Coupling a minimal stochastic lattice gas model of a cloud system to an atmospheric general circulation model

Francesco Ragone,[a]* Klaus Fraedrich,[a,b] Hartmut Borth[a] and Frank Lunkeit[a]

[a] *Meteorologisches Institut, Universität Hamburg, Germany*
[b] *Max-Planck-Institut für Metereologie, Hamburg, Germany*

*Correspondence to: F. Ragone, Meteorologisches Institut, Universität Hamburg, Bundesstrasse 55, D-20146 Hamburg, Germany. E-mail: francesco.ragone@zmaw.de

We propose a strategy to couple a stochastic lattice gas model of a cloud system to a rather general class of convective parametrization schemes. As proposed in similar models recently presented in the literature, a cloud system in a grid box of a general circulation model (GCM) is modelled as a subgrid lattice of N elements that can be in one of S states, each corresponding to a different convective regime. The time evolution of each element of the lattice is represented as a Markov process characterized by transition rates dependent on large-scale fields and/or local interactions. In order to make application to GCMs computationally feasible, we propose a reduction method leading to a system of $S - 1$ stochastic differential equations with multiplicative noise. The accuracy of the reduction method is tested in a minimal version of the model. The coupling to a convective scheme is performed in such a way that, in the limit of space- and time-scale separation, the modified stochastic parametrization converges to the original deterministic version of the host scheme. Experiments with a real GCM are then performed, coupling the minimal version of the stochastic model to the Betts–Miller scheme in an aquaplanet version of the Planet Simulator. In this configuration, the stochastic extension of the parametrization keeps the climatology of its deterministic limit but strongly impacts the statistics of the extremes of daily convective precipitation.

Key Words: stochastic parametrization; atmospheric convection; lattice gas models

Received 8 February 2013; Revised 26 December 2013; Accepted 3 January 2014; Published online in Wiley Online Library 03 April 2014

1. Introduction

Due to their nonlinear nature, atmospheric processes at different space- and time-scales interact with each other. In a numerical model of the atmosphere, the evolution equations are discretized and therefore filtered in space and time by the size of the grid and time step respectively. Thus, it is necessary to represent the effect of the unresolved processes on the resolved scales. Under the assumption of the existence of a separation between the active unresolved and resolved scales, the unresolved processes can be considered to be in statistical equilibrium once averaged over the truncation scales and their mean effect can be represented at the zeroth order as a deterministic function of the resolved variables (parametrization).

For processes like organized convection, the scale separation assumption is not realistic for applications to most of the scales of interest, leading standard deterministic parametrizations to misrepresent some of the statistical properties of the system. In order to cure this problem, apart from drastic changes in the approach to the representation of the unresolved processes (superparametrization), first-order corrections can be represented with the inclusion of stochastic terms, the statistical properties of which will depend on the physics of the parametrized processes and their interaction with the resolved dynamics (stochastic parametrization).

In general, the idea of introducing stochastic terms into a climate model in order to represent variability due to fast, unresolved processes dates back to the seminal work of Hasselmann (1976). Since then, it has been applied to a number of geophysical models. Chekroun *et al.* (2011) have recently provided an introduction to random dynamical systems theory addressed to the geophysical community, showing how concepts of classical dynamical systems theory can be extended in order to provide deeper insights into the statistical properties of nonlinear stochastic–dynamical models. Most of the earlier studies were focused on simplified, low-dimensional descriptions of the climate system or specific climatic processes. Because of the increase in resolution of operational numerical models of weather and climate, the interest in introducing stochastic parametrizations in full general circulation models (GCMs) in order to represent subgrid variability due to unresolved processes (which could then feed back through the nonlinearities of the system, with potentially large impacts on the mean state and on the higher order statistics of the system) has gained momentum in recent years, with particular attention devoted to the parametrization of atmospheric convection (Neelin *et al.*, 2008; Palmer and Williams, 2010).

The representation of unresolved atmospheric convection in GCMs is still one of the crucial problems of climate modelling (Frank, 1983; Arakawa, 2004; Randall *et al.*, 2007). Many different parametrization schemes have been developed in the last decades. Classically, they are divided into three families: moisture budget schemes (Kuo, 1965, 1974), adjustment schemes (Manabe *et al.*, 1965; Betts and Miller, 1986) and mass-flux schemes developed in different versions by many authors (Arakawa and Schubert, 1974; Bougeault, 1985; Tiedtke, 1989; Gregory and Rowntree, 1990; Kain and Fritsch, 1990; Gregory, 1997; Bechtold *et al.*, 2001; Kain, 2004), although with some theoretical issues in some of these implementations (Plant, 2010). Despite being built starting from different points of view and physical considerations, all these schemes present similarities in their design and impact on the dynamics (Arakawa, 2004) and to some extent can be derived from a common approach (Fraedrich, 1973). The crucial common feature of all these schemes is that they realize a negative feedback, which efficiently dampens the vertical destabilization of the atmosphere due to radiation, advection and surface fluxes, in most cases by reducing a vertically integrated measure of the buoyancy at an exponential rate (Yano *et al.*, 2000).

This common property is introduced basically by design in the first and second family (from this point of view, the Kuo scheme can be interpreted as an adjustment scheme (Arakawa, 2004)), while in the third family it is realized by one of the many possible implementations of the quasi-equilibrium (QE) hypothesis originally introduced by Arakawa and Schubert (1974). The general definition of QE is basically equivalent to the existence of a time-scale separation between the large-scale dynamics and the convective activity, allowing the effects of convection to be parametrized as a response to the destabilization enslaved by the large-scale dynamics. It can therefore be easily extended to include the kind of justifications on which the adjustment (including Kuo) schemes are based and thus it can be considered for the sake of simplicity the conceptual basis of all the parametrization schemes available. Part of the current debate on the parametrization of convection is focused on giving a modern interpretation of the QE principle (Neelin *et al.*, 2008; Yano and Plant, 2012) at both fundamental and operational levels.

At a fundamental level, Mapes (1997) introduced the concept of activation control as opposite to the classical idea of equilibrium control (basically a rephrasing of QE), indicating that the latter is an adequate representation of the nature of tropical convection only on global climate scales. Yano *et al.* (2001, 2004) showed that the intermittent, pulse-like nature of tropical convective activity leads to the presence of $1/f$ spectra for characteristic quantities over a broad range of scales, so that the usual picture of QE as a smooth adjustment based on a time-scale separation is indeed questionable. A substantial body of observational works (Peters *et al.*, 2002, 2009, 2010; Peters and Neelin, 2006, 2009; Neelin *et al.*, 2009) showed that convection in the Tropics presents many features typical of systems undergoing a phase transition or in a state of criticality, leading the authors to propose the concept of self-organized criticality (SOC) to explain the transition to precipitating convection (Neelin *et al.*, 2008). Although the proposed framework supports some aspects of the QE idea, in the sense that a system featuring SOC indeed adjusts itself to the neighbourhood of a critical point, thus dampening deviations from it, the physical interpretation is radically different and, if valid, would imply that important statistical properties of convection are not captured by parametrizations based on classical formulations of the QE principle. Despite the debate on the validity of the QE hypothesis being ongoing for quite some time, no substantial improvements have been made so far in proposing a new conceptual framework robust enough to lead to the definition of a new generation of parametrization schemes.

Less fundamental criticisms address the practical implementation of the QE principle, noting that the concept holds strictly only in an ensemble average sense, integrating over an area hosting a large number of independent convective events and over a time interval larger than the typical length of their life cycle. That is, even supposing that a scale separation exists, the concept is practically useful only if the truncation scales of the GCM are indeed much larger than the characteristic scales of the parametrized process, although one could note that the way space-scale separation is involved in the QE principle is not as clear as time-scale separation (see Yano, 1999; Adams and Renno, 2003). The typical resolution of a state-of-the-art GCM for climatic applications is 100–200 km in space and 10–30 min in time. At these scales, only a few active convective elements (clouds) are present in a grid box and their life cycle (initiation, growth and dissipation) is far from being exhausted in such a short period of time. This has led some authors to suggest that convective parametrizations should at least consider first-order corrections to the classical theory in order to take into account finite size effects and aspects of the temporal evolution of the ensemble of convective events. The first issue has been basically ignored in the classical approaches to convective parametrization, while, in order to address the second issue, some prognostic schemes have been proposed in the literature (Randall and Pan, 1993; Randall *et al.*, 1997; Pan and Randall, 1998). In addition, it has to be noted that atmospheric convection shows features of spatial (Peters *et al.*, 2009) and temporal (Mapes *et al.*, 2006) organization, both of which are thought to be involved in determining properties of tropical variability from daily to intraseasonal scales.

These issues were the basis for the suggestion of introducing stochastic components into pre-existing convective parametrization schemes. First attempts to design a stochastic parametrization of atmospheric convection were basically sensitivity studies (Neelin *et al.*, 2008). Buizza *et al.* (1999) developed a perturbed physics scheme to take into account model uncertainties in the context of ensemble prediction. Lin and Neelin (2000, 2002) perturbed the heating term due to convective precipitation with an AR1 process in the Betts–Miller (BM) scheme and in a mass-flux scheme (Lin and Neelin, 2003), showing sensitivity of the tropical activity to the autocorrelation time of the noise. A different approach has been followed by Berner *et al.* (2005), who introduced a stochastic forcing to the stream function of a GCM with a spatial pattern given by a cellular automaton mimicking in a simple way the organization of mesoscale convective systems. Plant and Craig (2008) developed a stochastic parametrization scheme coupling the deterministic Kain–Fritsch scheme (Kain and Fritsch, 1990) to a probabilistic model for the distribution of the cloud-base mass flux based on equilibrium statistics (Craig and Cohen, 2006), which had shown good agreement with cloud-resolving models (Cohen and Craig, 2006). For a more comprehensive review on the topic see Neelin *et al.* (2008) and Palmer and Williams (2010).

Among these attempts, some attention has been devoted recently to using subgrid stochastic lattice gas models in order to describe the dynamics of a cloud population in a GCM grid box (Majda and Khouider, 2002; Khouider *et al.*, 2003, 2010; Frenkel *et al.*, 2012). A stochastic lattice gas model consists of a collection of N elements spatially organized following a certain geometry (for example, on a regular square lattice in which each site has four first neighbours), each of which can be in one of S states. The time evolution of each element on the set of the S states is determined by probabilistic rules dependent on the state of the element and its neighbours (in order to represent local interactions) and/or on external fields. In applications to convective parametrization, the N sites correspond to places in which convection may or not occur, while the S states correspond to different convective regimes or cloud types (Majda and Khouider, 2002; Khouider *et al.*, 2003, 2010; Frenkel *et al.*, 2012). Considering a lattice model nested in each grid box of a GCM, the stochastic model would determine the fraction of each cloud type in the grid box, thus modulating the amount of convective activity. In turn, the GCM would provide the large-scale fields determining the transition rates of the lattice model (e.g. CIN, CAPE, precipitable water), realizing in this way a full two-way coupling (Khouider *et al.*,

2010) between the small- and large-scale dynamics. The proposed models were devised in order to represent finite size effects and properties of the initiation and life cycle of tropical convection (Mapes *et al.*, 2006) and were coupled to simplified models of the tropical dynamics. Recently, Stechmann and Neelin (2011) have proposed a conceptual stochastic model for the transition to strong convection, suggesting that it could be used to inform the transition rules of similar models. On similar lines, Plant (2012) has proposed a general framework for using subgrid individual-level models (ILM) in the context of mass-flux parametrization, making use of the van Kampen system size expansion approach (van Kampen, 2007).

Despite some growing interest in this approach to stochastic parametrization of atmospheric convection, no attempts have been made so far regarding coupling to a full GCM. Here, we derive a systematic way to treat such models in the context of convective parametrization in real GCMs. Two ingredients are then needed: (i) a numerically treatable, and possibly to some extent analytically understandable, formulation of the evolution of the lattice model and (ii) a robust coupling strategy in order to include the stochastic model into a pre-existing parametrization. In particular, the first problem has already been tackled by previous works by means of a coarse-graining technique, reducing the model to a system of S birth–death stochastic processes, which are then simulated with the Gillespie method in Khouider *et al.* (2003, 2010) and Frenkel *et al.* (2012), while Plant (2012) has proposed using the van Kampen system size expansion approach. This is in general a common problem in population dynamics (McKane and Newman, 2004; Tome and de Oliveira, 2009; Black and McKane, 2012, and references therein), where some standard approaches have been developed that could also be used in the convection community. Here we propose a method partially on these lines that could be useful in order to extend this approach to the 'stochasticization' of a convective parametrization to real GCMs.

The first problem is due to the fact that, given the size of a GCM grid box and the individual convective elements, N is supposed to be quite large, so that a direct simulation of all N processes in each grid box would be numerically burdensome. Here we propose a reduction method leading to a system of $S-1$ stochastic differential equations for the evolution of the macrostate of the lattice model in the limit of large N. Note that N corresponds to the number of sites where convection can occur and not to the actual number of active convective elements or clouds, which in general will be a small fraction of it. The reduction method therefore does not need a large ensemble of clouds and makes it possible to deal with cases in which there are a few clouds in a grid box, provided that individually they are much smaller than the grid-box size. The intensity of the noise scales with $N^{-1/2}$, thus showing strong similarities with the van Kampen system size expansion, but our derivation is formulated in a way that could make the application to spatial models (i.e. models featuring spatial interactions) more intuitive and results to be equivalent in the final equations to strategies presented for specific spatial predatory–prey models (Tome and de Oliveira, 2009).

The second problem stems from the fact that many parametrization schemes do not feature the cloud or updraught fraction explicitly in their formulation, so that it is not always possible to introduce the cloud fraction provided by the stochastic model into these schemes automatically. Here we suggest a way to introduce the information given by the lattice model into a generic pre-existing deterministic parametrization scheme, in such a way that, in the limit of an infinite number of convective elements and transition rates much faster than the rate of change of the large-scale fields determining the transitions themselves (i.e. perfect space- and time-scale separation phrased for this particular description of unresolved convection), the modified stochastic parametrization converges to the standard deterministic version of the scheme. Our approach therefore consists of introducing first-order corrections to the QE-based implementation of a convective scheme and does not introduce a fundamental change of perspective regarding the basic philosophy of parametrizing convection (although, as discussed above, it is likely that something in this direction will be needed in the future).

According to the idea of defining a systematic way to introduce this kind of model in a GCM, we perform first sensitivity experiments in an environment as controlled as possible. We set up the lattice model in its minimal version, a binary system with only empty or occupied sites, no spatial interactions and transition rates independent of the state of the large-scale model. This version of the model will therefore introduce in the parametrization only the effects coming from considering a demographic description of the cloud system. In this simple version, the model is treatable not only numerically but also analytically, so that the role of its parameters in determining its statistical properties is known. The stochastic model is coupled to a simplified version of the BM scheme in the Planet Simulator (Fraedrich *et al.*, 2005; Fraedrich, 2012), an intermediate-complexity GCM with a full set of physical parametrizations. The model is set up in aquaplanet conditions with fixed sea-surface temperature (SST) and no seasonal or diurnal cycle, following Neale and Hoskins (2001), so that no time-dependent forcings are acting on the system. This idealized set-up, together with the good computational performances of the Planet Simulator, allows for a robust exploration of the parameter space of the stochastic model and a systematic investigation of its impact on the statistics of the large-scale model. Particular attention is devoted to the statistics of the extreme events of daily convective precipitation in the Tropics.

The article is organized as follows. In section 2 we derive our reduction method and analyze in detail the minimal version of the model. In section 3 we propose a strategy to couple the general version of the model to a generic parametrization scheme and discuss as a practical example the coupling of the minimal version of the model to the BM scheme. In section 4 we present the results of the idealized aquaplanet experiments performed with the Planet Simulator, in which the minimal model developed in section 2 is coupled to the BM parametrization scheme as described in section 3. In section 5 we eventually present conclusions and possible future lines of research.

2. Stochastic population dynamics on a subgrid lattice

Considering a grid-box size of order 100 km and typical sizes of convective elements ranging from 100 m–10 km (from individual cumulus clouds to mesoscale systems, depending on the definition of the model), we expect N to be in the range 10^6–10^2. Since the full evolution of the stochastic model would require casting an equivalent amount of random numbers at each time step for each grid box of the GCM, it is clear that a direct simulation of the system as the sum of all N individual processes would be impractical even in the best case. The problem has already been tackled in previous works by means of a coarse-graining technique, reducing the model to a system of S birth–death stochastic processes, which are then simulated with the Gillespie method (Khouider *et al.*, 2010; Frenkel *et al.*, 2012). With a different approach, Plant (2012) has applied the van Kampen system size expansion approach to a joint model of the number of clouds and the mass flux of the system, in the spirit of McKane and Newman (2004). In this article, we present an alternative method able to reduce the number of degrees of freedom of a stochastic lattice gas model, leading to a treatable system of a few stochastic differential equations.

2.1. Reduced stochastic model

Let us describe a cloud system as a collection of N elements or sites that can be in one of S states, each identifying a different convective regime or cloud type. For the sake of simplicity, we can consider them to be organized on a regular square lattice, even if this is not crucial for the results developed in the following. Let us represent the state of element n at time t with a S-dimensional vector $\boldsymbol{\sigma}^{nt}$, the components of which are the occupation numbers of the S states, i.e. $\sigma_s^{nt} = 1$ if element n is in state s at time t, $\sigma_s^{nt} = 0$ otherwise. The time evolution of each element can be described as in Khouider *et al.* (2010) as a Markov process characterized by transition rates $R_{ss'}^{nt}$, defined so that, for sufficiently small values of the time increment dt,

$$p_{ss'}^{nt}(\mathrm{d}t) = R_{ss'}^{nt}\,\mathrm{d}t, \tag{1}$$

where $p_{ss'}^{nt}(\mathrm{d}t)$ is the conditional probability of finding element n in state s at time $t + \mathrm{d}t$, given that it was in state s' at time t. It is practical to introduce the transition or intensity matrix $\boldsymbol{R}^{nt}$ for element n by defining

$$R_{ss}^{nt} = -\sum_{s' \neq s}^{S} R_{ss'}^{nt}, \tag{2}$$

so that the probability of remaining in the same state is by definition (from now on avoiding showing explicitly the dependence of the transition probabilities on dt)

$$p_{ss}^{nt} = 1 - \sum_{s' \neq s}^{S} p_{ss'}^{nt} = 1 - \sum_{s' \neq s}^{S} R_{ss'}^{nt}\mathrm{d}t = 1 + R_{ss}^{nt}\,\mathrm{d}t. \tag{3}$$

The transition matrix has been defined consistently with the convention of right-hand matrix multiplication for the evolution of the Markov process, so that the vector $\mathbf{p}^{nt}$, the components of which are the absolute probabilities of finding the element n in state s at time t, evolves according to the master equation

$$\frac{\mathrm{d}\mathbf{p}^{nt}}{\mathrm{d}t} = \boldsymbol{R}^{nt}\mathbf{p}^{nt}. \tag{4}$$

As suggested in previous works (Majda and Khouider, 2002; Khouider *et al.*, 2010; Stechmann and Neelin, 2011; Plant, 2012), the coefficients of the transition matrix will depend in general on some large-scale fields, like CIN, CAPE, measures of dryness of the atmospheric column and/or precipitable water. This would reflect the influence of the large-scale conditions on the probability of activating convection: for example, we expect the probability of having deep precipitating convection to increase with larger CAPE and vice versa. The dependence of the transition rates on these fields will therefore be the same for each element of the lattice. In addition, we could imagine that the transition rates for each element will also depend on the state of its neighbourhood, due to the fact that the existence of a cloud at a certain point of the lattice will activate processes influencing the probability of having other clouds in the area nearby, in either a cooperative or a competitive sense. Clustering effects are indeed observed in studies of cumulus cloud life cycles, as a consequence of mesoscale processes leading to positive near-neighbour feedbacks (Mapes, 1993; Redelsperger *et al.*, 2000; Tompkins, 2001; Houze, 2004; Moncrieff and Liu, 2006, and references therein). Let us in general write the transition rate matrix for element n as

$$\boldsymbol{R}^{nt}\left(\mathbf{x}^t, \boldsymbol{\sigma}^t\right) = \boldsymbol{F}^t\left(\mathbf{x}^t\right) + \boldsymbol{J}^{nt}\left(\boldsymbol{\sigma}^{n't} \in \Lambda^n\right), \tag{5}$$

where $\boldsymbol{F}^t\left(\mathbf{x}^t\right)$ is the same for each element of the lattice and represents the effect of the large-scale conditions defined by the GCM state vector $\mathbf{x}^t$, while $\boldsymbol{J}^{nt}\left(\boldsymbol{\sigma}^{n't} \in \Lambda^n\right)$ represents the effects of possible interactions between the element n and the elements of its neighbourhood Λ^n, the range of which will depend on the nature of the processes involved. Let us now define the cloud area fraction vector as (Khouider *et al.*, 2010; Stechmann and Neelin, 2011)

$$\boldsymbol{\sigma}^t = \frac{1}{N}\sum_{n=1}^{N} \boldsymbol{\sigma}^{nt}. \tag{6}$$

We would like to have an evolution equation for process $\boldsymbol{\sigma}^t$ that does not involve the computation of all individual processes $\boldsymbol{\sigma}^{nt}$.

This is a common problem in population dynamics, and some common strategies are known in order to solve it (McKane and Newman, 2004; Tome and de Oliveira, 2009; Black and McKane, 2012). The simplest approach is to neglect the details of the spatial configuration of the lattice by taking an approximation, similar to what is done in the simple (deterministic) mean-field approximation (Tome and de Oliveira, 2009). Subsequently one can apply an expansion method for large N (implying a certain degree of space scale separation) like the one of the van Kampen approach (van Kampen, 2007) in order to obtain a first-order correction to the deterministic mean-field equations in the form of a set of standard stochastic differential equations (Tome and de Oliveira, 2009). The method presented here differs in the second step but recovers the same final result and can be considered equivalent to the one of Tome and de Oliveira (2009).

The mean-field approximation is a standard tool in statistical mechanics and population dynamics and is based on the assumption that, as long as we are interested in the properties of a macroscopic quantity, the contributions due to the correlations between the individual processes can be neglected, provided that we can replace each local interaction term with a mean-field term (constant over the lattice) that takes into account the collective contribution of all the interactions. In practical terms, it consists of neglecting the correlations between the individual processes

$$\langle \sigma_s^{nt}\sigma_{s'}^{n't}\rangle \approx \langle \sigma_s^{nt}\rangle\langle \sigma_{s'}^{n't}\rangle \quad \forall s, s', \qquad n \neq n', \tag{7}$$

and of substituting in the interaction terms the state of the individual processes $\boldsymbol{\sigma}^{nt}$ with the average value over the lattice $\boldsymbol{\sigma}^t$, defining in this way a new mean-field interaction term $\boldsymbol{J}^t\left(\boldsymbol{\sigma}^t\right)$ that replaces each $\boldsymbol{J}^{nt}\left(\boldsymbol{\sigma}^{n't} \in \Lambda^n\right)$. In this way, we obtain a mean-field transition matrix

$$\boldsymbol{R}^t\left(\mathbf{x}^t, \boldsymbol{\sigma}^t\right) = \boldsymbol{F}^t\left(\mathbf{x}^t\right) + \boldsymbol{J}^t\left(\boldsymbol{\sigma}^t\right) \tag{8}$$

valid for each element of the lattice. In the thermodynamic limit of infinite N, this allows us to represent the system with a set of deterministic differential equations, known as the mean-field equations of the system (McKane and Newman, 2004). The mean-field approximation yields, of course, exact results in the trivial case of a system of non-interacting elements (where it is not an approximation at all); in general its applicability will depend on the specific form of the interaction terms.

However, it is possible to derive a stochastic differential equation for the time evolution of the cloud area fraction $\boldsymbol{\sigma}^t$ representing a first-order correction to the mean-field equations. Let us suppose that we know the state of the system at time t. Given the Markovian nature of the model, the statistical properties of the increment $\mathrm{d}\boldsymbol{\sigma}^t$ from time t to time $t + \mathrm{d}t$ are then uniquely determined. Since in the mean-field approximation the system has been approximated as a collection of uncorrelated Markov processes, each with the same transition matrix $\boldsymbol{R}^t$, it is reasonable to approximate the process $\mathrm{d}\boldsymbol{\sigma}^t$ with a Gaussian process if N is sufficiently large, with fluctuations scaled by a factor $N^{-1/2}$. For example, a similar central limit argument for the properties of the first-order fluctuations is taken in the van Kampen method for expanding the master equation of the system. Taking this approximation, the process $\mathrm{d}\boldsymbol{\sigma}^t$ is completely described by its expectation value $\langle \mathrm{d}\boldsymbol{\sigma}^t\rangle$ and its covariance matrix $\boldsymbol{C}^t$ and can be written at each t as

$$\mathrm{d}\boldsymbol{\sigma}^t = \langle \mathrm{d}\boldsymbol{\sigma}^t\rangle + \epsilon_N \boldsymbol{\eta}^t, \tag{9}$$

where $\epsilon_N = N^{-1/2}$ and $\boldsymbol{\eta}^t$ is a Gaussian random vector with zero expectation value and covariance matrix $\boldsymbol{C}^t$. Again making use of the mean-field approximation, one can show (see the Appendix at the end of the article) that both $\langle \mathrm{d}\boldsymbol{\sigma}^t \rangle$ and $\boldsymbol{C}^t$ scale with dt and can be expressed as functions of $\boldsymbol{R}^t$ and $\boldsymbol{\sigma}^t$, namely $\langle \mathrm{d}\boldsymbol{\sigma}^t \rangle = \boldsymbol{R}^t \boldsymbol{\sigma}^t \mathrm{d}t$ and $\boldsymbol{C}^t = \boldsymbol{D}^t \mathrm{d}t$, where

$$D^t_{ss'} = -R^t_{s's}\sigma^t_s - R^t_{ss'}\sigma^t_{s'} + \delta_{ss'} \sum_{i=1}^{S} R^t_{si}\sigma^t_i. \quad (10)$$

The process $\boldsymbol{\eta}^t$ can therefore be written as

$$\boldsymbol{\eta}^t = \boldsymbol{G}^t \mathrm{d}\boldsymbol{W}^t, \quad (11)$$

where $\mathrm{d}\boldsymbol{W}^t = \boldsymbol{W}^{t+\mathrm{d}t} - \boldsymbol{W}^t$ is the increment of a standard multivariate Wiener process $\boldsymbol{W}^t$, so that it is a multivariate Gaussian process with covariance matrix $\boldsymbol{I}\,\mathrm{d}t$ and $\boldsymbol{G}^t$ is a suitable matrix transforming the covariance matrix of the process in $\boldsymbol{C}^t = \boldsymbol{D}^t\,\mathrm{d}t$. From simple linear algebra considerations, we find that it is sufficient to define $\boldsymbol{G}^t = \boldsymbol{E}^t\sqrt{\boldsymbol{\Lambda}^t}$, where $\boldsymbol{E}^t$ and $\boldsymbol{\Lambda}^t$ refer to the diagonal representation $\boldsymbol{D}^t\boldsymbol{E}^t = \boldsymbol{E}^t\boldsymbol{\Lambda}^t$. We can therefore write the process $\mathrm{d}\boldsymbol{\sigma}^t$ as

$$\mathrm{d}\boldsymbol{\sigma}^t = \boldsymbol{R}^t\boldsymbol{\sigma}^t\,\mathrm{d}t + \epsilon_N \boldsymbol{G}^t\,\mathrm{d}\boldsymbol{W}^t. \quad (12)$$

Equation (12) is the standard form of a stochastic differential equation for the time evolution of the process $\boldsymbol{\sigma}^t$. Formally dividing by dt, we can write it in the alternative form

$$\frac{\mathrm{d}\boldsymbol{\sigma}^t}{\mathrm{d}t} = \boldsymbol{R}^t\boldsymbol{\sigma}^t + \epsilon_N \boldsymbol{G}^t\boldsymbol{\xi}^t, \quad (13)$$

where $\boldsymbol{\xi}^t = \mathrm{d}\boldsymbol{W}^t/\mathrm{d}t$ is the usual Gaussian white noise with $\langle \boldsymbol{\xi}^t \rangle = 0$ and lagged covariance matrix $\langle \boldsymbol{\xi}^t \boldsymbol{\xi}^{t'} \rangle = \delta(t-t')\boldsymbol{I}$. Equation (13) consists of a system of S scalar stochastic differential equations for the fraction of each cloud type, which are reduced to $S-1$ independent equations because of the constraint $\sum_{s=1}^{S} \sigma^t_s = 1$.

The Fokker–Planck equation related to Eq. (13), the solution of which is the probability distribution function ρ^t of the process $\boldsymbol{\sigma}^t$, is

$$\begin{aligned}\frac{\partial \rho^t}{\partial t} = &- \sum_{s=1}^{S} \frac{\partial}{\partial \sigma_s}\left(\left[\boldsymbol{R}^t\boldsymbol{\sigma}^t\right]_s \rho^t\right) \\ &+ \frac{\epsilon_N^2}{2}\sum_{s=1}^{S}\sum_{s'=1}^{S}\frac{\partial^2}{\partial\sigma_s\partial\sigma_{s'}}\left(\boldsymbol{D}^t_{ss'}\rho^t\right).\end{aligned} \quad (14)$$

The Fokker–Planck equation is totally equivalent to the one obtained by Tome and de Oliveira (2009) for a specific spatial predator–prey model, applying the van Kampen expansion to the master equation of the system in order to build a first-order stochastic correction of the mean-field description of the system. Our approach (based on the crucial approximation (9)) gives the same final equations, but is presented here in a more general form and in a way that could be more intuitive for applications outside the population dynamics community.

We have therefore derived an equation for the time evolution of the macrostate of the lattice model for large N and interactions suitable for describing in the mean-field approximation. In a typical convective parametrization scheme, only a few cloud types are considered, typically non-precipitating shallow convection and precipitating deep convection. The computational burden of the numerical integration of Eq. (13) is therefore several orders of magnitudes smaller than the direct simulation of the lattice model, which makes possible the inclusion of a real GCM in the convective parametrization.

Comparing with the methods already proposed in the literature, in terms of computational cost our method is equivalent to the coarse-graining technique of Khouider *et al.* (2003, 2010). Apart from this, the formulations of the two methods are rather different. It has to be noted that the coarse-graining technique of Khouider *et al.* (2003, 2010) does not require N to be large (while still being able to represent local interactions, as in Khouider *et al.* (2003)). Our method instead does require N to be large (as in Plant, 2012), and therefore requires a certain degree of space-scale separation that is not necessary in Khouider *et al.* (2003, 2010). Moreover, the method of Khouider *et al.* (2003, 2010) conserves the Hamiltonian dynamics of the lattice model, which is very advantageous if the dependence of the transition rates on large-scale fields and local interactions is formulated in that framework (Khouider *et al.*, 2003). However, while the method of Khouider *et al.* (2003, 2010) results in a set of probabilistic rules for the evolution of the process not represented in an analytical form, our method has the attractiveness of providing a set of explicit stochastic differential equations (SDE) for the cloud fractions. This is quite an attractive feature, since it makes it possible to derive some general properties of the process from the form of Eq. (14), which may be useful in understanding to some extent the impact it could have on the large-scale dynamics and in performing experiments in a controlled and systematic way.

2.2. *Minimal version of the model: binary system with fixed transition rates*

Let us consider the minimal case of a two-state system, so that $\boldsymbol{\sigma}^t = (\sigma_1(t), \sigma_2(t))$, with fixed transition rates, without any dependence on external fields and in the absence of local interactions. We can interpret it as a model for an on/off description of convection, with $\sigma_1(t)$ and $\sigma_2(t)$ representing respectively sites that are convectively inactive (clear sky) and active (clouds). In the perspective of applications to a real convective parametrization, the assumption of constant transition rates is clearly unrealistic, as in general we expect the birth and death rates of deep convective clouds to depend on the state of the atmospheric column. This simplification is, however, attractive in a first, explorative phase of the study of the impact of this kind of model on a convective parametrization, since it introduces only the effects coming from a demographic description of the cloud system and, as is shown in the following, it leads to a fully analytically treatable form of the SDE, which will be useful in order to perform experiments in a controlled way. More realistic cases attempting to model the relationship between the state of the atmospheric column and the onset of deep convection will be the target of future works.

Under these assumptions, the transition matrix will in general be (from now on dropping the time dependence in the notation)

$$\boldsymbol{R} = \begin{pmatrix} -b & d \\ b & -d \end{pmatrix}, \quad (15)$$

where b and d are, respectively, the (constant) birth and death rate of the clouds or convective plumes quantified by σ_2 (conversely for σ_1). In this specific, case the matrix $\boldsymbol{D}$ defined by Eq. (10) becomes

$$\boldsymbol{D} = D\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad (16)$$

where $D = \sigma_1 b + \sigma_2 d$. The diagonalization yields

$$\begin{aligned}\boldsymbol{E} &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \\ \boldsymbol{\Lambda} &= D\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}, \\ \boldsymbol{G} &= \sqrt{D}\begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}.\end{aligned} \quad (17)$$

The resulting evolution equation is then

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\sigma_1\\ \sigma_2\end{pmatrix} = \begin{pmatrix}-b & d\\ b & -d\end{pmatrix}\begin{pmatrix}\sigma_1\\ \sigma_2\end{pmatrix} + \epsilon_N\sqrt{D}\begin{pmatrix}0 & 1\\ 0 & -1\end{pmatrix}\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}, \tag{18}$$

where ξ_1 and ξ_2 are independent Gaussian white noises. It is easy to see that Eq. (18) satisfies the condition

$$\frac{\mathrm{d}}{\mathrm{d}t}(\sigma_1 + \sigma_2) = 0, \tag{19}$$

so that, setting $\sigma_1(0) + \sigma_2(0) = 1$, the constraint will be satisfied at any following t. This comes from the structure of the matrix $\boldsymbol{G}$, specifically from the fact that the first column contains only zeros and therefore ξ_1 is not involved in the equations for σ_1 and σ_2 and that the the second column is such that ξ_2 contribute to the equations for σ_1 and σ_2 with the same magnitude and opposite sign. Simplifying the notation to $\sigma_1 = 1 - \sigma$, $\sigma_2 = \sigma$, $\xi_2 = \xi$, the system is fully described by the evolution equation for the cloud area fraction σ, i.e.

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = (1-\sigma)b - \sigma d + \epsilon_N\sqrt{(1-\sigma)b + \sigma d}\,\xi. \tag{20}$$

The system can therefore be modelled by a single stochastic differential equation with a linear deterministic drift term and a multiplicative stochastic forcing, the intensity of which depends on the size of the system.

In order to understand the properties of the solution of the model, it is more convenient to express Eq. (20) in terms of the parameters (σ_0, τ), where

$$\sigma_0 = \frac{b}{b+d}, \tag{21}$$

$$\tau = \frac{1}{b+d}. \tag{22}$$

In this way, we have

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = \frac{\sigma_0 - \sigma}{\tau} + \epsilon_N\sqrt{\frac{\sigma_0 + \sigma(1-2\sigma_0)}{\tau}}\,\xi. \tag{23}$$

Defining the scaled drift and diffusion coefficients $D_1' = \tau D_1 = \sigma_0 - \sigma$ and $2D_2' = 2\tau D_2 = \epsilon_N^2[\sigma_0 + (1-2\sigma_0)\sigma]$, the Fokker–Planck equation associated with Eq. (23) can be written as

$$\tau\frac{\partial\rho}{\partial t} = \frac{\partial\rho D_1'}{\partial\sigma} + \frac{\partial^2\rho D_2'}{\partial\sigma^2}. \tag{24}$$

Since the scaled coefficients D_1' and D_2' do not depend on τ, the stationary solution ρ_s of Eq. (24), which corresponds to the equilibrium probability distribution of the process σ, also does not depend on τ, but only on σ_0 and ϵ_N.

Following Risken (1989), if the variable σ has a lower bound (in our case 0), ρ_s is given by

$$\rho_s \propto \frac{1}{D_2'}\exp\left(\int_0^{\sigma}\frac{D_1'}{D_2'}\,\mathrm{d}\sigma\right), \tag{25}$$

where the proportionality constant is set by the normalization condition. In our case, this results in

$$\rho_s \propto [\sigma_0 + (1-2\sigma_0)\sigma]^{\frac{4}{\epsilon_N^2}\frac{\sigma_0(1-\sigma_0)}{(1-2\sigma_0)^2}-1}\exp\left(-\frac{2}{\epsilon_N^2}\frac{\sigma}{1-2\sigma_0}\right). \tag{26}$$

We see that ρ_s is strongly non-Gaussian. In particular, the upper tail is exponential, so that large deviations from the equilibrium value σ_0 induced by the demographic noise are (relatively) likely to be observed.

On the other hand, τ is the characteristic time-scale of the process σ, since it disappears from the equations if taken as unity by a rescaling of time, and it uniquely describes the memory of the process: the (non-normalized) autocorrelation function of the process σ is

$$r(t) = \langle\sigma^{t_0}\sigma^{t_0+t}\rangle - \langle\sigma^{t_0}\rangle\langle\sigma^{t_0+t}\rangle, \tag{27}$$

where t_0 can take any value. Taking the time derivative of Eq. (27) and using Eq. (20), we have

$$\frac{\mathrm{d}r}{\mathrm{d}t} = -\frac{r}{\tau}. \tag{28}$$

Therefore, the autocorrelation function of a specific realization of the process is an exponential decay on the time-scale τ, basically because of the linearity of the deterministic drift term in Eq. (20). The process σ is therefore memoryless and has a white spectrum. Note that, introducing interactions among the lattice elements, the transition rates become functions of σ and the deterministic drift term nonlinear, thus leading to (possibly) more complicated memory properties, even in the absence of time-dependent external fields.

2.3. *Numerical test*

To evaluate the accuracy of the reduction method, we take as a test case the binary system with time-independent transition rates described in the previous section. We compare, for different values of the parameters σ_0, τ and N, the stationary distributions and correlation functions resulting from the direct simulation of the full lattice model (DS) and the iteration of the correspondent SDE with two approaches (M1 and M2, see below), as well as the expected theoretical results derived in the previous section. We perform 64 simulations with the following values of the parameters: $\sigma_0 = (0.001, 0.01, 0.05, 0.1)$, $\tau = (3, 6, 12, 24)$ h and $N = (100, 225, 400, 1000)$. These values are compatible with application to the description of a cloud system inside a GCM grid box in the Tropics, where the typical value of the cloud fraction is supposed to be small, typical time-scales of evolution of convective events are of the order of a few hours and typical horizontal length-scales of the individual cumulus clouds are of the order of a few km. Considering a coarse, $T42$ resolution, equivalent to tropical grid-box linear dimensions of about 300 km, the selected values of N correspond to linear sizes of the convective elements in the range 10–30 km. These are large numbers for observed convection in the Tropics, but not totally unreasonable. Both the direct simulations of the lattice model and the iterations of the SDE are performed for a time period $T = 3$ yr with time step $\Delta t = 15$ min, starting from the initial condition $\sigma = \sigma_0$, so that no transient behaviour has to be taken into account. The SDE is integrated using the equivalent of the first-order Eulerian integration scheme for SDEs.

For small values of σ_0 and N, it is likely that a fluctuation of the noise term could lead to a negative value of σ during the iteration of the SDE. Those are cases in which we are at the edge of the applicability of the reduction method. We can avoid negative values of σ in two possible ways. In the first set of simulations (M1), we set $\sigma = 0$ every time the iteration of the SDE would lead to a negative value. In the second set of simulations (M2), we recast the random number every time a fluctuation would lead to a negative value, until $\sigma \geq 0$. The first method has the disadvantage of artificially increasing the probability of $\sigma = 0$, while the second method has the disadvantage of increasing the computational time required to iterate the SDE. The same procedure is applied in order to avoid values larger than 1.

Figure 1 shows how the stationary distribution changes for different values of N, keeping $\sigma_0 = 0.05$ and $\tau = 3$ h, with the

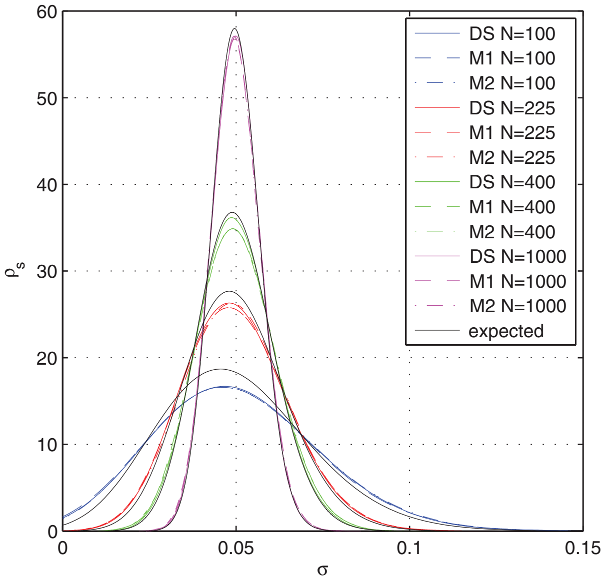

Figure 1. Stationary distributions of the process σ for $\sigma_0 = 0.05$, $\tau = 3$ h and different values of N, as the result of the direct simulation of the lattice model DS (solid line) and of the iteration of the equivalent SDE with two different algorithms M1 and M2 (dashed and dot–dashed lines). All the runs are performed for a time period $T = 3$ years with time step $\Delta t = 15$ min, starting from the initial condition $\sigma = \sigma_0$, so that no transient behaviour has to be taken into account. The distributions have been filtered with a window of length $1/N$. The black solid lines represent the expected behaviour following the solutions of the corresponding Fokker–Planck equation.

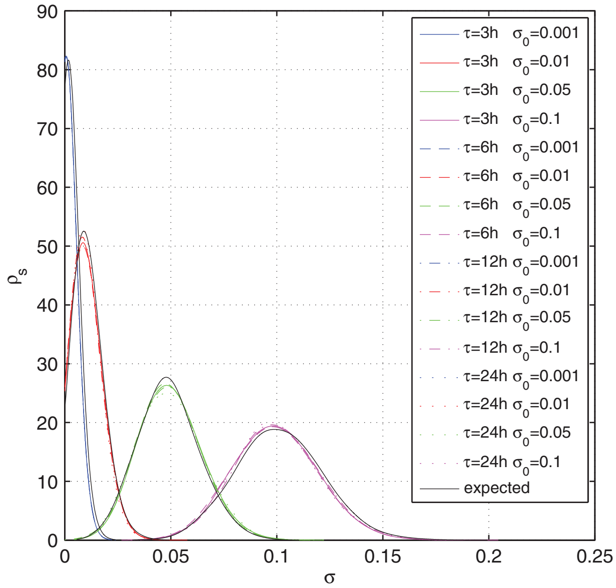

Figure 2. Stationary distributions of the process σ for $N = 225$ and different values of σ_0 and τ, as the result of the direct simulation of the lattice model. All the runs are performed for a time period $T = 3$ years with time step $\Delta t = 15$ min, starting from the initial condition $\sigma = \sigma_0$, so that no transient behaviour has to be taken into account. The distributions have been filtered with a window of length $1/N$. The black solid lines represent the expected behaviour following the solutions of the corresponding Fokker–Planck equation.

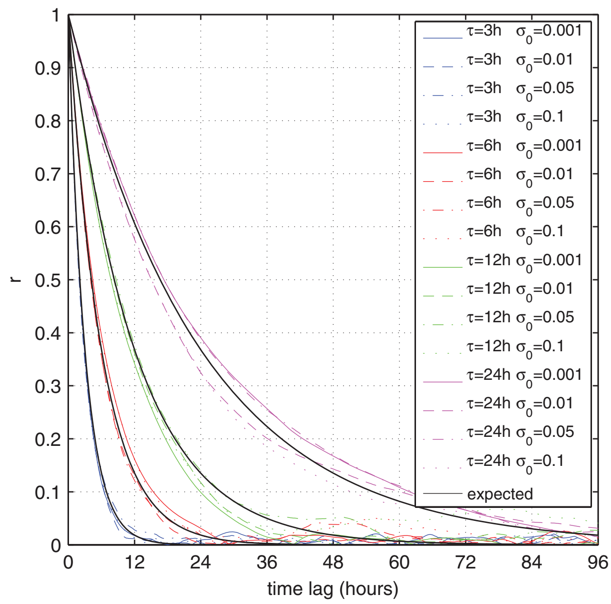

Figure 3. Autocorrelation functions of the process σ for $N = 225$ and different values of σ_0 and τ, as the result of the direct simulation of the lattice model. All the runs are performed for a time period $T = 3$ years with time step $\Delta t = 15$ min, starting from the initial condition $\sigma = \sigma_0$, so that no transient behaviour has to be taken into account. The distributions have been filtered with a window of length $1/N$. The black solid lines represent the expected behaviour following the solutions of the corresponding Fokker–Planck equation.

direct simulation of the lattice model DS and the iteration of the SDE with methods M1 and M2 (solid, dashed and dash–dotted lines respectively). The distributions have been filtered with a window of width $1/N$ (see below). In black, we have the expected behaviour from the solution of the corresponding Fokker–Planck equation.

In general, an important difference between the lattice model and the SDE reduction is that in the first case σ belongs to a discrete domain (the integer multiples of $1/N$), while in the second case it belongs to a continuous domain. This difference is particularly strong when the system hosts few active elements, i.e. for small values of σ_0 and N. Still, if the distributions are filtered by a factor $1/N$ we can see that the curve for the DS experiment collapses on the curves for the M1 and M2 experiments (which are almost indistinguishable from each other), as well as on the theoretical one. The reduced model is therefore able to represent the statistical properties of the lattice model, apart from the digital nature of the signal in the full simulation. This deficiency is in our opinion not of major concern in practical applications, the reason being that this digitalization of the cloud fraction is indeed an artefact of the representation of the cloud system as a regular lattice model, in which all the convective elements have exactly the same size ($1/N$ in units of grid-box area). Instead, in a real cloud system each element will have a different size. This does not mean that the continuous version of the signal obtained with the SDE is closer to reality, but at least the missing feature is not a real physical property of the system. We can therefore consider ourselves satisfied with the numerical performance of our reduction method. We can also see that no appreciable differences are present between methods M1 and M2, so that the faster method M1 can be considered as our best candidate for application in a GCM.

As expected, on increasing the size of the system, the range of fluctuations around σ_0 is reduced and vice versa. When the distribution is broad enough to interact with the lower boundary, the agreement between theory and numerical results becomes worse, but is still acceptable for the range of values considered here. For some reason, in these cases the iteration of the SDE (with both M1 and M2) follows the DS experiment (which is the 'true' system we want to model) better than the theory, so that this disagreement is of even less concern.

Figures 2 and 3 show the stationary distributions and correlation functions obtained from the DS experiment and the theory (the results from M1 and M2 do not give additional information and are not shown), varying σ_0 and τ and keeping $N = 225$. We can see that the full lattice mode shows the properties highlighted in the previous section: the stationary distributions for different values of τ collapse on each other and follow the expected form, as do the autocorrelation functions for different values of σ_0. In this range of values, changing σ_0 both shifts the centre of the distribution and modifies its width, with larger fluctuations for larger σ_0.

These results show that our stochastic model is a good representation of the original system (given the issues discussed above) and can therefore be used in order to 'stochasticize' a convective parametrization with a minimal demographic description of a cloud system. There are no theoretical reasons to expect a worse numerical accuracy when considering cases with more than two possible states and/or time-dependent transition rates. On the contrary, the applicability and numerical accuracy of the reduction method in the presence of interactions will depend on the degree of applicability of the mean-field approximation and will have to be considered case by case. In this case, of course, the mean-field approximation plays no role, since there are no interactions at all in the original system. Still, this simple model, which stems simply from the fact of 'counting' the clouds (hence the term 'demographic') and considering fixed characteristic time-scales for their birth and death rates presents non-trivial statistics, which could already have an interesting impact when introduced in the convective parametrization of a GCM.

3. Coupling strategy

Once the geometry of the system and the nature of the transitions are defined, Eq. (13) determines the cloud fraction of each cloud type considered. The question is now how the cloud fractions should enter into the description of unresolved convection in a GCM. Since the aim of this article is partly to define a strategy as general as possible to introduce the kind of model described in the previous section in a GCM, we would like to define a coupling strategy that is, as much as possible, independent of the specific parametrization scheme used in the host GCM. We also have to take into account the fact that many convective parametrization schemes do not explicitly include the cloud or updraught fraction in their formulations and therefore it is not possible in general to substitute it directly with the cloud fraction given by the stochastic model. In addition, we like to define a controlled environment for testing the introduction of the stochastic model, in the sense that we would like the modified parametrization to conserve its skills in representing the bulk statistics of the convective activity, while affecting mainly higher order properties.

3.1. *Stochastic extension of host deterministic parametrization*

We consider the general case described in section 2.1, in which the stochastic model can have an arbitrary number of states and the transition rates can depend on the large-scale conditions and on local interactions. Again, let $\mathbf{x}$ be the state vector of the resolved variables of a GCM (we do not show the time dependence explicitly from now on). We can represent its time evolution in general as $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \hat{\boldsymbol{\alpha}})$, where $\hat{\boldsymbol{\alpha}}$ is a vector containing the parameters of the parametrizations of unresolved processes present in the system.

The idea is that the stochastic model modifies the value of some relevant parameters, so that in general

$$\begin{cases} \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}, \boldsymbol{\alpha}(\boldsymbol{\sigma})), \\ \frac{\mathrm{d}\boldsymbol{\sigma}}{\mathrm{d}t} = \boldsymbol{R}(\mathbf{x}, \boldsymbol{\sigma})\boldsymbol{\sigma} + \epsilon_N \boldsymbol{G}(\mathbf{x}, \boldsymbol{\sigma})\boldsymbol{\xi}. \end{cases} \tag{29}$$

If the size of the grid box is much larger than the size of the individual convective elements ($N \to \infty$, space-scale separation), $\epsilon_N \to 0$. If the length of the time step of the GCM is much larger than the largest characteristic time-scale of the transitions ($\Delta t \min(R_{ij}) \to \infty$, time-scale separation), $\mathrm{d}\boldsymbol{\sigma}/\mathrm{d}t \to 0$. For these conditions, the equation of the stochastic model reduces to $\boldsymbol{R}\boldsymbol{\sigma} = \mathbf{0}$. In the case in which $\boldsymbol{R}$ does not depend on $\boldsymbol{\sigma}$ then this is a linear system (where $\mathbf{x}$ takes the role of a fixed set of parameters). Assuming that the matrix $\boldsymbol{R}$ is ergodic, so that $\det(\boldsymbol{R}) = 0$, the system always has one and only one solution different from the trivial one. Recalling Eq. (4), we see that this solution $\hat{\boldsymbol{\sigma}}$ is the invariant distribution of the Markov process defined by the mean-field transition matrix. In the case of the two-state system described in the previous section, the solution is $\hat{\boldsymbol{\sigma}} = (1 - \sigma_0, \sigma_0)$. In the case in which $\boldsymbol{R}$ depends on $\boldsymbol{\sigma}$, the equation for $\hat{\boldsymbol{\sigma}}$ becomes nonlinear, thus possibly leading to more complicated stationary solutions, such as multiple fixed points or limit cycles.

Avoiding these problematic cases, assuming the existence of a single fixed point and in the limit of space- and time-scale separation, the system becomes

$$\begin{cases} \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}, \boldsymbol{\alpha}(\hat{\boldsymbol{\sigma}})), \\ \boldsymbol{R}(\mathbf{x}, \hat{\boldsymbol{\sigma}})\hat{\boldsymbol{\sigma}} = \mathbf{0}. \end{cases} \tag{30}$$

Since the original deterministic version of the parametrization is supposed to be designed exactly in the case of perfect space- and time-scale separation, we will require that in this limit the modified stochastic parametrization converges to the original version of the scheme. This can be obtained by defining the functional form of the dependence of the parameters by the state of the stochastic model, so that $\boldsymbol{\alpha}(\hat{\boldsymbol{\sigma}}) = \hat{\boldsymbol{\alpha}}$. In this way, deviations from the fixed point of the deterministic limit of the stochastic model will correspond to first-order corrections to the original, deterministic version of the host scheme already implemented and tested.

As stated before, this strategy works only when it is possible to identify a single stationary solution for the deterministic limit of the stochastic model. When this is not the case, a different coupling strategy would be needed. In contrast, the relation $\boldsymbol{\alpha}(\hat{\boldsymbol{\sigma}}) = \hat{\boldsymbol{\alpha}}$ is also well-defined when $\boldsymbol{R}$ depends on time through the dependence on $\mathbf{x}$, as we will simply have different values of $\hat{\boldsymbol{\alpha}}$ at different times.

Note that this coupling strategy essentially results in a more complex random-parameter approach. In more simple-minded random-parameter approaches (Lin and Neelin, 2000; Bright and Mullen, 2002; Bowler *et al.*, 2008), a parameter is represented as a first-order autoregression process, with prescribed mean value and autocorrelation time (and often prescribed minimum and maximum thresholds in order to avoid unphysical values). This is quite similar to the final result of the approach described here, with the important differences that in our case (i) the range, distribution and autocorrelation function of the resulting process are not prescribed but determined by the nature of the transition rules, (ii) in the multidimensional case (more than one cloud type) several parameters are perturbed in a mutually correlated way, where the correlations again depend on the nature of the transition rules, and (iii) there is potentially a coupling between the statistical properties of the resulting process and the state of the GCM.

3.2. *Coupling the minimal model to the BM scheme*

We now take as an example the coupling of the minimal version of the stochastic model described in section 2.2 to the BM parametrization. We recall the basic design of the BM scheme; more details can be found in Betts and Miller (1986). In the usual BM scheme, the state of the atmosphere is relaxed towards a reference profile on a prescribed convective relaxation time-scale. The tendencies due to convection for the temperature T and moisture q are given respectively by the apparent heat source Q_1 and apparent moisture sink Q_2:

$$\begin{cases} Q_1 = C_\mathrm{p}\frac{T_\mathrm{c} - T}{\tau_0}, \\ Q_2 = L\frac{q_\mathrm{c} - q}{\tau_0}, \end{cases} \tag{31}$$

where C_p is the heat capacity and L the latent heat. The reference profiles T_c and q_c are computed by an iterative algorithm (which uses the pseudoadiabatic profile as a first guess) in order to guarantee the conservation of the vertically integrated moist static energy, so that the vertical integral of $Q_1 + Q_2$ equals zero. The vertical integral of Q_1 (or $-Q_2$) is proportional to the total amount of convective precipitation produced by the convective scheme, since it is the intensity of conversion of latent into sensible heat through condensational heating. When the vertical integral of Q_1 turns out to be negative (which would then lead to negative precipitation), convection is supposed to be of non-precipitating shallow nature and the reference profiles are recomputed in order to realize a mixing of T and q from the cloud base to a reference pressure level. The relaxation time-scale τ_0 is normally set to 1–2 h for a deep convection case and 3–4 h for a shallow convection case (the actual values depend on the resolution of the model).

For consistency with the formulation of the two-state stochastic model, a simplified version of the BM scheme is considered in which shallow convection is not allowed. The BM scheme, even if not explicitly designed using a QE assumption, realizes in practice an exponential decay of measure $A = \int_0^\infty C_p(T_c - T)\,dz$ of the vertically integrated buoyancy as defined by T_c (the equivalent of the cloud work function of Arakawa and Schubert, 1974), i.e.

$$\frac{dA}{dt} = -\frac{A}{\tau_0}. \tag{32}$$

The crucial parameter of the scheme is therefore the relaxation time-scale τ_0, which controls the intensity of the negative feedback realized by unresolved convection on the growth of the instability measured by A. A sensible way of introducing the stochastic model into this parametrization is to define a new relaxation time-scale

$$\tau_S = \frac{\sigma_0}{\sigma}\tau_0. \tag{33}$$

Note that this definition is equivalent to the one in Lin and Neelin (2000) and Khouider *et al.* (2010). This satisfies the condition of convergence to the original parametrization in the deterministic limit of the stochastic model and represents the effect of convection being stronger when there are more active convective elements than in the limit case and vice versa. Note that in practical applications τ_S has to be larger than the time step in order to avoid numerical instability (and physical inconsistency). In order to avoid this problem, which does not occur in the deterministic scheme, we simply truncate the range of possible values of τ_S with the time step of the GCM as a lower bound. This approach can easily be extended to the Kuo parametrization (which, in its original formulation, is formally equivalent to Eq. (31), with a different definition of the relaxation time-scales and reference profiles) and to mass-flux schemes making use of a CAPE-like closure.

4. Aquaplanet experiments

4.1. Experimental design

The numerical model applied for this study is the Planet Simulator (Fraedrich *et al.*, 2005; Fraedrich, 2012), an intermediate-complexity GCM developed at the University of Hamburg and freely available at http://www.mi.uni-hamburg.de/plasim. The dynamical core is based on the Portable University Model of the Atmosphere (PUMA: Fraedrich *et al.*, 1998), which has already been used in testing stochastic parametrization techniques (Seiffert *et al.*, 2006). The primitive equations are solved by a spectral transform method (Eliasen *et al.*, 1970; Orszag, 1970). Parametrizations include long- and short-wave radiation (Sasamori, 1968; Lacis and Hansen, 1974) with interactive clouds (Stephens, 1978; Stephens *et al.*, 1984; Slingo and Slingo, 1991). A horizontal diffusion according to Laursen and Eliasen (1989) is applied. Formulations for boundary-layer fluxes of latent and sensible heat and for vertical diffusion follow Louis (1979), Louis *et al.* (1981) and Roeckner *et al.* (1992). Stratiform precipitation is generated in supersaturated states. In the standard set-up, the Kuo scheme (Kuo, 1965, 1974) is used for deep moist convection, while shallow cumulus convection is parametrized as vertical diffusion. In this work, we use instead the BM scheme for deep convection as described in the previous section, so that shallow convection is not considered. The deterministic relaxation time-scale is $\tau_0 = 2$ h. Shallow convection is an important process in determining the properties of tropical dynamics, but our aim here is to check how the coupling with the stochastic model impacts the basic statistics of the tropical activity in the simplest possible conditions and we are not concerned about realism or specific aspects of tropical dynamics, which remain the subjects of later works.

As experimental settings, we run the model in aquaplanet conditions with $T42$ horizontal resolution and 10 σ levels in the vertical. The SST of the model is fixed following the same profile as the CTRL scenario in Dahms *et al.* (2011), which follows the control distribution of Neale and Hoskins (2001):

$$T_s(\lambda,\phi) = \begin{cases} 27\left(1 - \sin^2\left(\frac{3\phi}{2}\right)\right): & -\frac{\pi}{3} < \phi < \frac{\pi}{3}, \\ 0: & \text{otherwise}, \end{cases} \tag{34}$$

where λ and ϕ represent longitude and latitude, respectively. The model is run under perpetual equinoctial conditions and without a daily cycle, so that no time-dependent forcings act on the dynamics. An important property of the Planet Simulator is that the simulated circulation remains zonally symmetric if the model is initialized in a zonally symmetric state and driven by zonally symmetric boundary conditions. In each experiment, the model is run for 30 years of integration and the analysis is limited to the last 25 years in order to account for the spin-up. This is a long spin-up time for an aquaplanet experiment with fixed SST, but since, as stated in the Introduction, our analysis will focus on changes in the statistics of extremes, we prefer to stay on the safe side, given the delicate nature of the investigated properties.

The stochastic model has three parameters: σ_0 and τ, which depend on the the transition rates, and N, which depends on the geometry of the system. The parametric exploration can be simplified considering that the convective cloud fraction in a GCM box (which will take values fluctuating around σ_0) is typically supposed to be small, of the order of a few percent. The proposed coupling consists of multiplying the amount of convective precipitation by the factor $\gamma = \sigma/\sigma_0$. If $\sigma_0 \ll 1$, the evolution equation for γ can be approximated as

$$\frac{d\gamma}{dt} \approx \frac{1-\gamma}{\tau} + \epsilon'_N\sqrt{\frac{1+\gamma}{\tau}}\xi, \tag{35}$$

where $\epsilon'_N = \sigma_0^{-1/2}\epsilon_N$ (i.e. the noise scales with the inverse of the square root of the expected cloud number $N\sigma_0$). This means that, in the regime $\sigma_0 \ll 1$ (i.e. the physically interesting one for us), changing σ_0 is almost the same as modulating the noise amplitude. Therefore, in testing the sensitivity of the model we fix σ_0 to a small value ($\sigma_0 = 0.05$) in all experiments and tune only τ and N. The values considered for the other parameters are $\tau = (3, 6, 12, 24)$ h and $N = (100, 225, 400, 1000)$. For $T42$ resolution, these values of N correspond to sizes of the convective elements in the range 10–30 km, which are too large but not totally unreasonable numbers for observed convection in the Tropics.

In general, the chosen set-up configures a rather abstract and simplified experiment, but at this stage we just want to perform a sensitivity analysis of the coupling of the simplest possible version of the model to a GCM, exploring its parameter space for ranges of values compatible with the order of magnitudes observed for tropical cloud systems, without any claim of realism. The analysis of the impact of introducing the stochastic parametrization is focused here only on the statistical properties of the convective precipitation, which is the quantity directly modified by the stochastic term. We will consider changes in the climatology of the convective precipitation, i.e. mean and standard deviation as function of the latitude, well as changes in the extremes of its daily values. Before showing the results, we recall the basics of extreme value theory (EVT) and describe how we performed the statistical analysis.

4.2. EVT analysis of daily extremes of convective precipitation

A particularly interesting aspect of stochastic parametrizations is the impact that they could have on the statistics of the extremes (Stechmann and Neelin, 2011) of a GCM. Deterministic parametrizations currently in use in state-of-the-art GCMs are more or less able to reproduce the climatology of the system in terms of its bulk statistics, although with different levels of geographical detail and performance, depending on the complexity of the scheme and of the GCM itself. The extent to which they are able to represent higher order statistics and extremes in particular is less clear. This is an extremely important topic, considering that one of the main concerns regarding the climate-change problem is how the statistics of intense, extreme events will change in a changing climate. A summary on the topic of extreme-event analysis in a geophysical context can be found in Ghil *et al.* (2011) and Sura (2011).

The most common approach to extreme-value analysis is the so-called block-maxima approach. It consists of dividing a time series of an observable into bins and picking the maximum value in each of them. An asymptotic theorem due to Gnedenko (1943) states that, under certain conditions, the sample of maxima converges to the so called generalized extreme-value distribution (GEV). The GEV distribution is a three-parameter distribution with cumulative distribution function

$$F(x;\mu,\sigma,\xi) = \exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}. \tag{36}$$

The location and scale parameters μ and σ can be reduced to 0 and 1 respectively by a rescaling of the data. The shape parameter ξ is more fundamental and determines the domain of the probability distribution function. Depending on the value of ξ, the family of distributions is divided into three subfamilies. When ξ=0, the distribution is of Gumbel or type I kind; when ξ >0, the distribution is of the Frechét or type II kind; and when ξ <0, the distribution is of the Weibull or type III kind. Although less important from a mathematical point of view, the location and scale parameters are extremely important in practical applications, since they represent the typical value and typical range of variability of the extreme events (in a loose sense they are a type of 'mean' and 'standard deviation' of the distribution of extreme events).

The estimation of the GEV parameters from a sample of data is not trivial. The problem of convergence of the empirical distribution of extremes obtained with the block-maximum approach to the theoretical GEV distribution has been widely explored (see Coles *et al.*, 1999; Faranda *et al.*, 2011, and references therein). The main problem is that a reliable estimation of the parameters of the asymptotic distribution requires a very large amount of data and in any case the convergence properties differ substantially from system to system. In order to increase the size of our sample, we have taken advantage of the zonal and hemispheric symmetry of the Planet Simulator in aquaplanet set-up. In this condition, each grid point on the same latitudinal circle (in both hemispheres) is statistically equivalent. We can therefore consider the time series of daily convective precipitation in each grid point on the same latitudinal circle as independent realizations of the same process and put them together in order to increase the size of the sample. Of course, we can do this only if the time series are not correlated. Computing the spatial correlation function of daily convective precipitation in the zonal direction shows that picking one grid point in every four is sufficient to having nearly uncorrelated time series. This is, of course, only a linear correlation analysis, which does not contain all the information on the mutual dependence of the time series, but it should be sufficient for our purposes. In this way, our sample of data consists of 576 000 daily values. Defining a block length of 720 days, we have 800 maxima to perform a robust analysis.

4.3. Analysis of results

Figures 4 and 5 show the zonal mean and standard deviation of daily convective precipitation over 25 years after five

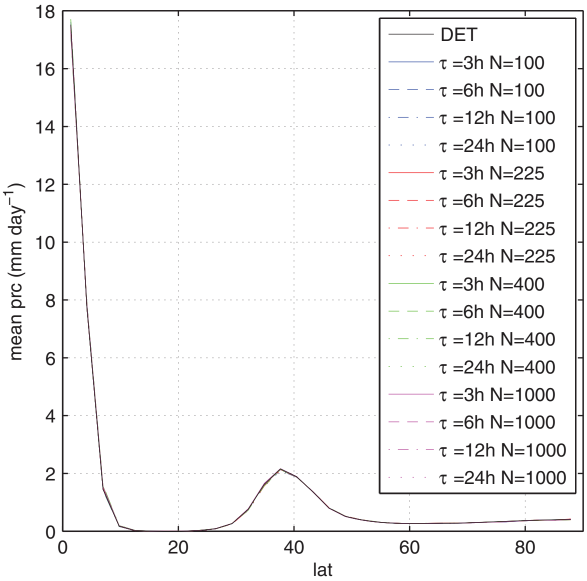

Figure 4. Mean of daily convective precipitation (climatology over 25 years) for Aquaplanet experiments in the absence of a diurnal or seasonal cycle. Because of the hemispheric and zonal symmetry of the dynamics, only the zonal quantities in the Northern Hemisphere are shown. The black solid line represents the standard deterministic case with the BM scheme without shallow convection, while the other lines (coloured in the online article) represent the experiments coupling the BM deep convection scheme with the two-state stochastic model, for $\sigma_0 = 0.05$ and different values of τ and N.

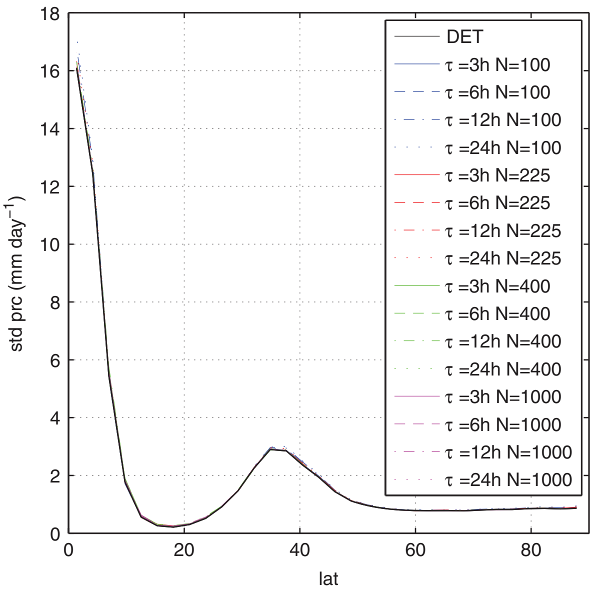

Figure 5. Standard deviation of daily convective precipitation (climatology over 25 years) for Aquaplanet experiments in the absence of a diurnal or seasonal cycle. Because of the hemispheric and zonal symmetry of the dynamics, only the zonal quantities in the Northern Hemisphere are shown. The black solid line represents the standard deterministic case with the BM scheme without shallow convection, while the other lines (coloured in the online article) represent the experiments coupling the BM deep convection scheme with the two-state stochastic model, for $\sigma_0 = 0.05$ and different values of τ and N.

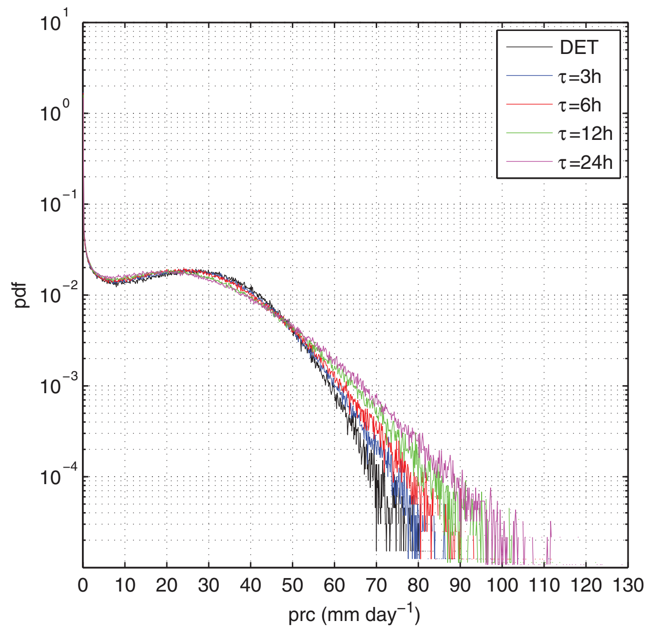

Figure 6. Probability distribution function of daily convective precipitation at 1.4°N for Aquaplanet experiments in the absence of a diurnal or seasonal cycle. The black solid line represents the standard deterministic case with the BM scheme without shallow convection, while the other lines (coloured in the online article) represent the experiments coupling the BM deep convection scheme with the two-state stochastic model, for $\sigma_0 = 0.05$, $N = 100$ and different values of τ.

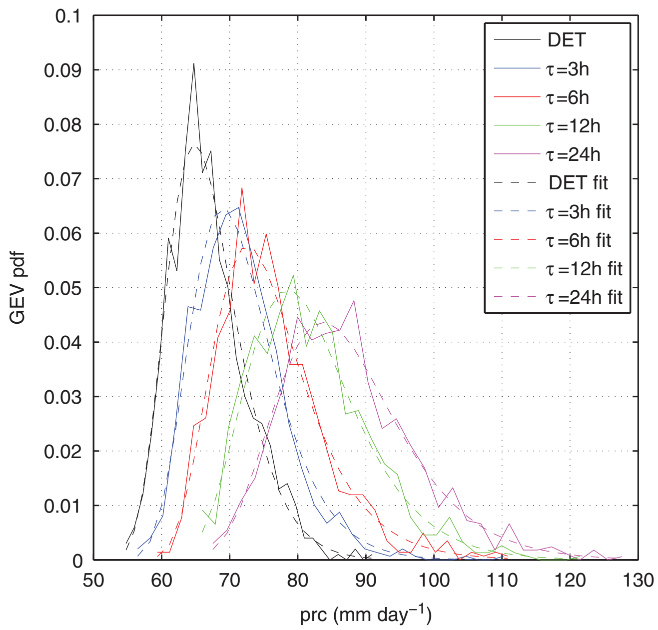

Figure 7. Probability distribution function of maxima of daily convective precipitation at 1.4°N for Aquaplanet experiments in the absence of a diurnal or seasonal cycle. The black solid line represents the standard deterministic case with the BM scheme without shallow convection, while the other solid lines (coloured in the online article) represent the experiments coupling the BM deep convection scheme with the two-state stochastic model, for $\sigma_0 = 0.05$, $N = 100$ and different values of τ. The maxima have been selected following the block maxima approach as explained in the text. The dashed lines (coloured in the online article) represent the best fit to a GEV distribution obtained with the maximum-likelihood method.

years of spin-up, for the standard deterministic run and for the experiments with the stochastic model. Because of the hemispheric and zonal symmetry of the dynamics, only the zonal quantities in the Northern Hemisphere are shown. Surprisingly, the introduction of the stochastic model does not affect the climatology of daily convective precipitation. Instead, the impact of the stochastic parametrization is concentrated on the higher moments of precipitation.

Figure 6 shows the probability distribution function of the daily convective precipitation at the closest grid point to the Equator (ca. 1.4°N), where convective precipitation is at its maximum. The curves refer to the deterministic run and the stochastic runs for $N = 100$ (the case with larger noise amplitude). We can see that the distributions differ substantially only in the upper tails, with larger values for larger autocorrelation times of the stochastic forcing. Qualitatively, the same result (not shown) is obtained fixing the autocorrelation time and tuning the amplitude of the noise (of course with larger values for larger noise amplitudes).

Figure 7 shows the GEV distributions for the same experiments. The empirical distributions have been fitted with the maximum-likelihood method, with good results. We can see that for larger autocorrelation times the distributions of extreme values of convective precipitation become broader and shift towards higher values. We can also see that the range of the GEV distributions coincides with the range over which the probability distribution functions of Figure 6 differ substantially: it appears that only the extreme values (in the proper statistical sense) are affected by the introduction of the stochastic model.

In order to make the analysis more quantitative, Figures 8–10 show the estimates of the location, scale and shape parameters as a function of the autocorrelation time, for different values of the noise amplitude. We can see that larger autocorrelation times lead to larger values of the location and scale parameters, with more pronounced sensitivity with larger values of the noise. For both parameters the increase is roughly logarithmic. The shape parameter shows no sensitivity to changes in the parameters among the uncertainties, so that the nature of the GEV is not affected. Figures 8–10 represent a possible parametrization of extremes of daily convective precipitation through the parameters of our stochastic model.

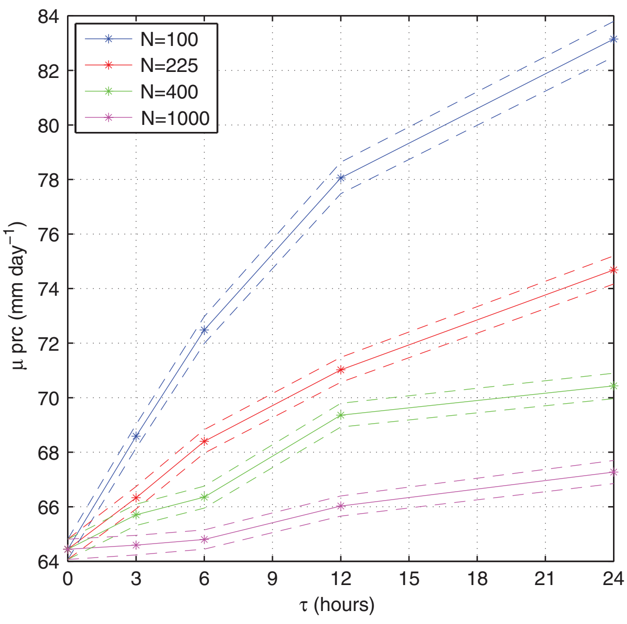

Figure 8. Location parameter of the GEV distribution of maxima of daily convective precipitation at 1.4°N versus values of τ, for $\sigma_0 = 0.05$ and different values of N, for Aquaplanet experiments in the absence of a diurnal or seasonal cycle. The point at $\tau = 0$ represents the standard deterministic case with the BM scheme without shallow convection. The maxima have been selected following the block maxima approach, using the methodology explained in the text. The dashed lines represent the 95% confidence interval of the estimate of the GEV parameters obtained with the maximum-likelihood method.

5. Summary and discussion

In this article, we have presented a rather general framework to include subgrid stochastic lattice gas models for the population dynamics of an ensemble of convective events in a host deterministic parametrization scheme. The proposed formalism is along the lines of models previously presented in the literature (Majda and Khouider, 2002; Khouider *et al.*, 2010; Stechmann

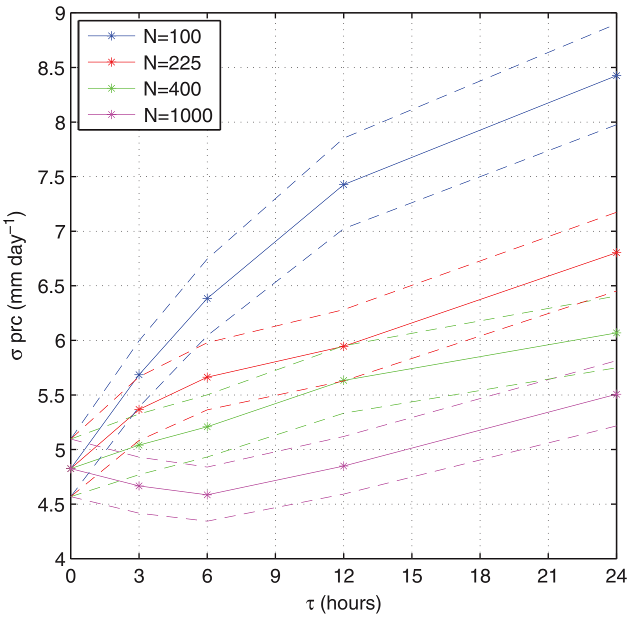

Figure 9. Scale parameter of the GEV distribution of maxima of daily convective precipitation at 1.4°N versus values of τ, for $\sigma_0 = 0.05$ and different values of N, for Aquaplanet experiments in the absence of a diurnal or seasonal cycle. The point at $\tau = 0$ represents the standard deterministic case with the BM scheme without shallow convection. The maxima have been selected following the block maxima approach, using the methodology explained in the text. The dashed lines represent the 95% confidence interval of the estimate of the GEV parameters obtained with the maximum-likelihood method.

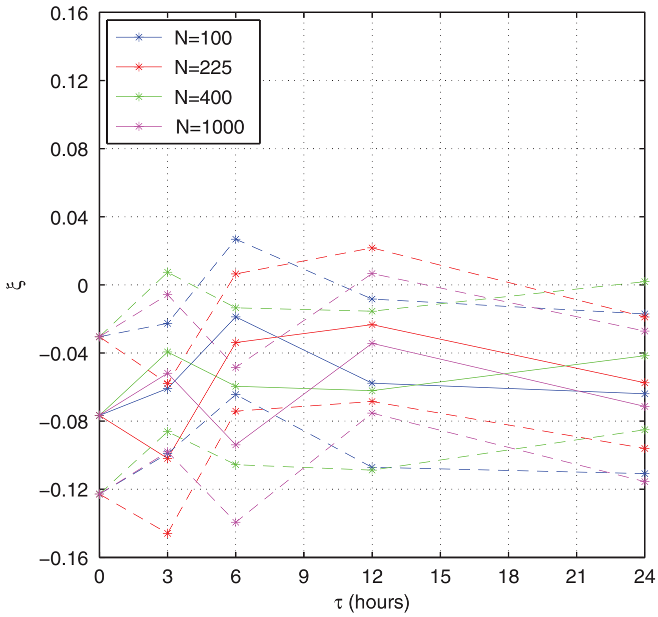

Figure 10. Shape parameter of the GEV distribution of maxima of daily convective precipitation at 1.4°N versus values of τ, for $\sigma_0 = 0.05$ and different values of N, for Aquaplanet experiments in the absence of a diurnal or seasonal cycle. The point at $\tau = 0$ represents the standard deterministic case with the BM scheme without shallow convection. The maxima have been selected following the block maxima approach, using the methodology explained in the text. The dashed lines represent the 95% confidence interval of the estimate of the GEV parameters obtained with the maximum-likelihood method.

and Neelin, 2011; Frenkel *et al.*, 2012) and partially on the lines of Plant (2012) and could be used in order to bridge those approaches to the world of operational GCMs.

In order to to make the application to real GCMs numerically treatable, we have derived a reduced model for the time evolution of the macrostate (cloud fraction) of the lattice model. The method consists of a first-order correction to the mean-field limit of the system, in a fashion similar to Tome and de Oliveira (2009).

We have studied the properties of the minimal version of the stochastic model, a binary system with fixed transition rates, in some detail. In this formulation the stochastic model reduces to a single SDE that is analytically treatable. The SDE corresponds to an exponential decay to an equilibrium value forced by a multiplicative noise term. The model has three parameters: the intensity of the noise, which scales with the inverse of the square root of the size of the system, and two parameters that depend on the transition (birth and death) rates, which are the equilibrium cloud fraction and the relaxation time-scale. Analysis of the Fokker–Planck equation shows that the stationary distribution of the model depends only on the intensity of the noise and the equilibrium cloud fraction and is independent of the relaxation time-scale. In contrast, the autocorrelation function of the process depends uniquely on the relaxation time-scale and consists of an exponential decay on the same time-scale. The process, therefore, has no memory and a white spectrum. This analysis shows that it is possible to tune the transition rates in order to tune independent properties of the system, allowing for a systematic exploration of the behaviour of the model in applications with a GCM.

We have analyzed the numerical accuracy of the reduction method for the minimal version of the stochastic model, comparing direct simulations of the lattice model with iterations of the reduced SDE for different values of the parameters. The parameter space of the model has been explored using a range of values compatible with application to the representation of a cloud system in a GCM grid box. We have shown that the reduction method reproduces the properties of the system remarkably well, even when tested close to the limits of its applicability. Regarding more complex applications, there is no reason to expect a worse numerical accuracy in systems featuring multiple states and/or time-dependent transition rates. In the case of systems featuring local interactions and, therefore, state-dependent transition rates, the validity and accuracy of the reduction method depend on the nature and strength of the interactions and have to be tested case by case. Overall, the reduction method seems to be promising for the kind of application we have in mind.

We have defined a coupling strategy in order to include the stochastic model in a pre-existing, host deterministic parametrization scheme. The state of the stochastic model modifies a relevant parameter of the parametrization in such a way that when perfect space- and time-scale separation is achieved we retrieve the usual value used in the deterministic version of the parametrization. In this way we define a robust coupling, which introduces first-order corrections due to the finite size and time evolution of the ensemble of convective events, around the zeroth-order description given by the original deterministic version of the parametrization. Simplified representations of the conditional dependence of the activation and decay of convective events on large-scale conditions and mutual interactions can be added through the definition of the transition rates. The strategy has been applied to the coupling of the minimal version of the stochastic model to a simplified version of the BM scheme.

We have performed numerical experiments with an aquaplanet version of the Planet Simulator, an intermediate complexity atmospheric general circulation model (AGCM). The experiments have been performed with a fixed zonally symmetric distribution of the SST without seasonal and daily cycles, in order to study the impact of the introduction of the stochastic model on a zonally symmetric dynamics in the absence of time-dependent forcings. The stochastic model is set up as a binary system with fixed transition probabilities, coupled to the BM convective scheme as described above. This formulation of the model is minimal in the sense that it introduces into the GCM only the effects coming from considering a demographic description of the cloud system. We have performed a limited exploration of the parameter space of the stochastic model, in ranges of values compatible with the observed properties of tropical convection. We have studied the sensitivity to changes of size of the lattice model (which determines the intensity of the noise and the shape of the stationary distribution, without affecting the memory of the stochastic process) and to the intrinsic time-scale of the model (which controls the memory of the process without affecting the stationary distribution).

The analysis has focused on convective precipitation, which is the quantity directly modified by the stochastic term. In these settings, the stochastic extension of the parametrization conserves the climatology of its deterministic limit, thus confirming that the coupling has been defined in a robust way. The analysis of the distribution of the daily convective precipitation in tropical areas shows that the inclusion of the stochastic term impacts only the upper tail of the distribution, without affecting the bulk statistics. We have performed a detailed analysis of the changes in the extreme statistics using EVT. The location and scale parameters of the GEV distribution of tropical daily convective precipitation turn out to be highly sensitive to both the noise intensity and the autocorrelation time of the stochastic forcing. They increase logarithmically with larger noise intensity and larger autocorrelation time. This means larger and more spread typical values for the daily extremes of convective precipitation. In the limit of vanishing noise intensity and autocorrelation time, the parameters converge to the values of the deterministic case. The shape parameter seems to be insensitive to changes in any parameter.

These findings suggest that the coupling behaves as expected in terms of robustness. The bulk statistics of convection is not affected by the introduction of the stochastic term and only high-order moments are modified. The changes introduced in the extreme statistics tend to zero, increasing the number of convective elements and decreasing the characteristic time-scale of the process, i.e. approaching space- and time-scale separation respectively. While the increase in the typical value and range of the extremes of daily convective precipitation with increasing amplitude of fluctuations of the stochastic process is somehow expected, the reason that these should increase with larger autocorrelation times of the noise is less clear. Lin and Neelin (2000, 2002, 2003) have already shown the sensitivity of tropical variability to the autocorrelation time of a stochastic forcing, although in a very different kind of analysis.

Our results also constitute an instructive example of the fact that a parametrization calibrated on the climatology of a process is not necessarily a good parametrization for studying the extreme statistics of that process. We have given a practical example of a parametrization that, for a large range of values of some of its parameters, reproduces the same climatology of a characteristic quantity, while showing large differences in the extremes for that range of values. In our case, the parametrization is stochastic and has been derived in order to represent specific features of atmospheric convection, but the principle holds in general.

Starting from this work, several future lines of research can be proposed. A first obvious experiment is to make the transition rates dependent on the state of the large-scale model, in order to realize an effective two-way coupling. Previous works have introduced CIN, CAPE and measures of the dryness of the atmosphere in order to characterize the transition rates (Majda and Khouider, 2002; Khouider *et al.*, 2003, 2010; Frenkel *et al.*, 2012). A promising alternative could be the proposal of Stechmann and Neelin (2011) to make the transitions between inactive and active convective states dependent on a critical value of the precipitable water. The idea of making the activation of convection dependent on a critical value of the moisture content of the atmospheric column is indeed not new: in many implementations of the Kuo-like moisture convergence closure, it is common to introduce a critical value of the relative humidity of the atmospheric column below which convection is shut down. For example, Frierson *et al.* (2011) have shown that tuning this critical value (corresponding to constraining the convective activity differently) has an impact on the intraseasonal variability of the model.

Another possibility could be to introduce multiple convective regimes and the correspondent life cycle of convective events. This will be needed in particular in the application to more realistic convective parametrizations. Operational convective parametrizations typically have at least two kinds of convective states (shallow and deep convection), sometimes a few more. In employing multiple convective regimes, it could be particularly interesting to define the dependence of the transition rates on the large-scale fields in order to capture preconditioning processes, along the lines of Khouider *et al.* (2010) and Frenkel *et al.* (2012), who already obtained promising results in applications with simplified models of tropical dynamics.

Finally, it could be interesting to introduce simple interaction rules for the lattice elements. As stated already, clustering of convective events is indeed observed in studies of tropical dynamics. The nature of the interactions between clouds is nevertheless still unclear. Therefore, at this stage, tentative rules for local interactions should be introduced in a very crude form, without pretending to give a realistic, quantitative representation of the phenomenon.

Acknowledgements

Thanks are due to referees and editor for their helpful comments. The authors also thank Edilbert Kirk and Frank Sielmann for technical support. FR acknowledges financial support from Clisap and SICSS and is thankful to Salvatore Pascale for discussions on the content of the article. FR and KF acknowledge support by a fellowship of the Max Planck Society.

Appendix: Derivation of SDE reduction

In this Appendix we derive the relations $\langle \mathrm{d}\boldsymbol{\sigma}^t \rangle = \boldsymbol{R}^t \boldsymbol{\sigma}^t\,\mathrm{d}t$ and $\boldsymbol{C}^t = \boldsymbol{D}^t\,\mathrm{d}t$, which are used in section 2 to derive the SDE reduction. Knowing that the state of the system at time t is $\boldsymbol{\sigma}^t$ (and knowing the configuration of the lattice, i.e. the value of each $\boldsymbol{\sigma}^{nt}$), the expectation value of $\mathrm{d}\boldsymbol{\sigma}^t$ is componentwise given by (neglecting terms $\mathcal{O}(\mathrm{d}t^2)$)

$$\begin{aligned}\langle \mathrm{d}\sigma_s{}^t \rangle_{\boldsymbol{\sigma}^t} &= \frac{1}{N}\sum_{n=1}^{N} \langle \mathrm{d}\sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t} \\ &= \frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{S} (p_{si}^{nt} - \delta_{si})\sigma_i^{nt} = \frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{S} R_{si}^{nt}\sigma_i^{nt}\,\mathrm{d}t, \end{aligned} \tag{A1}$$

where $\langle \bullet \rangle_{\boldsymbol{\sigma}^t}$ represents the expectation value of a quantity conditional on the knowledge of the state of the system at time t (where by this we mean knowing the exact configuration of the entire lattice, although we use the symbol $\boldsymbol{\sigma}^t$ for simplicity of notation). In the mean-field approximation, the transition rates are constant over the entire lattice once we replace the local interaction terms R_{si}^{nt} with the mean-field term R_{si}^t, so that

$$\langle \mathrm{d}\sigma_s{}^t \rangle_{\boldsymbol{\sigma}^t} = \sum_{i=1}^{S} R_{si}^t \sigma_i^t\,\mathrm{d}t \tag{A2}$$

or, in vectorial notation,

$$\langle \mathrm{d}\boldsymbol{\sigma}^t \rangle_{\boldsymbol{\sigma}^t} = \boldsymbol{R}^t \boldsymbol{\sigma}^t\,\mathrm{d}t. \tag{A3}$$

The components of the covariance matrix $\boldsymbol{C}^t$ are, by definition, given by

$$\begin{aligned}\frac{1}{N}C_{ss'}^t &= \langle \mathrm{d}\sigma_s{}^t \mathrm{d}\sigma_{s'}^t \rangle_{\boldsymbol{\sigma}^t} - \langle \mathrm{d}\sigma_s{}^t \rangle_{\boldsymbol{\sigma}^t} \langle \mathrm{d}\sigma_{s'}^t \rangle_{\boldsymbol{\sigma}^t} \\ &= \frac{1}{N^2}\sum_{n=1}^{N}\sum_{n'=1}^{N}\left[\langle \mathrm{d}\sigma_s^{nt} \mathrm{d}\sigma_{s'}^{n't} \rangle_{\boldsymbol{\sigma}^t} - \langle \mathrm{d}\sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t} \langle \mathrm{d}\sigma_{s'}^{n't} \rangle_{\boldsymbol{\sigma}^t} \right]. \end{aligned} \tag{A4}$$

Thanks to the mean-field approximation, for $n \neq n'$ we have

$$\langle \mathrm{d}\sigma_s^{nt} \mathrm{d}\sigma_{s'}^{n't} \rangle_{\boldsymbol{\sigma}^t} \approx \langle \mathrm{d}\sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t} \langle \mathrm{d}\sigma_{s'}^{n't} \rangle_{\boldsymbol{\sigma}^t}, \tag{A5}$$

so that we retain only the terms for which $n = n'$,

$$C_{ss'}^t = \frac{1}{N} \sum_{n=1}^{N} \left[\langle \mathrm{d}\sigma_s^{nt} \mathrm{d}\sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t} - \langle \mathrm{d}\sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t} \langle \mathrm{d}\sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t} \right]. \tag{A6}$$

Rearranging some terms, the first term in the sum is given by

$$\begin{aligned} \langle \mathrm{d}\sigma_s^{nt} \mathrm{d}\sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t} =& \langle \sigma_s^{nt+\mathrm{d}t} \sigma_{s'}^{nt+\mathrm{d}t} \rangle_{\boldsymbol{\sigma}^t} - \langle \sigma_s^{nt} \sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t} \\ & - \langle \sigma_s^{nt} \mathrm{d}\sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t} - \langle \sigma_{s'}^{nt} \mathrm{d}\sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t}. \end{aligned} \tag{A7}$$

Using the fact that in general $\sigma_s^{nt} \sigma_{s'}^{nt} = \delta_{ss'} \sigma_s^{nt}$, where $\delta_{ss'}$ is the usual Kronecker delta, for the first two terms we have the result

$$\begin{aligned} & \langle \sigma_s^{nt+\mathrm{d}t} \sigma_{s'}^{nt+\mathrm{d}t} \rangle_{\boldsymbol{\sigma}^t} - \langle \sigma_s^{nt} \sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t} \\ & \quad = \delta_{ss'} \langle \sigma_s^{nt+\mathrm{d}t} \rangle_{\boldsymbol{\sigma}^t} - \delta_{ss'} \langle \sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t} = \delta_{ss'} \langle \mathrm{d}\sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t}. \end{aligned} \tag{A8}$$

It is easy to see that the second two terms are given by

$$\langle \sigma_s^{nt} \mathrm{d}\sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t} + \langle \sigma_{s'}^{nt} \mathrm{d}\sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t} = p_{s's}^{nt} \sigma_s^{nt} + p_{ss'}^{nt} \sigma_{s'}^{nt}. \tag{A9}$$

Therefore, making use of Eq. (38) and Eq. (1),

$$\begin{aligned} & \langle \mathrm{d}\sigma_s^{nt} \mathrm{d}\sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t} \\ & \quad = -R_{s's}^{nt} \sigma_s^{nt} \mathrm{d}t - R_{ss'}^{nt} \sigma_{s'}^{nt} \mathrm{d}t + \delta_{ss'} \sum_{i=1}^{S} R_{si}^{nt} \sigma_i^{nt} \mathrm{d}t. \end{aligned} \tag{A10}$$

Making use of Eq. (38), the second term of Eq. (42) can be neglected, since

$$\begin{aligned} & \langle \mathrm{d}\sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t} \langle \mathrm{d}\sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t} \\ & \quad = \left(\sum_{i=1}^{S} R_{si}^{nt} \sigma_i^{nt} \mathrm{d}t \right) \left(\sum_{j=1}^{S} R_{sj}^{nt} \sigma_j^{nt} \mathrm{d}t \right) = \mathcal{O}\left(\mathrm{d}t^2\right) \end{aligned} \tag{A11}$$

and the terms in Eq. (46) are $\mathcal{O}(\mathrm{d}t)$. Therefore, we are left with

$$C_{ss'}^t = \frac{\mathrm{d}t}{N} \sum_{n=1}^{N} \left(-R_{s's}^{nt} \sigma_s^{nt} - R_{ss'}^{nt} \sigma_{s'}^{nt} + \delta_{ss'} \sum_{i=1}^{S} R_{si}^{nt} \sigma_i^{nt} \right) \tag{A12}$$

and, making use of the mean-field approximation, we can eventually write

$$\boldsymbol{C}^t = \boldsymbol{D}^t \mathrm{d}t, \tag{A13}$$

defining the matrix $\boldsymbol{D}^t$ as

$$D_{ss'}^t = -R_{s's}^t \sigma_s^t - R_{ss'}^t \sigma_{s'}^t + \delta_{ss'} \sum_{i=1}^{S} R_{si}^t \sigma_i^t, \tag{A14}$$

as we have introduced it in Eq. (11) of section 2. We see therefore that both $\langle \mathrm{d}\boldsymbol{\sigma}^t \rangle_{\boldsymbol{\sigma}^t}$ and $\boldsymbol{C}^t$ scale with $\mathrm{d}t$ and can be written as functions of $\boldsymbol{R}^t$ and $\boldsymbol{\sigma}^t$, thus leading to the SDE reduction proposed in section 2.

References

Adams DK, Renno NO. 2003. Remarks on quasi-equilibrium theory. *J. Atmos. Sci.* **60**: 178–181.

Arakawa A. 2004. The cumulus parameterization problem: Past, present and future. *J. Clim.* **17**: 2493–2525.

Arakawa A, Schubert WH. 1974. Interaction of a cumulus cloud ensemble with the large-scale environment, part I. *J. Atmos. Sci.* **31**: 674–701.

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

Berner J, Shutts G, Palmer T. 2005. 'Parameterising the multiscale structure of organised convection using a cellular automaton'. In *ECMWF Workshop on Representation of Sub-grid Processes Using Stochastic–Dynamic Models*. Reading, UK.

Betts AK, Miller MJ. 1986. A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX, ATEX and Arctic air-mass data sets. *Q. J. R. Meteorol. Soc.* **112**: 693–709.

Black AJ, McKane AJ. 2012. Stochastic formulation of ecological models and their applications. *Trends Ecol. Evol.* **27**: 337–345.

Bougeault P. 1985. A simple parameterization of the large-scale effects of cumulus convection. *Mon. Weather Rev.* **113**: 2108–2121.

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

Bright DR, Mullen SL. 2002. Short-range ensemble forecasts of precipitation during the southwest monsoon. *Weather and Forecasting* **17**: 1080–1100.

Buizza R, Miller M, Palmer T. 1999. Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Q. J. R. Meteorol. Soc.* **125**: 2887–2908.

Chekroun MD, Simonnet E, Ghil M. 2011. Stochastic climate dynamics: Random attractors and time-dependent invariant measures. *Physica D.* **240**: 1685–1700.

Cohen BG, Craig GC. 2006. Fluctuations in an equilibrium convective ensemble. Part II: Numerical experiments. *J. Atmos. Sci.* **63**: 2005–2015.

Coles S, Heffernan J, Tawn J. 1999. Dependence measures for extreme value analyses. *Extremes* **2**: 339–365.

Craig GC, Cohen BG. 2006. Fluctuations in an equilibrium convective ensemble. Part I: Theoretical formulation. *J. Atmos. Sci.* **63**: 1996–2004.

Dahms E, Borth H, Lunkeit F, Fraedrich K. 2011. ITCZ splitting and the influence of large-scale eddy fields on the tropical mean state. *J. Meteorol. Soc. Jpn* **89**: 399–411.

Eliasen E, Machenhauer B, Rasmussen E. 1970. '*On a numerical method for integration of the hydrodynamical equations with a spectral representation of the horizontal fields*', Report 2. Institute for Theoretical Meteorology, Copenhagen University: Copenhagen.

Faranda D, Lucarini V, Turchetti G, Vaienti S. 2011. Numerical convergence of the block-maxima approach to the generalized extreme value distribution. *J. Stat. Phys.* **145**: 1156–1180.

Fraedrich K. 1973. On the parameterization of cumulus convection by lateral mixing and compensating subsidence. Part 1. *J. Atmos. Sci.* **30**: 408–413.

Fraedrich K. 2012. A suite of user-friendly global climate models: Hysteresis experiments. *Eur. Phys. J. Plus* **127**: 1–9.

Fraedrich K, Kirk E, Lunkeit F. 1998. '*PUMA: Portable University Model of the atmosphere*', Technical Report 16. Deutsches Klimarechenzentrum: Hamburg, Germany.

Fraedrich K, Jansen H, Luksch U, Lunkeit F. 2005. The planet simulator: Towards a user friendly model. *Meteorol. Z.* **14**: 299–304.

Frank WM. 1983. The cumulus parameterization problem. *Mon. Weather Rev.* **111**: 1859–1871.

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

Frierson DMW, Kim D, Kang I, Lee M, Lin J. 2011. Structure of AGCM-simulated convectively coupled Kelvin waves and sensitivity to convective parameterization. *J. Atmos. Sci.* **68**: 26–45.

Ghil M, Yiou P, Hallegatte S, Malamud BD, Naveau P, Soloviev A, Friederichs P, Keilis-Borok V, Kondrashov D, Kossobokov V, Mestre O, Nicolis C, Rust HW, Shebalin P, Vrac M, Witt A, Zaliapin I. 2011. Extreme events: Dynamics, statistics and prediction. *Nonlinear Processes Geophys.* **18**: 1–56.

Gnedenko BV. 1943. Sur la distribution limite du terme maximum d'une srie alatorie. *Ann. Math.* **44**: 423–453.

Gregory D. 1997. The mass flux approach to the parameterization of deep convection. In *The Physics and Parameterization of Moist Atmospheric Convection*, Smith RK. (ed.): 297–319. Kluwer Academic Publishers: Dordrecht, Netherlands.

Gregory D, Rowntree PR. 1990. A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure. *Mon. Weather Rev.* **119**: 1483–1506.

Hasselmann K. 1976. Stochastic climate models, part 1: Theory. *Tellus* **28**: 473–485.

Houze RA. 2004. Mesoscale convective systems. *Rev. Geophys.* **42**: RG4003.

Kain JS. 2004. The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteorol.* **43**: 170–181.

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

van Kampen NG. 2007. *Stochastic Processes in Physics and Chemistry* (3rd edn). Elsevier: Amsterdam, The Netherlands.

Khouider B, Majda AJ, Katsoulakis MA. 2003. Coarse-grained stochastic models for tropical convection and climate. *Proc. Natl. Acad. Sci. U.S.A.* **100**: 941–946.

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

Kuo HL. 1965. On formation and intensification of tropical cyclones through latent heat release by cumulus convection. *J. Atmos. Sci.* **22**: 40–63.

Kuo HL. 1974. Further studies of the parameterization of the influence of cumulus convection on large-scale flow. *J. Atmos. Sci.* **31**: 1232–1240.

Lacis AA, Hansen JE. 1974. A parameterization for the absorption of solar radiation in the Earth's atmosphere. *J. Atmos. Sci.* **31**: 118–133.

Laursen L, Eliasen E. 1989. On the effects of the damping mechanisms in an atmospheric general circulation model. *Tellus* **41A**: 385–400.

Lin JW, Neelin JD. 2000. Influence of a stochastic moist convective parameterization on tropical climate variability. *Geophys. Res. Lett.* **27**: 3691–3694, doi: 10.1029/2000GL011964.

Lin JW, Neelin JD. 2002. Considerations for stochastic convective parameterization. *J. Atmos. Sci.* **59**: 959–975.

Lin JW, Neelin JD. 2003. Toward stochastic moist convective parameterization in general circulation models. *Geophys. Res. Lett.* **30**: 1162, doi: 10.1029/2002GL016203.

Louis J. 1979. A parametric model of vertical eddy fluxes in the atmosphere. *Boundary-Layer Meteorol.* **17**: 187–202.

Louis J, Tiedtke M, Geleyn JF. 1981. 'A short history of the operational PBL-parameterization at ECMWF'. In *Proceedings, ECMWF Workshop on Planetary Boundary Layer Parameterization*. Reading, UK.

Majda AJ, Khouider B. 2002. Stochastic and mesoscopic models for tropical convection. *Proc. Natl. Acad. Sci. U.S.A.* **99**: 1123–1128.

McKane AJ, Newman TJ. 2004. Stochastic models in population biology and their deterministic analogs. *Phys. Rev. E* **70**: 902.

Manabe S, Smagorinsky J, Strickler RF. 1965. Simulated climatology of a general circulation model with a hydrological cycle. *Mon. Weather Rev.* **93**: 769–798.

Mapes BE. 1993. Gregarious tropical convection. *J. Atmos. Sci.* **50**: 2026–2037.

Mapes B. 1997. Equilibrium vs. activation controls on large-scale variations of tropical deep convection. In *The Physics and Parameterization of Moist Atmospheric Convection*, Smith RK. (ed.): 321–358. Kluwer Academic Publishers: Dordrecht, Netherlands.

Mapes B, Tulich S, Lin J, Zuidema P. 2006. The mesoscale convection life cycle: Building block or prototype for large-scale tropical waves? *Dyn. Atmos. Oceans* **42**: 3–29.

Moncrieff MW, Liu C. 2006. Representing convective organization in prediction models by a hybrid strategy. *J. Atmos. Sci.* **63**: 3404–3420.

Neale R, Hoskins B. 2001. A standard test for AGCMs including their physical parametrizations. I: The proposal. *Atmos. Sci. Lett.* **1**: 101–107.

Neelin JD, Peters O, Lin JW-B, Hales K, Holloway CE. 2008. Rethinking convective quasi-equilibrium: Observational constraints for stochastic convective schemes in climate models. *Philos. Trans. R. Soc. London, Ser. A* **366**: 2581–2604.

Neelin JD, Peters O, Hales K. 2009. The transition to strong convection. *J. Atmos. Sci.* **66**: 2367–2384.

Orszag S. 1970. Transform method for the calculation of vector-coupled sums: Application to the spectral form of the vorticity equation. *J. Atmos. Sci.* **27**: 890–895.

Palmer T, Williams P. 2010. *Stochastic Physics and Climate Modelling*. Cambridge University Press: Cambridge, UK.

Pan DM, Randall DA. 1998. A cumulus parameterization with prognostic closure. *Q. J. R. Meteorol. Soc.* **124**: 949–981.

Peters O, Neelin JD. 2006. Critical phenomena in atmospheric precipitation. *Nat. Phys.* **2**: 393–396.

Peters O, Neelin JD. 2009. Atmospheric convection as a continuous phase transition: Further evidence. *Int. J. Mod. Phys. B* **23**: 5453–5465.

Peters O, Hertlein C, Christensen K. 2002. A complexity view of rainfall. *Phys. Rev. Lett.* **88**: 18 701.

Peters O, Neelin JD, Nesbitt SW. 2009. Mesoscale convective systems and critical clusters. *J. Atmos. Sci.* **66**: 2913–2924.

Peters O, Deluca A, Corral A, Neelin JD, Holloway CE. 2010. Universality of rain event size distributions. *J. Stat. Mech.* **2010**(11): P11 030.

Plant RS. 2010. A review of the theoretical basis for bulk mass flux convective parameterization. *Atmos. Chem. Phys.* **10**: 3529–3544.

Plant RS. 2012. A new modelling framework for statistical cumulus dynamics. *Philos. Trans. R. Soc. London, Ser. A* **370**: 1041–1060.

Plant R, Craig GC. 2008. A stochastic parameterization for deep convection based on equilibrium statistics. *J. Atmos. Sci.* **65**: 87–105.

Randall DA, Pan DM. 1993. Implementation of the Arakawa–Schubert cumulus parameterization with a prognostic closure. In *The Representation of Cumulus Convection in Numerical Models*, Emanuel KA, Raymond DJ. (eds.) Meteorological Monographs **46**: 137–144. American Meteorological Society: Boston, MA.

Randall DA, Pan DM, Ding P, Gripe DG. 1997. Quasi-equilibrium. In *The Physics and Parameterization of Moist Atmospheric Convection*, Smith RK. (ed.): 359–385. Kluwer Academic Publishers: Dordrecht, Netherlands.

Randall DA, Wood RA, Bony S, Colman R, Fichefet T, Fyfe J, Kattsov V, Pitman A, Shukla J, Srinivasan J, Stouffer RJ, Sumi A, Taylor KE. 2007. Climate models and their evaluation. In *Climate Change 2007: The Physical Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL. (eds.). Cambridge University Press: Cambridge, UK.

Redelsperger J, Guichard F, Mondon S. 2000. A parameterization of mesoscale enhancement of surface fluxes for large-scale models. *J. Clim.* **13**: 402–421.

Risken H. 1989. *The Fokker–Planck Equation*, Springer Series in Synergetics **18**. Springer-Verlag: Berlin and Heidelberg, Germany; New York, NY.

Roeckner E, Arpe K, Rockel B, Bengtsson L, Duemenil L, Esch M, Kirk E, Lunkeit F, Schubert S, Windelband M, Schlese U, Brinkop S, Ponater M, Sausen R. 1992. '*Simulation of the present day climate with the ECHAM model: Impact of model physics and resolution*', Report 93. Max-Planck-Institut fur Meteorologie: Hamburg, Germany.

Sasamori T. 1968. The radiative cooling calculation for application to general circulation experiments. *J. Appl. Meteorol.* **7**: 721–729.

Seiffert R, Blender R, Fraedrich K. 2006. Subscale forcing in a global atmospheric circulation model and stochastic parameterisation. *Q. J. R. Meteorol. Soc.* **132**: 1627–1643.

Slingo A, Slingo JM. 1991. Response of the National Center for Atmospheric Research Community Climate Model to improvements in the representation of clouds. *J. Geophys. Res.* **96**: 341–357, doi: 10.1029/91JD00930.

Stechmann SN, Neelin JD. 2011. A stochastic model for the transition to strong convection. *J. Atmos. Sci.* **68**: 2955–2970.

Stephens GL. 1978. Radiation profiles in extended water clouds. II: Parameterization schemes. *J. Atmos. Sci.* **35**: 2123–2132.

Stephens GL, Ackerman S, Smith EA. 1984. A shortwave parameterization revised to improve cloud absorption. *J. Atmos. Sci.* **41**: 687–690.

Sura P. 2011. A general perspective of extreme events in weather and climate. *Atmos. Res.* **101**: 1–21.

Tiedtke M. 1989. A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Weather Rev.* **117**: 1779–1800.

Tome T, de Oliveira MJ. 2009. Role of noise in population dynamics cycles. *Phys. Rev. E* **79**: 061 128.

Tompkins AM. 2001. Organization of tropical convection in low vertical wind shears: The role of water vapor. *J. Atmos. Sci.* **58**: 529–545.

Yano JI. 1999. Scale-separation and quasi-equilibrium principles in Arakawa and Schuberts cumulus parameterization. *J. Atmos. Sci.* **56**: 3821–3823.

Yano J-I, Plant RS. 2012. Convective quasi-equilibrium. *Rev. Geophys.* **50**: RG4004.

Yano J-I, Grabowski WW, Roff GL, Mapes BE. 2000. Asymptotic approaches to convective quasi-equilibrium. *Q. J. R. Meteorol. Soc.* **126**: 1861–1887.

Yano J-I, Fraedrich K, Blender R. 2001. Tropical convective variability as $1/f$ noise. *J. Clim.* **14**: 3608–3616.

Yano J-I, Zhang C, Blender R, Fraedrich K. 2004. $1/f$ noise and pulse-like events in the tropical atmospheric surface variabilities. *Q. J. R. Meteorol. Soc.* **130**: 1697–1721.

