions for deep and shallow convection are placed in different sections to facilitate the comprehension. We start in section 3 with the simpler case of deep convection; this case is easier to treat not because of the nature of deep convection per se but because the distribution of the cloud mass fluxes and the cloud lifetime in the stochastic scheme for deep convection in PC08 are formulated in a mathematically more simple way, which allows to present the steps of the methodology more transparently. In section 4, with a didactic aim a system of an intermediate complexity is explored. Finally, in section 5, the system of SDEs for shallow convection is presented. Each of those latter three sections is completed with the respective results of the comparison of the coarse- and fine-grained schemes made by means of numerical experiments. The conclusions are given in section 6.

For the sake of comprehensibility, the list of the most important notations used in the paper is provided here.

- \( t \) time (s),
- \( \tau \) cloud lifetime (s),
- \( m \) mass flux at cloud base of a single cloud (kg/s),
- \( m_b \) mean cloud mass flux of generated clouds (kg/s),
- \( N \) total number of living clouds in a grid box (dimensionless),
- \( M \) total mass flux at cloud base of living clouds in a grid box (kg/s),
- \( \lambda \) cloud generation rate \( (s^{-1}) \),
- \( p_b (m) \) probability density function of mass flux of generated clouds \( (s/kg) \),
- \( p (m) \) probability density function of mass flux of living clouds \( (s/kg) \),
- \( p_d (m) \) probability density function of mass flux of disappearing clouds \( (s/kg) \),
- \( n(m) \) number density of living clouds with a particular mass flux \( (s/kg) \),
- \( \langle \cdot \rangle \) ensemble mean of a random variable,
- \( \bar{\cdot} \) mean of a single-cloud characteristics (mass flux or lifetime) over all clouds in a grid box,
- \( \alpha \) a prefactor in the relation of the cloud lifetime to its mass flux \( (s^{\alpha+1}/kg^{-\beta}) \),
- \( \beta \) an exponent in the relation of the cloud lifetime to its mass flux (dimensionless),
- \( k \) shape parameter of a Weibull distribution

### 2. Brief Description of the Stochastic Convection Schemes

As was described in section 1, both stochastic convection schemes addressed in the present study share in substance the same basic principles, although the differences between them are also remarkable. In this section, the mathematical properties of the schemes are described that will be needed in the rest of the paper. For the comprehensive description of the schemes, the reader is referred to CC06, PC08, and SSH15.

The assumptions behind the PC08 scheme consist of

- the independence of the mass fluxes of individual clouds of each other,
- the invariance of the mean mass flux per cloud \( m_b \) to the strength of forcing,
- an exponential distribution of mass fluxes of living clouds

\[
p(m) = \frac{1}{m_b} e^{- \frac{m}{m_b}},
\]

- a constant lifetime of clouds \( \tau_0 \), set to \( \tau_0 = 45 \) min in the original formulation of PC08, and
- a Poisson process for the cloud generation in time, with the parameter \( \lambda \), the cloud generation rate determined by the large-scale convective forcing. The reader is referred to appendix to see why these are the assumptions independently adopted in the PC08 scheme.
Given an exponential distribution of mass fluxes of living clouds and the constant cloud lifetime, it follows that the distribution of newborn cloud mass fluxes $p_b(m)$ should also be exponential

$$p_b(m) = \frac{1}{m_b} e^{-\frac{m}{m_b}},$$

with the same parameter $m_b$ as in the distribution of living clouds. This parameter has the meaning of the mean mass flux per newborn (or living) cloud in a grid box.

In contrast to the PC08 scheme, the scheme of SSH15 explicitly considers the process of the cloud generation. It distinguishes between the distributions of the mass flux of generated and living clouds, where only the former is specified and the latter appears in a natural way according to the former. The analysis of the LES data obtained for the Rain in Cumulus over the Ocean (RICO) case, a precipitating shallow convective case in the trade wind region, suggests that the distribution of generated clouds may be fitted by a mixed Weibull distribution with two modes of “active” and “passive” clouds

$$p_b(m) = f \frac{k}{m_{b1}} \left( \frac{m}{m_{b1}} \right)^{k-1} e^{-\left(\frac{m}{m_{b1}}\right)^{k}} + (1 - f) \frac{k}{m_{b2}} \left( \frac{m}{m_{b2}} \right)^{k-1} e^{-\left(\frac{m}{m_{b2}}\right)^{k}},$$

where $f$ is the fraction of clouds belonging to one of the two cloud modes, $k$ is the shape parameter, and $m_{b1}$ and $m_{b2}$ are the respective scale parameters of the two modes (denoted by $\theta_1$ and $\theta_2$ in SSH15). Using the nonlinear least square fitting for the same LES data, the cloud lifetime $\tau$ is found to be a power law function of cloud mass flux $m$ as

$$\tau = a m^\beta,$$

where $\alpha$ and $\beta$ are the empirically determined parameters. It is also assumed that clouds appear according to a Poisson process with the parameter $\lambda$ characterizing large-scale conditions.

As was mentioned before, the cloud generation rate $\lambda$ is determined by the large-scale forcing. Both schemes recognize that atmospheric conditions in a single grid element may not be representative for a smoothly changing large-scale environment. Instead, to characterize the large-scale forcing, a spatial scale is introduced, which is, on the one hand, larger than a single grid box but which, on the other hand, is sufficiently small for the convective forcing to be regarded as homogeneous in space. This means that the convective forcing in each grid box itself is regarded as a fluctuation around a homogeneous equilibrium value which is used to determine $\lambda$. This homogeneous equilibrium forcing is computed then as running means of the model fields over some adjacent grid boxes. In PC08, the averaging scale of $\sim 320$ km is proposed. In the implementation of the PC08 scheme into the Met Office Regional Ensemble Prediction System (MOGREPS), the area with the side of 168 km over ocean and 72 km over land is used (Keane et al., 2016). In both PC08 and SSH15 schemes implemented into the global ICON NWP model used operationally at Deutscher Wetterdienst, the averaging is performed over the neighbors of each grid box and the neighbors of the neighbors, which results in the averaging area with the side of approximately 65 km (Keane et al., 2014; Sakradzija et al., 2016). In all these examples, the temporal averaging over the time interval of 10 min to 1 hr is also used.

3. The Deep Convection Case: The Scheme of PC08

As was mentioned in section 1, it seems to be impossible to precisely represent the PC08 scheme in terms of two ordinary SDEs for $N$ and $M$, and this is for the following reasons. In the time series of the cloud number $N$ generated by the PC08 scheme, the increments $\Delta N$ over certain distinct time intervals $\Delta t$ are not independent random variables. For each time interval $\Delta t$, the decrement of the number of living clouds $N$ over this interval is equal to the number of clouds born between $\tau_0$ and $\tau_0 + \Delta t$ time units ago, that is, to one of the past increments, which creates a correlation between the increments of $N$ over $\Delta t$ and $\Delta t - \tau_0$. On the other hand, the solution of a first-order (with respect to time) SDE should be fully determined by the knowledge of the present state and be independent of the exact details of the evolution in the past that has lead to this present state (Markov property) and, in particular, be independent of past increments (an independent increment process). Therefore, the trajectories of $N$ in the PC08 scheme, which do not constitute an independent increment process, cannot be modeled by means of an SDE. Similar considerations apply, of course, to $M$ as well. It may also be noted that the sets of Markov processes and processes with independent increments coincide up to some rather exotic examples, so that in what follows, it will be distinguished between Markovian (as an SDE system) and non-Markovian (as the original PC08 scheme) processes.
Despite the impossibility to precisely replace a non-Markovian process through an SDE, by means of an SDE one can obtain a Markov process which is quite close to that produced by the PC08 scheme. These stochastic equations are the exact solution of another system that may be regarded as a modification of the PC08 scheme. In the modified PC08 system, which will be termed Markovian PC08 scheme in what follows, cloud lifetime is not a constant $\tau_0$ but is distributed exponentially with the mean $\langle \tau \rangle = \tau_0$. It may seem that such modification is quite crude from the physical point of view, in particular if one realizes that in this Markovian system the lifetime of a cloud and its mass flux at the cloud base are mutually independent, so that a large cloud may have a short lifetime and vice versa. But since, as will be shown later, the properties of the resulting random processes of $N$ and $M$ are in fact close to those of the original PC08 system and since $M$ is the only variable required as an output from the PC08 scheme, the proposed stochastic equations may indeed be used as a reasonable approximation for the PC08 scheme. Furthermore, later in section 4 another kind of possible modification will be pointed out which also admits an exact solution with respect to $N$.

The modeling of the cloud number $N$ in the system with the exponentially distributed lifetime represents a special case of a well-known problem which appears in many technical, engineering, and biological applications (see, e.g., Daley & Vere-Jones, 2003; Taylor & Karlin, 1998, for numerous examples). This classical problem concerns various statistical properties of a population of some identical objects, which appear and disappear (die, go out of service, etc.) according to some prescribed stochastic laws; the respective stochastic process is referred to as a birth-death process. In the integral and most general form sample paths of such process have the form of the difference of two independent random processes describing the birth and the death of objects

$$N(t) = B(t) - D(t),$$

where $B(t)$ and $D(t)$ are the cumulative number of appeared and disappeared objects up to the time $t$, respectively. An important special case that is relevant to the problem considered here is the case where the birth rate is constant and the lifetime of objects is random and distributed exponentially. In this case both $B(t)$ and $D(t)$ are Poisson processes with a constant birth rate $\lambda$ and a variable death rate $\mu = N(\langle \tau \rangle)$, respectively (see Taylor & Karlin, 1998). According to the definition of a Poisson process, its parameter—say, the birth rate $\lambda$—has the meaning of the ensemble mean number of newborn objects per unit of time. Equivalently, $\lambda dt$ is the probability that an object is generated during an infinitesimal time increment $dt$. Thus, equation (2) may be represented in the differential form as

$$dN = P(\lambda dt) - P\left(\frac{N}{\langle \tau \rangle} dt\right),$$

where $P(a)$ denotes a random variable with a Poisson distribution with the parameter $a$. In our case $\lambda$ is the rate of the appearance of clouds (the mean number of newborn clouds per time unit) which is determined by large-scale conditions. This is the same parameter that is used by the original PC08 scheme to produce clouds. The relation $\mu = N(\langle \tau \rangle)$ for the death rate is based on the fact that for an exponential lifetime distribution, the so-called failure, or mortality, rate of an object (expressed in the number of hypothetical failures per unit of time if an object were immediately restored after each “failure”) is a constant in time being equal to $1/\langle \tau \rangle$. Then the total death rate of all currently living objects is the failure rate of a single object times the object number $N$. In general, for an arbitrary lifetime distribution, $\mu$ is also the sum of failure rates over all living objects. However, only an exponential distribution of lifetime possesses the property that the failure rate of an object is constant in time. At that point, it is easy to understand why only a system with exponentially distributed lifetimes admits a representation of $N(t)$ as realizations of a single SDE. Indeed, if the failure rate of a single object were not constant, but a function of time, then to determine the total death rate $\mu$, it would be necessary to know the “life stage” of each cloud; in other words, the complete history of birth times of currently living clouds should be kept. Thus, it would be impossible to uniquely determine the death rate $\mu$ using the information of the current value of $N(t)$ only. As an example, it can be noticed that in the original PC08 formulation, the failure rate of a cloud coincides with the lifetime PDF itself and is merely the delta function $\delta(t - \tau_0)$; thus, the scheme has to track the history back to the time $t - \tau_0$. For the same reason, a Gaussian distribution for the lifetimes centered at $\tau_0$, which would be a closer substitute for the constant lifetime than an exponential one, cannot lead to a single differential equation.

The relation for the death rate $\mu = N(\langle \tau \rangle)$ is also easy to understand intuitively as expressing the fact that the probability that an object disappears during $dt$ increases with the increasing number of the living objects and the decreasing mean lifetime.
One may ask whether it is necessary to use a Poisson process to model the decrement term $D(t)$ and whether it would be possible to use another sort of stochastic process in order for the model to agree better with the original PC08 scheme. It may be noticed that in the original PC08 scheme, the death process $D(t)$ represents precisely a Poisson process, being merely equal to the given Poisson birth process $B(t)$ taken with the negative sign and shifted by $\tau_0$ forward in time. Thus, the use of a Poisson process to model the reduction of $N$ is justified. The difference between the representation of this term in the original PC08 scheme and in equation (3) is that in the original PC08 scheme, the process $D(t)$ depends on the past of the process $B(t)$, whereas in equation (3), the processes $B(t)$ and $D(t)$ are mutually independent. In the Markovian PC08 scheme, the death process $D(t)$ also represents a Poisson process. On the one hand, the death process $D(t)$ in this system is also uniquely determined by the past of the process $B(t)$, because in this system only the cloud birth is random and the rest behavior is deterministic, as in the original formulation. On the other hand, due to the randomness of cloud lifetimes, this system is statistically indistinguishable from a system without memory such as equation (3). “Statistically” means that although, given the same input birth process, the exact reproduction of each sample path is not possible by equation (3), all statistics produced by equation (3) coincide with those obtained by the Markovian PC08 scheme.

At each time instant, the increments $dB(t)$ and $dD(t)$ may take values of either 0 or 1, according to the definition of a Poisson process which states that only one event may occur at a time instant. It is worth providing here the average of equation (3) over sample paths,

$$\frac{d\langle N \rangle}{dt} = \lambda - \langle N \rangle (\tau).$$

(4)

The system equilibrates at

$$\langle N \rangle = \lambda (\tau).$$

(5)

Recall now that the quantity we are interested in is not the cloud number $N$ but the total mass flux $M$ of all the clouds in a grid box. Since the mass flux of a single cloud at the cloud base is also a random variable, it is necessary to extend the problem and to supplement equation (3) with another SDE for $M$, which can be derived in the following way. Clearly, it should also have two terms on the right-hand side describing gain and loss of $M$ commensurating with gain and loss of $N$

$$dM = dM^+ - dM^-.$$

(6)

The random variables $dM^+$ and $dM^-$ should be mutually independent, as $B$ and $D$ are (birth and death processes are not directly related to each other). In contrast to $N$, which has a point distribution confined to whole numbers, the random variable $M$, as well as $dM^+$ and $dM^-$, is continuously distributed over $[0, +\infty)$. Let us find the probability that the random variable $dM^+$ has a value between $m$ and $m + dm$ during an infinitesimal time increment $dt$. As was already mentioned, only one event is allowed to occur, that is, only one cloud may appear or disappear, at each time instant. Thus, $dM^+$ is the mass flux of at most only one newborn cloud (similarly for $dM^-$). The event “a cloud with the mass flux between $m$ and $m + dm$ during the time $dt$ is born” is the composition of two events: first, a cloud is born (with point probability $\lambda dt$) and second, this cloud has the mass flux between $m$ and $m + dm$ (this is determined by the prescribed probability density $p_{\delta}(m) = \frac{1}{m_b}e^{-m/m_b}$). Taken together, the probability density of $dM^+$ is

$$p_{dM^+}(m) = \begin{cases} \frac{\lambda dt}{m}e^{-m/m_b}, & m > 0, \\ (1 - \lambda dt)\delta(m), & m = 0. \end{cases}$$

(7)

Here $\delta(m)$ is the Dirac delta function. To find the distribution of $dM^-$, one can start from similar considerations. The probability that a cloud disappears during $dt$ is $Ndt/(\tau)$ where $N$ is the instantaneous cloud number to be taken from equation (3). Further, to characterize the event “the disappearing cloud has the mass flux between $m$ and $m + dm$,” instead of the mass flux distribution of newborn clouds $p_{\delta}(m)$, the probability density of the mass flux of clouds “ready-to-disappear” should be used, which will be denoted as $p_{\delta}(m)$.

The determination of this probability density presents the greatest difficulty in the derivation. In contrast to the distribution of newborn clouds, which is constant in time and the same for all sample paths (because it is a prescribed input to the problem and is independent of the problem’s internal variables), the distribution of the mass fluxes of disappearing clouds evolves in time individually for each sample path (because it
depends on the current set of living clouds and their lifetimes). This distribution is itself random, or saying it more exactly, this is a two-variate random function \( p_d(m, t) \) evolving in time and keeping the integral over \( m \) equal to unity. There seems to be no general analytical expression for following manner. The stationary and ensemble-averaged distribution of disappearing clouds is closed...

Assume now that the instantaneous and realization-specific...

Thus, in our case,

\[
\langle p_d(m) \rangle^{st} = \frac{1}{\langle m_d \rangle} e^{-\frac{m}{\langle m_d \rangle}}, \quad \langle m_d \rangle = m_s,
\]

where \( m_s \) is the actual mean mass flux of disappearing clouds and \( \langle m_d \rangle \) is its ensemble-averaged value. Although this result is exact, this is not yet the distribution that should be used in place of \( p_d(m) \). As was mentioned above, the instantaneous and realization-specific \( p_d(m) \) but not the stationary and ensemble-averaged \( \langle p_d(m) \rangle^{st} \) has to be used in equation (6). This is easier to understand, if one realizes that this situation resembles the case of the simple ordinary differential equation with a relaxation term on the right-hand side, such as \( \frac{dy}{dt} = -\gamma y + \ldots \). Clearly, the actual instantaneous value of \( y \) should enter the relaxation term but not its (time) average; otherwise, there would be no correct relaxation. Similarly, \( p_d(m) \) may also fluctuate around its equilibrium value \( \langle p_d(m) \rangle^{st} \), and if at some time instant there are eventually more small and less large clouds than on average, the structure of the distribution of disappearing clouds will adjust accordingly.

Assume now that the instantaneous \( p_d(m) \) keeps the functional form of \( \langle p_d(m) \rangle^{st} \), that is, it is also exponential, and only the parameter of the distribution, which is the reciprocal of its mean, is allowed to vary. That means that instead of the mean mass flux per newborn cloud \( m_s \), the actual instantaneous mean mass flux per cloud \( \overline{m} = \frac{M}{N} \) should be used,

\[
p_d(m) = \frac{1}{M/N} e^{-\frac{m}{M/N}},
\]

whereby both quantities \( M \) and \( N \) are available through their prognostic equations. In Figure 1, the computed by means of the Markovian PC08 scheme \( p_d(m) \) is compared with \( p_d(m) \) given by the expression (8) for two distinct values of the normalized quantity \( \frac{1}{m_m} \). The agreement between the observed and the approximating PDFs is so precise that one may anticipate that the expression (8) has something to do with the exact solution rather than being just an approximation.

With the expression (8), the term \( p_{\Delta M^-}(m) \) describing the probability density that a cloud with the mass flux \( m \) disappears during \( dt \) is closed

\[
p_{\Delta M^-}(m) = \begin{cases} \frac{N\delta}{(\tau)} p_d(m) = \frac{N\delta}{(\tau)} \frac{1}{M/N} e^{-\frac{m}{M/N}}, & m > 0, \\ \left(1 - \frac{N\delta}{(\tau)}\right) \delta(m), & m = 0. \end{cases}
\]
It can be noticed that equation (3) for $N$ is independent of equation (6) for $M$, but the latter equation is coupled to the former one through the “relaxation” term describing the disappearance of clouds. Thus, although the cloud number per se is not needed as an output of the PC08 scheme, the equation for $N$ should still be carried.

It is worth noting that this representation of the fluctuating mass flux by equation (6) is possible if the mass flux of a single cloud does not change in time from the cloud birth until its disappearance, which is the case in the PC08 scheme and in the basic version of the SSH15 scheme. If, as it is also proposed in SSH15, each cloud experiences a true life cycle with the cloud base mass flux evolving in time, such a representation is not possible. This somewhat resembles the situation with the representation of the cloud number by equation (3), which is possible only if the failure rate of a single cloud is constant in time. The pair of equations (3) and (6), together with the expressions (7) and (9), can be numerically solved and compared with the results obtained by the original and Markovian PC08 schemes. In the discretized numerical representation of equation (3), the terms on the right-hand side are the number of clouds that are born, $\Delta B$, and disappeared, $\Delta D$, during the time step $\Delta t$. They are represented by independent random variables sampled from the Poisson distribution with the respective parameters $\lambda \Delta t$ and $N \langle \tau \rangle \Delta t$. In equation (6), $dM^+ (dM^-)$ is the sum of $\Delta B (\Delta D)$ independent random variables sampled from the exponential distribution with the parameters $m_b$ and $M/N$, respectively.

For the cloud number $N$, equation (3) theoretically exactly corresponds to the Markovian PC08 scheme, thus, strictly speaking, an additional comparison of the numerical results for $N$ is redundant. However, for the modeling of the total mass flux $M$ an additional assumption on the distribution $p_d(m)$ has been invoked, and thus, the numerical comparison is desirable to justify the approximation. To this end, some considerations on how to compare two stochastic processes using numerical simulations are in order. Generally speaking, a stochastic process $X(t)$ is characterized by its infinite-dimensional joint distribution with respect to all time instants. It means that the metric in which the equivalence or the proximity of two stochastic processes can be established is, strictly speaking, the difference between the two infinite-dimensional joint distributions, which is obviously difficult to measure numerically. In practice, we may restrict the comparison to some first dimensions. We will consider the stationary state discarding the spin-up stage, so that the distributions do not depend on the absolute time. The one-dimensional distribution with respect to time, that is, the so-called marginal distribution, is the distribution of $X(t)$ at arbitrary time. It can in turn be characterized by some of

\[ p(N), \quad p(M) \]
their low-order moments that may be compared for two PDFs (alternatively, two PDFs may be compared using various metrics to measure the difference between PDFs). The two-dimensional distribution is in fact a one-parameter family of the joint distributions at two time instants spaced a varying time lag $\Delta t$. Among all its first and second moments, this is the autocovariance $\langle X(t)X(-\Delta t)t \rangle$ (or, equivalently, the autocorrelation function (ACF), which is the autocovariance normalized by $\langle X^2(t) \rangle$) only that adds a new information to what the marginal distribution already contains.

For the comparison, we have run three ensembles of realizations, one for the original PC08 scheme, one for the Markovian PC08 scheme, and one solving the pair of SDEs (3) and (6). Each ensemble contained 100 members, the clusters were run for 200 hr following the spin-up phase (5 hr) which is excluded from the analysis.

In Figure 2, the fine-grained original, the fine-grained Markovian, and coarse-grained marginal distributions of $N$ and $M$ (normalized with $m_b$) are shown for two contrast cases, one with the low birth rate ($\lambda = 0.001$ cloud per second) and one with the high one ($\lambda = 0.1$ cloud per second). As it follows from the stationary equation for the ensemble-averaged cloud number $\langle N \rangle$ (equation (5)), with the mean lifetime $\langle \tau \rangle$ of 45 min this amounts to the mean cloud number of 2.7 and 270 clouds, respectively. From Figure 2, first, it can be noticed that the marginal distributions of $N$ and $M$ for the Markovian PC08 scheme are statistically indistinguishable from the original PC08 scheme with the constant cloud lifetime. This is an interesting side result which probably may be explained by the independence of individual cloud mass fluxes, in particular, of individual clouds with different $m$, and by the independence of cloud lifetime of its mass flux. One can also note that the same result, the exact representation of the marginal $N$ and $M$ distributions, comes out for a similar system but with Gaussian-distributed lifetimes. This indicates that this is rather a common feature not characteristic for the only exponential distribution (but recall that only an exponential distribution leads to a Markovian system).

Further, it can be seen that the marginal distributions obtained by equations (3) and (6) coincide with those produced by both PC08 schemes.

In Figure 3, the fine-grained and the coarse-grained autocorrelation functions for $N$ are shown for the case $\lambda = 0.001$. The autocorrelation function does not depend on $\lambda$, so that for $\lambda = 0.1$ the results are identical to those shown in Figure 3. The autocorrelation functions of $M$ are also identical to that of $N$ shown in Figure 3. It can be seen that, in contrast to marginal distributions, the time behavior is not the same for the original and the Markovian PC08 schemes. In the original scheme, the autocorrelation function decreases linearly, reaching zero at the lag $\tau_0 = 45$ min and is exactly zero at larger lags. In the Markovian PC08 scheme, as well as in the stochastic equation system, the autocorrelation function is exponential with the decorrelation ("e-folding") time equal to $\langle \tau \rangle = \tau_0$. For equation (3), its exponential autocorrelation function may also be found analytically (Taylor & Karlin, 1998). The approximation of the linear decay by an exponential curve may be regarded as not too bad: it has the same slope in the vicinity of the zero lag, and the simulated process becomes decorrelated (the autocorrelation falls to $1/e$) at the lag at which the autocorrelation should be equal to zero according to the original formulation. The present discrepancy, mostly pronounced at the lags between $\sim \tau_0$ and $2\tau_0$, from the physical point of view means some additional weak "memory." More exactly, the Markovian PC08 scheme and the solution of equations (3) and (6) produce additional weak temporal correlations for $N$ and $M$ as compared to the original PC08 formulation. Whether this agrees better or not with the real behavior of natural clouds, is an open question. It may just be mentioned that, for example, in the practical implementation of the PC08 scheme into MOGREPS, the cloud lifetime $\tau_0$ was adjusted to 15 min instead of 45 min in the default formulation (Keane et al., 2016). Thus, the autocorrelation function of equation systems (3) and (6) falls well within the uncertainty in the autocorrelation function between both true PC08 scheme versions proposed in PC08 and implemented in the MOGREPS.

The normalized third-order moments $\langle N(t)N(t-\Delta t)/N^2(t) \rangle$ and $\langle M(t)M(t-\Delta t)/M^2(t) \rangle$ are identical to the respective autocorrelations for each of the three experiments and are not shown here.
4. The Intermediate Case: The Exponential Mass Flux Distribution and the Linear Relation Between the Lifetime and the Mass Flux

As was noted in section 2 where the brief description of two stochastic schemes is given, the PC08 scheme is built upon the assumption of an exponential distribution for living clouds. It is worth reminding that this is the independence of the cloud lifetime of cloud mass flux which implies the same exponential distribution for generated clouds. If the cloud lifetime were a function of its mass flux, the distributions for generated and living clouds would differ. Despite that these are living clouds, for which the cloud-resolving modeling data suggest an exponential distribution in the deep convection case (Cohen & Craig, 2006), for our purposes here it is convenient to consistently regard both schemes as built upon the assumption on the distribution for generated clouds. From the mathematical point of view, therefore, the stochastic scheme describing shallow convection, which will be considered in the next section, differs from PC08 scheme for deep convection in two points. First, instead of an exponential distribution of newborn clouds, a Weibull or a mixed Weibull distribution is assumed. Second, the lifetime of a cloud is assumed to be a power law function of its mass flux, instead of the constant lifetime in the PC08 scheme. Before going to that scheme that is mathematically more sophisticated than the PC08 scheme in several aspects at once, we demonstrate the applicability of the approach developed in the previous section to a hypothetical scheme which represents only one step toward the increasing complexity as compared to the PC08 scheme. Specifically, we introduce into the PC08 scheme the simplest linear relation between the cloud lifetime and its mass flux. Note that this might be another (and physically not unrealistic) way of the modification of the PC08 scheme which results in an exactly solvable problem, at least with respect to \( N(t) \). To show this, first, it is worth reminding the formula for the transformation of the probability distribution under a change of the independent variable. If a random variable \( x \) is transformed to a random variable \( y \) by means of a monotonic function \( y = f(x) \), their respective probability density functions are related to each other as

\[
p_Y(y) = \frac{p_X(x)}{|f'(x)|}.
\]  

(10)

Hence, if the mass fluxes of generated clouds are distributed exponentially

\[
p_b(m) = \frac{1}{m_b} e^{-\frac{m}{m_b}}
\]

and there is a linear relation between the cloud lifetime and its mass flux, \( \tau = am \), the cloud lifetimes are distributed according to (10) as

\[
p(\tau) = \frac{1}{am_b} e^{-\frac{\tau}{am_b}}.
\]

that is, also exponentially with the parameter \( am_b \), which is the mean cloud lifetime. As was already mentioned, in this case of the exponentially distributed lifetime (and only in this case) an exact representation of \( N(t) \) by a Markov stochastic process is possible, because only in this case the "failure rate" of each cloud is constant in time and the sink term in equation (3), which is equal to the sum of the current failure rates of all living clouds, does not explicitly depend on the history. The exact solution for the total mass flux \( M \) is, however, not guaranteed for the considered system; in this section the solutions for \( N \) and \( M \) under the application of the approach developed for the PC08 scheme are investigated. Moreover, we propose two equation systems and their solutions, each of which reproduces the original formulation exactly with respect to some characteristics and approximately with respect to others.

If, as is shown above, the cloud lifetime is distributed exponentially, the problem of the determination of \( N \) is the same as in the case of the Markovian PC08 system. Indeed, the equation for \( N \) is independent of the equation for \( M \); in other words, the equation that counts the number of objects does not know anything about an additional property (in our case mass flux) possessed by the objects. Thus, the only information of the exponentially distributed lifetimes suffices to arrive at equation (3) with respective parameters \( \lambda \) and \( \frac{N}{N} = \frac{N}{m_b} \). As to the equation for \( M \), the formulation (6) together with the expression for its gain term \( dM^* \) (7) are still valid. The loss term (9) or, more exactly, the distribution of disappearing clouds \( p_d(m) \) (the expression (8)), has however to be reformulated. We will proceed as in the previous section, that is, first, we find the stationary and ensemble-averaged distribution \( (p_d(m))^\tau \). Then, in order to determine an instantaneous and trajectory-specific \( p_d(m) \), we make a similar assumption as previously, namely, that the functional form of
\( (p_d(m))^a \) is kept and only its parameter value is expressed through the actual \( N \) and \( M \) instead of being stationary and ensemble averaged.

As it was noticed in the previous section, the requirement of stationarity means that the stationary and ensemble-averaged distribution of disappearing clouds is equal to the distribution of newborn clouds

\[
(\langle p_d(m) \rangle)^a = p_b(m). \tag{11}
\]

Thus, as in the previous section,

\[
(\langle p_d(m) \rangle)^a = \frac{1}{\langle m_d \rangle} e^{-\frac{m}{\langle m_d \rangle}}, \quad \langle m_d \rangle = m_b,
\]

where \( m_d \) is the actual mean mass flux of disappearing clouds and \( \langle m_d \rangle \) is its ensemble-averaged value. Now, in order to specify the instantaneous and trajectory-specific \( p_d(m) \), the actual mean mass flux of disappearing clouds \( m_d \) should be expressed through the actual available variables \( N \) and \( M \) or the mean actual mass flux of living clouds \( \bar{m} = \frac{M}{N} \). Recall that in the case of the PC08 scheme, where the cloud lifetime and the cloud mass flux were not correlated, all the three distributions, one of generated clouds and the stationary and ensemble-averaged distributions of living and disappearing clouds, were equal to each other; hence, \( \langle m_d \rangle = \langle \bar{m} \rangle = \langle \frac{M}{N} \rangle \), and we assumed that \( m_d = \frac{M}{N} \). Now, due to the relation between the cloud lifetime and its mass flux, the smaller clouds will disappear with preference over the larger ones and the distribution of disappearing clouds \( (p_d(m)) \) is no longer identical to the distribution of living clouds \( (p(m)) \); thus, the previous assumption may not be made. Now our strategy will be to find the relation between \( m_d = \langle m_d \rangle \) and \( \langle \bar{m} \rangle = \langle \frac{M}{N} \rangle \) and then to substitute \( \frac{M}{N} \) instead of \( \langle \frac{M}{N} \rangle \) in this relation. To find \( \langle \bar{m} \rangle \), we may invoke the distribution of living clouds \( (p(m)) \) which we will now seek for and then compute its mean.

In order to find the stationary and ensemble-averaged distribution of living clouds \( (p(m))^a \), consider first an ensemble-averaged prognostic equation for the density of the number of clouds with the lifetime \( \tau \), denoted by \( \langle n(\tau) \rangle \). Since all such clouds have the same fixed lifetime, the problem of the determination of their number resembles the problem of the modeling of the original PC08 scheme. As was shown in the previous section, this problem has no exact description in terms of a stochastic equation, but there exists an approximating equation, the ensemble mean of which reads (cf. equations (4) and (7))

\[
\frac{d\langle n(\tau) \rangle}{dt} = \lambda p_b(\tau) - \frac{\langle n(\tau) \rangle}{\tau}, \tag{12}
\]

where \( \lambda p_b(\tau) \) is the mean number density of clouds with the lifetime \( \tau \) generated during the time unit. Transforming this equation in terms of \( m \), one has

\[
\frac{d\langle n(m) \rangle}{dt} = \lambda p_b(m) - \frac{\langle n(m) \rangle}{am}. \tag{13}
\]

Its stationary solution is

\[
\langle n(m) \rangle^a = \lambda a m p_b(m), \tag{14}
\]

and therefore, since the probability density of the cloud mass flux is the normalized number density of clouds with different mass fluxes,

\[
\langle p(m) \rangle^a = \frac{\langle n(m) \rangle^a}{N} \sim m p_b(m), \tag{15}
\]

with an appropriate normalization prefactor that ensures that the probability density integrates to unity. It should be noted that, although equation (12) is approximate, its stationary solution (14), which we seek in order to determine \( \langle \bar{m} \rangle \), is nevertheless exact. We have already seen this in the previous section—replacing the constant lifetime with the exponentially distributed random lifetime distorts only the time behavior of the solution, but the stationary ensemble-mean value \( \langle (N)^a \rangle \) in the previous section, \( \langle n(m) \rangle^a \) here) is reproduced precisely.

According to (15), if \( p_b(m) \) is an exponential distribution with the parameter \( m_b \), \( \langle p(m) \rangle^a \) is then a gamma distribution with the parameters \( m_b \) and 2

\[
\langle p(m) \rangle^a = \frac{m e^{-\frac{m}{m_b}}}{\Gamma(2)m_b^2}, \tag{16}
\]
where $\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx$ is the gamma function. According to the properties of a gamma distribution, its mean $\langle m \rangle = 2m_b$. It is interesting to compare this result with the result obtained previously for the case where the lifetime of a cloud and its mass flux were not related to each other, where we had $\langle m \rangle = m_b$. As one can see, it is the relation between the lifetime of a cloud and its mass flux that shifts the mean of the mass flux of living clouds to the larger values as compared to the mean mass flux of newborn clouds: small clouds, although extensively appearing, die quickly; thus middle-sized clouds are able to contribute more to the average. Since $\langle m \rangle = \langle \frac{M}{N} \rangle$ and $\langle m_b \rangle = m_b$, one has $\langle m_d \rangle = m_b = \langle \frac{M}{N} \rangle = \frac{1}{2} \langle \frac{M}{N} \rangle$. Finally, we make an assumption on the instantaneous distribution parameter, $m_d = \frac{2}{7} = \frac{1}{2} \langle \frac{M}{N} \rangle$, which completes the formulations of the distribution of disappearing clouds (cf. the expression (8))

$$p_d(m) = \frac{1}{m/2} e^{-\frac{m}{2m_b}}$$

and of the loss term in the equation for $M$ (cf. the expression (9))

$$p_{\Delta M}^{\text{loss}}(m) = \begin{cases} \frac{N\lambda}{am_b} p_d(m) = \frac{N\lambda}{am_b} \frac{1}{m/2} e^{-\frac{m}{2m_b}}, & m > 0, \\ \left(1 - \frac{N\lambda}{am_b}\right) \delta(m), & m = 0. \end{cases}$$

This equation system will be referred to as system A.

A pair of the ensemble runs were performed, one realizing the true cloud system (exponential distribution of generated clouds and the linear dependence of the lifetime on the mass flux) and the second one solving the corresponding system of equations (3) and (6) together with the expressions (7) and (17). The cloud generation rate $\lambda$ in these runs was set to 0.01, and the mean mass flux per generated cloud $m_b$ and the parameter $\alpha$ were set to the values adopted in the scheme of SSH15, that is, to 29868.46 kg/s and 0.33 s$^2$/kg, respectively. This implies the mean lifetime of generated clouds of $\sim 165$ min, which, although not very realistic, does not matter for illustrative purposes, and the stationary mean cloud number $N = \lambda am_b = 99$ clouds. In Figures 4a and 4b the respective marginal distributions for $N$ and $M/m_b$ and in Figures 4c and 4d their autocorrelation functions are shown, in green (the true system) and in red (equations (3), (6), (7), and (17)). As one can see, both characteristics of the cloud number $N$ (Figures 4a and 4c) are reproduced precisely. For the marginal distribution of $M$ (Figure 4b), the system of stochastic equations reproduces correctly the form of the distribution but somewhat underestimates the variance of $M$ with respect to the
true stochastic scheme. The autocorrelation function of the total mass flux \( M \) (Figure 4d) is reproduced although not absolutely exactly but to a very good approximation.

Yet another way of thinking may be proposed for the derivation of the SDEs for \( N \) and \( M \) which leads to a slightly different result. One may think of the cloud number \( N \) as of the integral of the cloud number density over all mass fluxes (or lifetimes). Then, the approximate equation for \( N \) may be obtained by the integration of the approximate equation for the number density of clouds with the mass flux \( m \) over \( m \). The result is (cf. equation (13))

\[
dn(m) = P(\lambda p_b(m)dt) - P\left(\frac{n(m)}{am}dt\right).
\]

(The analogous derivation may be conducted in terms of the lifetime \( \tau_c \).) As to \( n(m) = N p(m) \) which enters (18), one may apply the same approximation as previously, that is, keep the form of the stationary ensemble-mean distribution (16) but take the sample-specific parameter \( m'_b \), so that in the considered case \( p(m) \) would be a gamma distribution with the parameters \( m'_b = \frac{2}{1} \frac{M}{2N} \) and 2

\[
p(m) = \frac{m \exp\left(-\frac{m}{\frac{M}{2N}}\right)\Gamma(\frac{2}{1} \frac{M}{2N})}{\Gamma(2)\left(\frac{M}{2N}\right)^2}.
\]

Multiplying (19) with \( N \), inserting \( n(m) = N p(m) \) into (18) and integrating the result over \( m \) yields

\[
dN = P(\lambda dt) - P\left(\frac{N}{\langle \tau_c \rangle}dt\right).
\]

where \( \langle \tau_c \rangle = a m'_b = \frac{M}{2N} \) is the sample-specific mean lifetime of disappearing clouds or, more exactly, conditional mean lifetime of disappearing clouds given the actual \( N \) and \( M \). Comparing equation (20) with equation (3), which was used in solution A (producing the red curves in Figure 4), one can see that the difference between them is in the specification of the lifetime of disappearing clouds. Whereas the solution A uses the overall mean lifetime of disappearing clouds, this time the mean lifetime given the actual \( N \) and \( M \) is used. In conformity with (20), the loss term \( dM^- \) in equation for \( M \) (6) should be changed to (cf. (17))

\[
P_{dM^-}(m) = \begin{cases} N\frac{\lambda dt}{\langle \tau_c \rangle} \frac{1}{m'_{b2}} e^{-\frac{m}{m'_{b2}}}, & m > 0, \\ \left(1 - \frac{N\lambda dt}{\langle \tau_c \rangle}\right)\delta(m), & m = 0. \end{cases}
\]

The gain term \( dM^+ \) given by (7) does not require any change. This equation system will be referred to as the solution B.

In Figure 4, the blue lines show the solution of equations (20), (6), (7), and (21). It is interesting to notice that this time the marginal distributions of both \( N \) and \( M \) are reproduced exactly, whereas the time behavior deviates from the time behavior of the fine-grained model. In summary, both solutions reproduce exactly the marginal distribution of the cloud number. The solution A correctly reproduces the time behavior of \( N \) and almost correctly that of \( M \) but has a somewhat smaller variance of the marginal distribution of \( M \). The solution B correctly reproduces both marginal distributions but is not exact in both autocorrelation functions, although both autocorrelation functions for \( N \) and \( M \) have nevertheless the right slope at the zero lag, which means the correct time behavior at short time scales.

Which of the two solutions is preferable, is difficult to judge. From the theoretical point of view, the equation system A seems to be more consistent than system B since, as was already mentioned, the equation for \( N \) should be unaware of the value of \( M \) in the case of the exponentially distributed lifetime. In the equation system B the information on current \( M \), however, enters the computation of the conditional mean lifetime of disappearing clouds \( \langle \tau_c \rangle \) which is used then in both equations. However, the good results provided by solution B—in particular, the precise reconstruction of both marginal distributions—makes it worthwhile to consider this system as an approximation to the original fine-grained system.
5. The Shallow Convection Case: The Scheme of SSH15

The intermediate complexity case elaborated in the previous section—-with an exponential distribution of newborn clouds and the linear relation between the lifetime of a cloud and its mass flux—does not correspond (so far) to any of the existing stochastic cloud scheme, although it might represent a plausible extension of the conventional PC08 scheme. Nevertheless, its consideration was instructive in that it has demonstrated that only one step in the direction of the increasing complexity as compared to the Markovian PC08 scheme—the introduction of the simplest linear relation between the lifetime and the mass flux—suffices for the approach developed in the section 3 to fail in the exact reproduction of the corresponding fine-grained model (recall that in the case of an exponential distribution of newborn clouds and in absence of any relation between the lifetime and the mass flux, the problem admits an exact solution). Starting from the beginning with the scheme of SSH15, which is mathematically more complex than the Markovian PC08 scheme in several aspects at once, it would not be clear, which of its properties is responsible for the ensuing discrepancies between the full scheme and its coarse-grained representation. In this section, we apply, step by step, the reasoning that has been developed for the intermediate case, to the scheme of SSH15. It may be anticipated that similar deviations from the exact solution will be obtained also for this scheme.

For the sake of comprehensibility, the analysis is performed assuming that the distribution of the mass flux of generated clouds is an ordinary Weibull distribution

\[ p_b(m) = \frac{k}{m_b} \left( \frac{m}{m_b} \right)^{k-1} e^{-\left(\frac{m}{m_b}\right)^k} \]  

(22)

instead of a mixed Weibull distribution given by (1). The generalization of the results to a mixed Weibull distribution does not present any essential difficulty.

As in the previous section, two possible variants of the equation system are suggested for the shallow convection scheme. They differ only in the mean lifetime of disappearing clouds used — the overall mean (model A) or the conditional mean given the actual \( M \) and \( N \) (model B). The first five steps are the same for both variants.

(1) The ensemble-averaged prognostic equation for the number density of clouds with the mass flux \( m \) (13) is now amended according to the power-law dependence of the lifetime on the mass flux

\[ \frac{d\langle n(m) \rangle}{dt} = \lambda p_b(m) - \langle n(m) \rangle \frac{\langle m(m) \rangle}{am^\beta}. \]  

(23)

Its stationary solution (cf. (15)), that is, the stationary and ensemble-averaged distribution of living clouds, in the case of a Weibull distribution for the mass fluxes of generated clouds (22) is a generalized gamma distribution with the parameters \( m_b, k + \beta \), and \( k \)

\[ \langle p(m) \rangle_{st} = \frac{k}{\Gamma(1 + \frac{\beta + 1}{k}) m_b^{k+\beta-1} e^{-\left(\frac{m}{m_b}\right)^k}}. \]  

(24)

It should be noted that now this expression appears to be in principle approximate as opposed to (16), because it is exact if the lifetime is distributed exponentially, which is now not the case. However, in the specific case of the given shallow convection scheme, it is very close to being exact; see the discussion on the lifetime distribution below.

(2) Find the mean of \( \langle p(m) \rangle_{st} \) as

\[ \langle \overline{m} \rangle = m_b \frac{\Gamma \left( 1 + \frac{\beta + 1}{k} \right)}{\Gamma \left( 1 + \frac{\beta}{k} \right)}. \]  

(25)

This expression provides us with the relation between the distribution parameter \( m_b \) and the stationary ensemble-averaged mean mass flux per cloud \( \langle \overline{m} \rangle = \langle \frac{M}{N} \rangle \).
(3) Make the assumption on the sample-specific parameter \( m'_b \) by inverting (25) with respect to \( m_b \) and substituting the actual \( \frac{M}{N} \) instead of \( \langle \frac{M}{N} \rangle \)

\[
m'_b = \frac{M}{N} \frac{\Gamma\left(1 + \frac{\beta}{k}\right)}{\Gamma\left(1 + \frac{\beta + 1}{k}\right)}.
\]

(4) Find the stationary ensemble-averaged distribution of disappearing clouds \( \langle p_d(m) \rangle^\text{st} \) as being equal to the distribution of the generated clouds

\[
\langle p_d(m) \rangle^\text{st} = p_s(m). \quad \text{(11) revisited}
\]

In this case this is a Weibull distribution (22).

(5) Use \( m'_b \) from step 3 instead of \( m_b \) in the \( \langle p_d(m) \rangle^\text{st} \) to obtain the sample-specific \( p_d(m) \),

\[
p_d(m) = \frac{k}{m'_b} \left( \frac{m}{m'_b} \right)^{k-1} e^{-\left(\frac{m}{m'_b}\right)^{\frac{1}{k}}}.
\]

The first equation system (the solution A) is then completed by

6a) Find the (unconditional) mean lifetime of disappearing clouds

\[
\langle \tau \rangle = \int_0^\infty \tau \langle p_d(m) \rangle^\text{st} \, dm = \int_0^\infty a m^\beta p_s(m) \, dm = a m_b^\beta \Gamma\left(1 + \frac{\beta}{k}\right)
\]

The complete equation system is then formulated as

\[
dN = P(\lambda dt) - P\left(\frac{N}{\langle \tau \rangle} dt\right) \quad \text{(3) revisited}
\]

\[
dM = dM^+ - dM^- \quad \text{(6) revisited}
\]

with

\[
p_{dM^+}(m) = \begin{cases} 
\lambda dt p_s(m) = \lambda dt \frac{k}{m'_b} \left( \frac{m}{m'_b} \right)^{k-1} e^{-\left(\frac{m}{m'_b}\right)^{\frac{1}{k}}}, & m > 0, \\
(1 - \lambda dt) \delta(m), & m = 0,
\end{cases}
\]

\[
p_{dM^-}(m) = \begin{cases} 
\frac{N dt}{\langle \tau \rangle} \frac{k}{m'_b} \left( \frac{m}{m'_b} \right)^{k-1} e^{-\left(\frac{m}{m'_b}\right)^{\frac{1}{k}}}, & m > 0, \\
\left(1 - \frac{N dt}{\langle \tau \rangle}\right) \delta(m), & m = 0,
\end{cases}
\]

and where \( m'_b \) and \( \langle \tau \rangle \) are determined through (26) and (27), respectively.

The derivation of the equation system B proceeds as follows.

(6b) Use the sample-specific parameter \( m'_b \) from step 3 instead of \( m_b \) in the stationary and ensemble-averaged distribution of living clouds \( p(m) \) from step 1 to obtain the sample-specific distribution of living clouds

\[
p(m) = \frac{k}{\Gamma\left(1 + \frac{\beta}{k}\right)} m'^{k+\beta} e^{-\left(\frac{m}{m'_b}\right)^{\frac{1}{k}}}.
\]

(7b) Substitute \( n(m) = N p(m) \) into the approximate equation for the number of clouds with the mass flux \( m \) (cf. equation (18))

\[
dn(m) = P \left( \lambda p_b(m) dt \right) - P \left( \frac{n(m)}{a m^\beta} dt \right)
\]

\[
(30)
\]
and integrate it over \( m \), obtaining equation (20) with the conditional mean lifetime

\[
\langle \tau_c \rangle = \alpha m'^{\beta} \Gamma \left( 1 + \frac{\beta}{k} \right).
\]  

(31)

In this case the equation system consists of

\[
dN = P(\lambda dt) - P \left( \frac{N}{\langle \tau_c \rangle} dt \right) \tag{20}\text{ revisited}
\]

\[
dM = dM^+ - dM^- \tag{6}\text{ revisited}
\]

with \( p_{dM^+}(m) \) given by (28) and with \( p_{dM^-}(m) \) given by (29) where \( \langle \tau \rangle \) is replaced with \( \langle \tau_c \rangle \) (31).

Three ensemble runs were performed, one realizing the stochastic shallow convection scheme of SSH15 and two solving the two stochastic equation systems described above. The SSH15 scheme was run with the parameter values taken from SSH15, \( m_b = 29868.46 \text{ kg/s}, \alpha = 0.33 \text{ s}^2/\text{kg}, \beta = 0.72, \) and \( k = 0.7 \). In Figure 5 the statistics of the respective stochastic processes are shown. The results are qualitatively similar to those obtained in the previous section. As for the intermediate case with an exponential distribution of the generated clouds and the linear dependence of the mass flux on the lifetime, here the marginal cloud number distribution (Figure 5a) is reproduced exactly by both methods. The marginal distribution of the total mass flux (Figure 5b) is reproduced exactly by system B (that with the lifetime of disappearing clouds responding to current conditions), whereas system A (with the constant lifetime of disappearing clouds), correctly simulating the form of the distribution, produces somewhat smaller variance. System A, however, reproduces exactly the autocorrelation function of the cloud number (Figure 5c), while using system B leads to longer weak correlations; the autocorrelation function has nevertheless the right slope at the zero lag, so that up to the value of 0.6 the autocorrelation of system B is still in good agreement with the truth. Here it is interesting to confront the exact reproduction of the statistical properties of \( N \) by the equation system A with the statement from the previous sections that the exact solution for \( N \) is possible only if the lifetime is distributed exponentially. It is not easy to see this latter fact in the formulation of the stochastic shallow convection scheme directly. In fact, if the mass flux of generated clouds has a Weibull distribution (22) and the cloud lifetime is related to its mass flux as \( \tau = am^\beta \), the application of the transformation rule (10) yields...
the distribution of the lifetime \( \tau \) as

\[
p_b(\tau) = \frac{k}{a m_b \beta} \left( \frac{\tau}{a m_b} \right)^{\frac{\alpha}{\beta}} e^{-\left( \frac{\tau}{a m_b} \right)^{\frac{1}{\beta}}}.
\]

By coincidence, the values of empirical parameters \( k \) and \( \beta \) in the stochastic shallow convection scheme of SSH15 are almost the same (\( k = 0.7 \) and \( \beta = 0.72 \)). It follows that the newborn cloud lifetime distribution in this scheme is indeed nearly exponential.

It is difficult to say which one of the two equation systems is better in terms of the total mass flux autocorrelation function (Figure 5d). System A is closer to the truth on the integral scale, while system B has the right slope at the zero lag and thus better agreement at short times. In summary, both equation systems are close to the exact solution. This is the user’s choice, which one of the two is preferable—the choice may also depend on what aspect, instantaneous mass flux distribution or its time behavior, a user finds more important. Taking into account the above analysis of the results, it may be argued that on an overall basis, system B is probably a more adequate practical representation of the shallow convection stochastic scheme (even if it seemingly contains more approximations!).

6. Conclusions

In the present study, an attempt is undertaken to translate the machinery of the stochastic convection schemes for deep (Plant & Craig, 2008) and shallow (Sakradžija et al., 2015) convection into the language of stochastic ordinary differential equations in order to save the computational costs required for the microscopic sampling schemes. In those schemes, the total mass flux at cloud base in a grid box is determined as the sum of mass fluxes of single clouds launched randomly in time. The rate at which clouds are generated is consistent with large-scale conditions and is determined by the total mass flux at cloud base computed by a deterministic convection scheme. To represent the fluctuating mass flux at cloud base coming from those schemes as a solution of an SDE, it is proposed to utilize the framework of birth-death stochastic processes. Restricting oneself to one first-order differential equation per variable, that is, to a system without memory, it seems to be impossible to represent the PC08 scheme with such equations, because the constant lifetime \( \tau_0 \) for all clouds means the dependence of the number of disappearing clouds on the number of newborn clouds \( \tau_0 \) time units ago, that is, a system with memory. It is, however, possible to reformulate the PC08 scheme—by replacing the constant lifetime with the exponentially distributed lifetime with the same mean—so that the resulting Markovian system admits an exact representation in the form of two ordinary SDEs for the cloud number and for the total mass flux, respectively. Interestingly, despite this seemingly crude modification, the resulting random processes describing the stochastic evolution of the cloud number and of the total mass flux are very close in properties to the dynamics of those quantities in the original PC08 scheme. The marginal distributions of both quantities are reproduced exactly, and the autocorrelation functions are recovered to a good approximation.

Within the proposed framework, it is also possible to reproduce the behavior of the shallow convection scheme of SSH15 to a good approximation. As shown in section 4, the discrepancies between the fine-grained models and their coarse-grained representations appear already with the introduction of a simplest linear dependence of the cloud lifetime on the cloud mass flux into the Markovian PC08 scheme. These discrepancies, being not large and acceptable in practical applications, are then inherited by the more complex system of the shallow convection scheme. For these cases, with a dependence of the lifetime on the mass flux, we propose two ways of solution both being intuitively plausible, each in its own way. The difference lies in the formulation of the mean lifetime of disappearing clouds—it is either the overall unconditional mean (system A) or the conditional mean given the current values of macroscopic variables \( N \) and \( M \) (system B).

Each of the two systems reproduces the original formulation exactly with respect to some characteristics and approximately with respect to others. A user may choose between two possibilities according to the emphasis of the viewpoint, but it seems that from the practical point of view, system B is preferable since it better catches the details of the desired statistical properties.

The strategy of the derivation of the SDEs described in section 5 is presented in a general form that allows to apply it to a wide spectrum of distributions of the mass flux of generated clouds with various dependences of the cloud lifetime on the cloud mass flux. Note, however, that in all cases examined in the present study
(Markovian PC08, intermediate and shallow convection schemes), the distribution of the new-born cloud lifetime was exponential, that is, memoryless, which implied the exact solution for \( N \) by the coarse-grained model A. It is difficult to say \textit{a priori}, how accurate the A and B solutions would be if lifetime distributions were not memoryless. It is possible, however, that the proposed system of SDEs (currently one equation per variable) may be extended by additional differential equations to make the system non-Markovian.

The presented process of the derivation of the SDEs was not always rigorous in the sense that sometimes right guesses played a role. It may also be that those right guesses are in fact strictly derivable statements, but this point remains unclear up to now. It may also be that the framework of birth-death processes is not a unique approach to the solution of the problem.

The proposed solutions are implemented as a research version into the shallow- and deep convection branches of the Bechtold-Tiedtke convection scheme (Bechtold et al., 2008; Tiedtke, 1989) within the ICON NWP model. The impact of the stochastic convection in this version on the global and regional weather prediction will be investigated in the future. The source code of the SDE solution is compact (∼20 lines per scheme) and may be obtained upon request from the authors.

Appendix A: Discussion on PC08 Assumptions

The starting points in the development of the PC08 scheme, as is stated in CC06, are four assumptions that (1) there is a separation between the spatial and temporal scales, on which convective forcing may be regarded as homogeneous, and the respective scales of a single cloud, (2) the invariance of the mean mass flux per cloud to the strength of forcing, (3) the independence of the mass fluxes of individual clouds of each other (mean-field approximation), and (4) equal \textit{a priori} probabilities of all possible mass flux distributions which are constrained only by large-scale conditions. In this appendix it will be shown that in fact some more independent assumptions are involved as the basics of the PC08 scheme.

Assumption 3 of the independence of the mass fluxes is used in CC06 to infer that the cloud mass fluxes are distributed exponentially. Next, we will show that this inference is not correct and thus an exponential distribution of the mass fluxes of individual clouds is an additional and independent assumption in the PC08 scheme.

It is stated in CC06 that Assumption 3 has a consequence that for each given mass flux \( M \) the number of clouds, contributing to it, is a Poisson-distributed random variable. To this end, an analogy with a Poisson process is invoked by considering a mass flux (\( M \)-) axis instead of a time axis (see Figure 1 in CC06 and Figure A1), so that the mass fluxes of individual clouds are represented as intervals on the \( M \) axis, divided by points (events in the terminology of the Poisson process theory). Then, the statement that the number of clouds for a given \( M \) is Poisson distributed is used then in CC06 to infer that the cloud mass fluxes are distributed exponentially. Here a well-known fact from the Poisson process theory is used that if on an arbitrary time interval (read: interval \((0, M]\)) the number of points falling into this interval is a Poisson-distributed random variable, the time intervals between adjacent points (read: mass fluxes of individual clouds) are distributed exponentially (Kingman, 1993; Taylor & Karlin, 1998).

Note, however, that a Poisson distribution of the number of clouds in an interval \((0, M]\) cannot be derived from the assumption on the independence of cloud mass fluxes of each other. To infer a Poisson distribution, not the intervals between the points dividing clouds on the \( M \) axis, that is, the mass fluxes of individual clouds, but the positions of these points must be independent, that is, sampled from a uniform distribution on \((0, M]\), which is not the same as the independence of the intervals between the points. Indeed, consider independent mass fluxes which are sampled from any but not exponential distribution, for instance, from a Gaussian distribution with a small variance. In this case mass fluxes are independent, but the points lay on the axis almost regularly and their number in \((0, M]\) is not Poisson distributed. Only the condition that the positions of points are completely random leads to a Poisson distribution of the number of points in some interval \((0, M]\). This in turn implies an exponential distribution of interval lengths. If, however, only the independence of intervals is assumed, as is done in CC06, nothing may be
said about the distribution of their lengths. They may be sampled from an arbitrary distribution and do not necessarily have an exponential one. There is one more trouble in using the analogy between clouds and events of a Poisson process, because the statement "the number of points falling into (0, M] is Poisson distributed," which is required by the theory, is not the same as "the number of clouds in a given M is Poisson distributed," since in the latter case the first and the last points should always coincide with the left and right boundaries of (0, M] for a given M, respectively, which is not possible due to the randomness of the mass fluxes of single clouds. This remark is, however, of the secondary importance, because the crucial point is that a Poisson distribution for the number of points in (0, M] does not follow from the independence of mass fluxes, as was shown above. In effect, one may say that an exponential distribution of the mass fluxes of individual clouds is additionally and independently assumed in the PC08 scheme. Despite that this is just an assumption, it is feasible, since it is justified by cloud-resolving modeling simulations (Cohen & Craig, 2006).

Further, the consideration of the distribution of points in some interval (0, M] says nothing about possible fluctuations of M, which is the final goal of the entire construction. This is because intervals and points were considered for some fixed M. Specifically, the assumption of an exponential distribution of intervals (cloud mass fluxes) leads to a Poisson distribution of the number of points dividing intervals in a given (0, M]. It does not help to specify the fluctuations of M. We dare to assume that in fact the flow of assumptions and inferences leading to the PC08 scheme was the following. First, the mean cloud number in a given grid box is found from both the given mean total mass flux and the mean mass flux per cloud as \( N = \langle M \rangle / \bar{m} \). Then, an assumption that \( N \) in this grid box is Poisson distributed with the mean \( \langle N \rangle \) is made. After that, an assumption of an exponential distribution of mass fluxes is adopted, so that \( N \) clouds with mass fluxes randomly sampled from this exponential distribution provide a fluctuating \( M \). Each of the two assumptions (a Poisson distribution of \( N \) and an exponential distribution of mass fluxes) is adopted independently; they do not follow from each other and from the independence of cloud mass fluxes.

Since the cloud lifetime is taken constant, the assumption of the Poisson-distributed cloud number in a grid box is equivalent to the assumption that clouds in the grid box appear according to a Poisson random process. The further assumption of the PC08 scheme, the equal a priori probabilities of mass flux distributions, represents another line of reasoning and seems actually not to be used in the PC08 scheme and, in particular, in the derivation of an exponential distribution of cloud mass fluxes as it is conducted in CC06. The analogy with a canonical ensemble for an ideal gas, with the resulting Boltzmann exponential distribution of molecule kinetic energies as the most probable, has served rather as a motivation for the PC08 scheme, as is stated in Appendix of CC06.

Acknowledgments
We thank George Craig for a fruitful discussion and Dmitrii Mironov, Judith Berner, one anonymous reviewer, and Associate Editor Peter Caldwell for a careful assessment of the manuscript and for providing many comments and suggestions that have helped to improve and clarify the paper.

References
Bechtold, P., Köhler, M., Jung, T., Doblas-Reyes, F., Leutbecher, M., Rodwell, M., et al. (2008). Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. Quarterly Journal of the Royal Meteorological Society, 134, 1337–1351.

Cohen, B. G., & Craig, G. C. (2006). Fluctuations in an equilibrium convective ensemble. Part II: Numerical experiments. Journal of the Atmospheric Sciences, 63, 2005–2015.

Craig, G. C., & Cohen, B. G. (2006). Fluctuations in an equilibrium convective ensemble. Part I: Theoretical formulation. Journal of the Atmospheric Sciences, 63, 1996–2004.

Daley, D. J., & Vere-Jones, D. (2003). An introduction to the theory of point processes (pp. 471). New York: Springer.

ECMWF (2011). Workshop on representing model uncertainty and error in numerical weather and climate prediction: Working Group Reports. https://www.ecmwf.int/en/learning/workshops/past-workshops

ECMWF/WWRP (2016). Workshop on model uncertainty: Working Group Reports. https://www.ecmwf.int/en/learning/workshops/past-workshops

Hasselmann, K. (1976). Stochastic climate models. Part I. Theory. Tellus, 28, 473–485.

Kain, J. S. (2004). The Kain-Fritsch convective parameterization: An update. Journal of Applied Meteorology, 43, 170–181.

Kain, J. S., & Fritsch, J. M. (1990). A one-dimensional entraining/detraining plume model and its application in convective parameterization. Journal of the Atmospheric Sciences, 47, 2784–2802.

Keane, R. J., Craig, G. C., Keil, C. h., & Zängl, G. (2014). The Plant-Craig stochastic convection scheme in ICON and its scale adaptivity. Journal of the Atmospheric Sciences, 71, 3404–3415.

Keane, R. J., & Plant, R. S. (2012). Large-scale length and time-scales for use with stochastic parameterization. Quarterly Journal of the Royal Meteorological Society, 138, 1150–1164.

Keane, R. J., Plant, R. S., & Tennant, W. J. (2016). Evaluation of the Plant-Craig stochastic convection scheme (v2.0) in the ensemble forecasting system MOGREPS-R (24 km) based on the Unified Model (v7.3). Geoscientific Model Development, 9, 1921–1935.

Kingman, J. F. C. (1993). Poisson processes (pp. 112). Oxford: Oxford Studies in Probability 3, Clarendon Press.
Plant, R. S., & Craig, G. C. (2008). A stochastic parameterization for deep convection based on equilibrium statistics. *Journal of the Atmospheric Sciences, 65*, 87–105.

Sakradžija, M., Seifert, A., & Dipankar, A. (2016). A stochastic scale-aware parameterization of shallow cumulus convection across the convective gray zone. *Journal of Advances in Modeling Earth Systems, 8*, 786–812. https://doi.org/10.1002/2016MS000634

Sakradžija, M., Seifert, A., & Heus, T. (2015). Fluctuations in a quasi-stationary shallow cumulus cloud ensemble. *Nonlinear Processes in Geophysics, 22*, 65–85.

Taylor, H. M., & Karlin, S. (1998). *An introduction to stochastic modeling* (pp. 631). San Diego: Academic Press.

Tiedtke, M. (1989). A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Monthly Weather Review, 117*, 1779–1800.