Stochastic parameterization of atmospheric convection - stochastic lattice models of cloud systems in PlaSim

Dissertation

Zur Erlangung des Doktorgrades

an der Fakultät für Mathematik, Informatik und Naturwissenschaften

Fachbereich Geowissenschaften

der Universität Hamburg

vorgelegt von

Francesco Ragone

aus Parma, Italien

Hamburg
2013

Als Dissertation angenommen vom Fachbereich Geowissenschaften der Universität Hamburg auf Grund der Gutachten von

Prof. Dr. Klaus Fraedrich und
Dr. Hartmut Borth

Hamburg, den

Prof. Dr. Christian Betzler
Leiter des Fachbereichs Geowissenschaften

Abstract

We propose a strategy to couple a stochastic lattice-gas model of a cloud system to a rather general class of convective parameterization schemes. The stochastic model consists in a sub-grid lattice of N elements which can be in one out of S states, each correspondent to a different cloud type. The time evolution of the elements of the lattice is represented as a Markov process on the set of the S states with transition rates dependent on large-scale (grid-box) fields and/or local interactions.

We derive a reduction method based on the mean-field approximation leading to a system of S-1 stochastic differential equations for the evolution of the macrostate (cloud fractions) of the lattice model in the limit of large N. The intensity of the noise scales with $N^{-1/2}$, consistently with the van Kampen system size expansion. The accuracy of the method is tested in a minimal version of the model.

We design a strategy to couple the lattice model to a generic parameterization scheme, so that in the limit of space and time scale separation the modified stochastic parameterization converges to the deterministic version of the host scheme.

We perform numerical experiments coupling the minimal version of the stochastic model to the Betts-Miller and Kuo schemes in the aqua-planet version of the Planet Simulator. After characterizing the climate produced by the standard deterministic model, we perform two set of experiments. In the first we consider constant birth and death rates of the cumulus clouds, in the second birth and death rates dependent on a critical value of the relative humidity of the atmospheric column.

In the first set of experiments we find that for both the Betts-Miller and Kuo schemes the stochastic extension of the parameterization preserves the bulk statistics of its deterministic limit. The impact instead is strong on the statistics of the extremes of daily convective precipitation analyzed with the extreme value theory.

In the second set of experiments we find that the inclusion of the critical dependence of the activation of convection on relative humidity has a different impact with the two schemes. Even in cases in which the climatology is relatively preserved, the representation of the structure of the tropical dynamics can be deteriorated by a conflict between the natural relationship between convection and relative humidity and the one induced by the stochastic model.

These results suggest that, although promising for tackling a number of problems related to the representation of sophisticated features of atmospheric convection, applications of this kind of models to a complex GCM require to be carefully designed for the specific characteristics of the host deterministic scheme.

Contents

Abstract **iii**

1 Introduction **1**
1.1 Stochastic parameterization of convection 1
1.2 Research questions and goals 6
1.3 Thesis outline 8

2 Parameterization of atmospheric convection **11**
2.1 Large-scale effects of atmospheric convection 11
2.2 Parameterization of atmospheric convection 13
2.2.1 General structure of a parameterization scheme 13
2.2.2 Adjustment schemes 14
2.2.3 Moisture budget schemes 15

3 Stochastic lattice-gas model of a cloud system **17**
3.1 Stochastic cloud population dynamics 17
3.1.1 Reduced stochastic model 17
3.1.2 Minimal model: binary system of non-interacting elements . . 23
3.1.3 Numerical test 26
3.2 Coupling strategy 30
3.2.1 Stochastic extension of host deterministic parameterization . . 30
3.2.2 Coupling to the BM scheme 32
3.2.3 Coupling to the Kuo scheme 33
3.3 Summary and plan of the experiments 34

4 Experiments with deterministic parametrizations **37**
4.1 Experimental settings 37
4.2 Results 40
4.3 Discussion and conclusions 47

5 Experiments with fixed transition rates **49**
5.1 Experimental settings 49
5.2 Results 53
5.3 Discussion and conclusions 57

6 Experiments with non-fixed transition rates **59**
6.1 Experimental settings 59
6.2 Results 63
6.3 Discussion and conclusions 71

7 Conclusions **73**
7.1 Summary and discussion 73
7.2 Future perspectives 77

Erklärung/Declaration **89**

List of Figures

3.1 Accuracy of reduction method for minimal model 27
3.2 Test properties minimal model 28

4.1 Fixed SST profile 38
4.2 Precipitation and evaporation zonal mean 40
4.3 Precipitable water and relative humidity zonal mean 41
4.4 Zonal wind at 850 hPa and CAPE zonal mean 42
4.5 Bidimensional pdf of precipitating events 43
4.6 Hovmöller diagram of convective precipitation 44
4.7 WK spectra of precipitation for BM scheme 45
4.8 WK spectra of precipitation for Kuo scheme 46
4.9 WK spectra of 850 hPa zonal wind for BM scheme 47
4.10 WK spectra of 850 hPa zonal wind for Kuo scheme 48

5.1 Climatology of convective precipitation 53
5.2 Pdf of convective precipitation 54
5.3 GEV pdf of convective precipitation 55
5.4 Location parameter of GEV pdf of convective precipitation 56
5.5 Scale parameter of GEV pdf of convective precipitation 57

6.1 Zonal mean and standard deviation of convective precipitation 63
6.2 Zonal mean and standard deviation of 850 hPa zonal wind 64
6.3 Zonal mean and standard deviation of relative humidity 65
6.4 Zonal mean and standard deviation of cloud fraction 66
6.5 Bidimensional pdf of precipitative events for the BM scheme 67
6.6 Bidimensional pdf of precipitative events for the Kuo scheme 68
6.7 WK spectra for the BM-H09-T6-N100 experiment 69
6.8 WK spectra for the KUO-H09-T6-N100 experiment 70

List of Tables

4.1 List of experiments for deterministic test 39

5.1 List of experiments with fixed transition rates 51

6.1 List of experiments with non-fixed transition rates 62

Chapter 1

Introduction

1.1 Stochastic parameterization of convection

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 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 (parameterization).

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

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). It has been since then applied to a number of geophysical models. Chekroun et al. (2011) have recently provided an introduction to the random dynamical systems theory addressed to the geophysical community, showing how concepts of the 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 of specific climatic processes. Because of the increase in resolution of operational numerical models of weather and climate, the interest in introducing stochastic parameterizations in full GCMs in order to represent subgrid variability due to unresolved processes (which could then feed back trough 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 parameterization 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 parameterization schemes have been developed in the last decades. Classically they are divided into three families: adjustment schemes (Manabe et al., 1965; Betts and Miller, 1986), moisture budget schemes (Kuo, 1965, 1974) and mass-flux schemes developed in different versions by many authors (e.g. 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 parameterized as a response to the destabilization enslaved by the large scale dynamics. It can therefore easily be extended to include the kind of justifications on which the adjustment-like (including Kuo) schemes are based, and thus it can be considered for sake of simplicity the conceptual basis of all the parameterization schemes avail-

able. Part of the current debate on the parameterization of atmospheric convection is focused on a reconsideration of the validity, interpretation and implementation of the QE principle (Yano and Plant, 2012), both at a fundamental and operational level. One branch of this debate has led to various attempts to design stochastic parameterizations of atmospheric convection, in order to represent deviations from the QE behavior (Neelin et al., 2008; Palmer and Williams, 2010).

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; Peters and Neelin, 2006, 2009; Peters et al., 2009; Neelin et al., 2009; Peters et al., 2010) 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 neighborhood of a critical point thus dampening deviations from it, the physical interpretation is radically different and if valid it would imply that important statistical properties of convection are not captured by parameterizations 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 parameterization 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 parameterized 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) and 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 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 parameterizations 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 parameterization, 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 at the basis for the suggestion of introducing stochastic components into pre-existing convective parameterization schemes. First attempts to design a stochastic parameterization 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 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 streamfunction 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 parameterization 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), Khouider et al. (2010), Frenkel et al. (2012)). A stochastic lattice-gas model consists of a collection of N elements spatially organized following a certain geometry (for a

example on a regular square lattice in which each site has four first neighbors), each of which can be in one out 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 of its neighbors (in order to represent local interactions) and/or on external fields. In applications to convective parameterization, 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), Khouider et al. (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 a similar line, Plant (2012) has proposed a general framework for using subgrid Individual-Level models (ILM) in the context of mass-flux parameterization making use of the van Kampen system size expansion approach (van Kampen, 2007).

In general, these kind of models could in principle be useful in order to tackle a number of unsolved problems concerning the representation of organized atmospheric convection and its interaction with the large-scale tropical dynamics, like the representation of preconditioning processes in the life cycle of convective systems, the representation of the criticality behavior of precipitation conditional on the moisture field, the presence of long term memory in observed convection-related quantities, the representation of the daily cycle of convection, the double ITCZ problem, and the representation of the intraseasonal variability and the Madden-Julian oscillations.

1.2 Research questions and goals

Despite some growing interest in this approach to stochastic parameterization of atmospheric convection, no attempts have been made so far to couple such models to a full GCM. The aims of this thesis are therefore:

1) to develop a systematic methodology to treat such models in the context of convective parameterization in real, operational GCMs;

2) to provide first examples of their application in such context by performing experiments with the Planet Simulator.

Regarding point 1, two issues have to be tackled. The first problem is due to the fact that the number of lattice elements N is given by the ratio between the size of the GCM grid box L and the size of the individual convective elements l. For a medium resolution GCM for climatic application L is on the order of 10^2 Km, while l is on the order $10^{-2} - 10^0$ Km, depending on which kind of cloud system the lattice model is supposed to represent. Therefore, $N = L/l$ is supposed to be quite large, on the order of $10^2 - 10^4$. The number of grid points in a medium resolution GCM for climatic applications is on the order of $10^4 - 10^5$. Therefore, computing at each grid point the evolution of the macrostate (the S cloud fractions) of the lattice model from the direct simulation of the evolution of its microstate (the N individual elements of the lattice) would require casting order $10^6 - 10^9$ random numbers at each time step, which would be numerically untreatable. The second problem stems from the fact that many parameterization schemes do not feature explicitly the cloud or updraft fraction in their formulation, so that it is not always possible to automatically introduce the cloud fraction provided by the stochastic model into these schemes. Two ingredients are therefore needed:

1a) to develop a numerically treatable and possibly to some extent analytically understandable formulation of the evolution of the macrostate of lattice model;

1b) to design a robust coupling strategy in order to include the stochastic model into a pre-existing parameterization.

Previous works have already partially tackled these issues: here we propose an alternative that could be useful in order to extend this approach to the "stochasticization" of a convective parameterization to real GCMs in a more general fashion.

Regarding point 2, being this line of research in a very preliminary phase, the numerical experiments we have performed with the Planet Simulator (the very first attempts in the literature with a full GCM) are designed in order to present a

demonstration of the feasibility of this strategy, and to obtain informations about the basic impact of the stochastic model in simplified settings, without any claim of realism or any aim of improving at this stage the representation of specific features of the tropical dynamics. We have tested the model in its simplest configuration: a binary system (S=2) representing sites convectively inactive (clear sky) or active (clouds) without local interactions. We couple this simple on/off description of convection to the aqua-planet version of the Planet Simulator in two different ways. In the first case, we consider a model with constant transition rates, that is considering constant birth and death rates for the clouds in each grid box. In this way the stochastic model introduces only the effects of considering a demographic description of the cloud system (litterally just the fact that we are "counting" the clouds). Therefore the state of the GCM does not affect the evolution of the lattice model, and the coupling is just one-way (lattice model $\rightarrow$ GCM). In the second case, we make the birth rate of the clouds dependent on a measure of the column-integrated moisture content of the grid point, mimicking the critical behavior of the onset of precipitating convection found in several observational works. In this case the state of the GCM does affect the evolution of the lattice model, and we have a full two-way coupling (lattice model $\leftrightarrow$ GCM). With these experiments we investigate the following questions:

2a) how the climate of the GCM is affected by the introduction of the stochastic model with the one-way coupling, in particular in terms of local statistics (climatology and extremes) of convective precipitation;

2b) how the climate of the GCM is affected by the introduction of the stochastic model with the two-way coupling, in particular in terms of tropical wave propagation and local relationship between convective precipitation and moisture field.

As said, we have focused here on the methodological aspects of the introduction of such models in a convective parameterization; more complex or realistic applications of this strategy remain subject of future works.

1.3 Thesis outline

In the following we present the structure of the thesis. Part of the results presented in **Chapter 3** and **Chapter 5** have been used to produce an article currently under review in the Quarterly Journal of the Meteorological Society (Ragone et al., 2013). The results shown in **Chapter 4** and **Chapter 6** will serve as a basis for an additional publication that is currently under preparation. Note that for consistency with the format of an academic thesis the content of Ragone et al. (2013) has been reworked and redistributed across different parts of the thesis.

- in **Chapter 2** we present a brief overview on the problem of the parameterization of unresolved atmospheric convection. Starting from the basic equations of conservation, we describe the ideas behind the most important approaches to convective parameterization, discussing their properties, limitations, and common assumptions. We focus in particular on the families of parameterization schemes to which belong the ones in use in the GCM that has been used in this thesis;

- in **Chapter 3** we describe the potential benefits of using stochastic lattice-gas models to represent the dynamics of cloud systems, referring to the limited existent literature on the subject. We propose a general method to derive an approximated set of stochastic differential equations for the time evolution of the macrostate of such models, making computationally possible the coupling to a convective parameterization of a GCM. We study in detail the minimal version of the model, and we test the numerical accuracy of the method in different set-up. We then propose a general strategy to couple such models to a generic parameterization scheme, so that in the limit of space and time scale separation between the large (GCM) and small (lattice model) dynamics we recover the original deterministic version of the host scheme;

- in **Chapter 4** we analyze the climatology and the properties of the tropical wave dynamics in the aqua-planet version of the Planet Simulator in its original deterministic version. In particular we study how the behavior of the system changes when the shallow convection is switched off, since in order to couple the parameterization to the minimal version of the model we can allow only two possible convective states (clear sky and deep convection). The analysis is performed with both the BM and Kuo parameterizations, and the differences and similarities between the climates resulting from using the two schemes are discussed;

- in **Chapter 5** we perform experiments with the Planet Simulator coupling the minimal version of the stochastic model with fixed transition rates to both the BM and Kuo parameterization schemes in aqua-planet conditions. We perform a limited but comprehensive exploration of the parameter space of the stochastic model, and we show how with both schemes the introduction of the stochastic model affects only the tails of the distribution of the convective precipitation, while keeping unaltered the bulk statistics. This in a sense confirms the robustness of the coupling strategy we have designed. By performing an analysis of the changes in the daily extremes of convective precipitation following the extreme value theory, we show how the impact on the extremes is instead substantial. This raises some interesting considerations on the use of the stochastic model and in general on the study of extremes of precipitation with a GCM;

- in **Chapter 6** we perform experiments introducing in a simple way a dependence of the activation of convection on the humidity content of the atmospheric column, consistently with recent results on the onset of deep precipitating convection. Again we perform the experiments with both the BM and Kuo schemes in aqua-planet conditions, showing how the impact in this case differs substantially between the two schemes. We analyze the basic properties of the tropical wave dynamics, showing how the interaction with the natural (not induced by the stochastic model) relationship between precipitation and moisture field realized by the convective parameterization is crucial in determining the response to the introduction of the stochastic model;

Eventually in **Chapter 7** we present the summary of the thesis, we draw our conclusions and we discuss possible future lines of research.

Chapter 2

Parameterization of atmospheric convection

2.1 Large-scale effects of atmospheric convection

In this Chapter we present a brief introduction to the general problem of the parameterization of atmospheric convection. We start from the basic equations of conservation of energy, momentum and water under Reynolds averaging, and we identify the terms related to atmospheric convection that need to be parameterized. We then present the ideas behind the approaches to convective parameterization that are at the basis of the schemes currently implemented in PlaSim (Fraedrich et al., 2005; Fraedrich, 2012), the GCM that we have used in order to perform the numerical experiments in this thesis.

The exposition follows closely the derivation of Yanai et al. (1973). We start with the basic equations of conservation for energy and water

$$\begin{cases} \dfrac{\partial s}{\partial t} + \nabla \cdot s\vec{\boldsymbol{v}} + \dfrac{\partial s\omega}{\partial p} = Q_R + L(c-e) \\ \dfrac{\partial q}{\partial t} + \nabla \cdot q\vec{\boldsymbol{v}} + \dfrac{\partial q\omega}{\partial p} = e - c \end{cases} \tag{2.1}$$

where we have introduced the basic state variable of the system: the dry static energy $s = C_p T + gz$ and the water vapor content q, while $\vec{\boldsymbol{v}}$ is the the horizontal velocity field and ω is the vertical velocity in pressure coordinate p. On the right hand side we have introduced the forcing and energy conversion terms acting on the system: the radiative forcing Q_R and the evaporation and condensation rates e and c. Note that the equation for the moist static energy $h = s + Lq$ results to be

$$\frac{\partial h}{\partial t} + \nabla \cdot h\vec{\boldsymbol{v}} + \frac{\partial h\omega}{\partial p} = Q_R \tag{2.2}$$

We define the Reynolds average and deviations of a generic field X over an area A such that

$$\begin{cases} \overline{X} = \dfrac{1}{A}\displaystyle\int_A X dA dt \\ X' = X - \overline{X} \end{cases} \tag{2.3}$$

with the usual properties $\overline{\overline{X}} = \overline{X}$ and $\overline{X'} = 0$. The conservation equations for the averaged variables can be rearranged to define the apparent heating source Q_1 and apparent moisture sink Q_2

$$\begin{cases} Q_1 = \dfrac{\partial \overline{s}}{\partial t} + \overline{\nabla \cdot s\vec{\boldsymbol{v}}} + \dfrac{\partial \overline{s}\,\overline{\omega}}{\partial p} = Q_R + L(c-e) - \dfrac{\partial \overline{s'\omega'}}{\partial p} \\ Q_2 = -L\left(\dfrac{\partial \overline{q}}{\partial t} + \overline{\nabla \cdot q\vec{\boldsymbol{v}}} + \dfrac{\partial \overline{q}\,\overline{\omega}}{\partial p}\right) = L(c-e) + L\dfrac{\partial \overline{q'\omega'}}{\partial p} \end{cases} \tag{2.4}$$

We assume that small-scale eddies in the horizontal components of the velocity field are not correlated with s' and q', while the presence of convection could cause strong correlations in the vertical and therefore vertical eddy transport of s and q. Note that for historical reasons Q_1 is defined as a source of dry static energy and Q_2 is defined as a sink of moisture.

The vertical integrals of $Q_1 - Q_R$ and Q_2 between the top and bottom pressure p_T and p_0 result to be (with g the gravity acceleration)

$$\begin{cases} \dfrac{1}{g}\displaystyle\int_{p_T}^{p_0} (Q_1 - Q_R)\, dp = \frac{L}{g}\int_{p_T}^{p_0} (c-e)\, dp - \frac{1}{g}\left(\overline{s'\omega'}\right)_{p=p_0} = LP + S \\ \dfrac{1}{g}\displaystyle\int_{p_T}^{p_0} Q_2 dp = \frac{L}{g}\int_{p_T}^{p_0} (c-e)\, dp + \frac{L}{g}\left(\overline{q'\omega'}\right)_{p=p_0} = L\,(P-E) \end{cases} \tag{2.5}$$

where P is the precipitation, S is the surface flux of sensible heat and E is the evaporation at the surface. These are integral relations linking Q_1 and Q_2 with quantities measured at one boundary (the surface).

The primary goal of a convective parameterization is to compute Q_1 and Q_2 as functions of the large scale variables. In the following we describe the general structure of a parameterization scheme and the most common approaches to the problem. Note that with the same considerations it is possible to derive a term Q_3 for the momentum, but for sake of simplicity we limit ourselves to the more fundamental Q_1 and Q_2.

Note that atmospheric convection is typically classified in three prominent modes of convection: deep precipitating cumulus clouds reaching the tropopause, cumulus congestus clouds penetrating the melting layer (located around 500 hPa in the Tropics), and shallow convection that penetrates to the boundary-layer inversion layer.

These three types of convection have different dynamical and microphysical properties, and require separate treatments in the design of a parameterization scheme. In the following we will mainly refer to the parameterization of deep precipitating cumulus convection, when not specified otherwise.

2.2 Parameterization of atmospheric convection

2.2.1 General structure of a parameterization scheme

The goal of a convective parameterization scheme is to compute the effect of an ensemble of convective clouds in a model column as a function of the grid-scale variables, quantified by the terms Q_1 and Q_2. A convective parameterization scheme consists in general of three steps:

1. determine whether or not there is convective activity in the considered grid column, and of which kind - this is often called **triggering** of convection;

2. determine the vertical structure of Q_1 and Q_2 - this is normally performed by the **plume** or **cloud model** of the parameterization scheme;

3. determine the total amount of convective activity, that is the magnitude of Q_1 and Q_2, which also implies determining the total amount of energy conversion due to latent heat release - this is normally based on some hypotesis on the relationship between convective and large scale activity which strongly characterize the scheme, and it is called **closure**.

Note that, in addition to the general theory we have exposed in Section 2.1, in practice a convective parameterization scheme meant to be implemented in a real GCM needs to fulfill a number of additional requirements and to take care of a number of technical issues that are not evident from the theory.

Beside determining the contribution of unresolved convection to the heating, moistening and momentum exchange, determining the total amount of convective precipitation and hence the contributions of unresolved convective activity to the water and energy cycle (and defining a crucial part of the connection between the two), and in general in producing a realistic climate, a convective parameterization is also needed to remove instabilities, both physical and numerical, at a sufficient fast rate so that the model does not explode. Moreover, a number of ad hoc procedures have to be applied in order to avoid numerical and physical inconsistencies that are caused by the rest of the model or by the parameterization itself, and that would severely deteriorate the representation of convective activity. A convective

parameterization scheme is typically a much more complex object than what can be expected from the description of the theory it relies on, and it is the result of the operational experience of the developers as much as of the theory. This is often the case with modules included in complex numerical models of multi-scale systems, developed across many years by a large number of people, as in the case of the GCMs that are in use nowadays.

Several parameterization schemes have been proposed and implemented since the beginning of the history of numerical weather prediction and climate simulation. They are classically divided into three families:

1. **adjustment** schemes;

2. **moisture budget** schemes;

3. **mass-flux** schemes.

We give here a brief description of the general properties of the first two families of schemes, to which belong the parameterization schemes available in PlaSim. Mass-flux schemes are today the most common choice for state-of-the-art GCMs, but even in their simplest "bulk" formulation (opposed to the more general "spectral" formulation) they are quite complex, therefore for their description we redirect to classical papers on the topic (Arakawa and Schubert, 1974; Bougeault, 1985; Tiedtke, 1989; Gregory and Rowntree, 1990; Kain and Fritsch, 1990; Gregory, 1997; Bechtold et al., 2001; Kain, 2004).

2.2.2 Adjustment schemes

The convective adjustment schemes are based on the idea that convection acts in order to adjust the state of the atmosphere towards a reference profile, that is normally prescribed or computed in order to match with observations of the mean state of the tropical atmosphere (that is where the effect of convection is most important). This is of course the simplest way of representing the effect of convection on the large scale dynamics, since the properties of the reference profile are prescribed instead of being generated by a (simplified) model of the physical process. This was indeed the first approach to convective parametrization proposed with the moist convective adjustment scheme of Manabe et al. (1965). Nowadays the prototype of adjustment scheme is the more sophisticated Betts-Miller penetrative adjustment scheme (Betts and Miller, 1986), that is still in use in operational climate and weather models.

Generally speaking, adjustment schemes compute tendencies for temperature and moisture as

$$\begin{cases} Q_1 = C_p \dfrac{T_c - T}{\tau_T} \\ Q_2 = -L \dfrac{q_c - q}{\tau_q} \end{cases} \tag{2.6}$$

The relaxation time scales τ_T and τ_q are taken as constants (typically with the same value, order of few hours), or feature a weak dependence on some large scale quantity, such as horizontal wind speed at the surface or others. The reference profiles T_c and q_c are computed by algorithms which ensure the conservation of moist static energy, which means requiring that

$$\frac{1}{g} \int_{p_T}^{p_0} (Q_1 - Q_2)\, dp = 0 \tag{2.7}$$

since the result would be given by the contribution from the radiative processes and the surface-atmosphere processes, that are calculated by other parameterizations. Despite the conceptual simplicity of the adjustment idea, these algorithms can be quite involved, so that typically T_c and q_c cannot be written in a analytical form as functions of T and q. Moreover, many empirical parameters are included in these algorithms. In the adjustment schemes the complexity of the parameterized process is therefore hidden in the algorithm computing the reference profiles, and resolved with a substantial amount of ad hoc tuning.

The Betts-Miller scheme is available in PlaSim in a formulation that follows closely the original Betts and Miller (1986), although it is not the default choice and the model has received little testing with this scheme.

2.2.3 Moisture budget schemes

Moisture budget schemes are variants of the scheme originally proposed by Kuo (1965, 1974). They are based on the idea that convection acts in order to precipitate a certain fraction $(1 - \beta)$ of the moisture converging in a atmospheric column, storing the remaining β in the column. These schemes therefore postulate a very strong link between large scale moisture convergence and small scale convective activity, that is to some extent confirmed by correlations between moisture convergence and precipitation in the Tropics found in several observational works. The total amount of convective precipitation generated in a time interval dt is prescribed to be

$$Pr = (1 - \beta)\, F_q\, dt \tag{2.8}$$

where F_q is the vertically integrated moisture convergence, including both horizontal transports of moisture by the large scale dynamics and surface evaporation.

The state of the atmosphere is then relaxed towards the pseudo-adiabatic profile satisfying the constraints given by the Kuo closure and by the moist static energy conservation, which implies $Q_2 = \beta F_q \, dt$. Formally this results in the same equations as for the adjustment schemes

$$\begin{cases} Q_1 = C_p \dfrac{T_c - T}{\tau_T} \\ Q_2 = -L \dfrac{q_c - q}{\tau_q} \end{cases} \tag{2.9}$$

where T_c and q_c are the pseudo-adiabatic profiles, and τ_T and τ_q instead of being constants are computed as

$$\begin{cases} \tau_T = \dfrac{\int_0^{+\infty} C_p (T_c - T) dz}{(1-\beta) F_q dt} \\ \tau_q = \dfrac{\int_0^{+\infty} L (q_c - q) dz}{\beta F_q dt} \end{cases} \tag{2.10}$$

It is clear that formally the original version of the Kuo scheme is not different from an adjustment scheme. In a sense it is a complementary version of the adjustment idea: while in the adjustment schemes the relaxation time scales are fixed to the same value, and the reference profiles are computed in order to satisfy the physical constraints, in the Kuo scheme the reference profiles are "fixed" to the moist adiabatic ones, and the relaxation time scales are computed in order to satisfy the physical constraints, including the moisture convergence closure. In the Kuo scheme τ_T and τ_q can take very different values, differing by orders of magnitude. In this sense Arakawa (2004) critically refers to the Kuo scheme as an asynchronous adjustment scheme. Other criticisms to the Kuo scheme address the fact that the scheme is built on the idea that convection consumes water instead of potential energy, leading to a possible positive feedback where more precipitation means more moisture convergence due to surface evaporation, which leads to more precipitation and so on. Despite these criticisms, the Kuo scheme performs quite well and it is still in use in many GCMs. In many of the current implementations of the scheme the reference profile is computed with a plume model based on the mass flux approach instead of being pseudo adiabatic, so that these schemes are hybrid mass flux schemes featuring a Kuo-like moisture convergence closure.

The Kuo scheme is the default convective parameterization scheme of PlaSim. For a detailed description of the implementation of the scheme the reader can refer to the Reference Manual of the model freely available together with the code at http://www.mi.uni-hamburg.de/plasim.

Chapter 3

Stochastic lattice-gas model of a cloud system

3.1 Stochastic cloud population dynamics

Considering a grid box size of order 100 km and typical sizes of convective elements ranging from 100 m to 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 the 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 paper 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 few stochastic differential equations.

3.1.1 Reduced stochastic model

Let us describe a cloud system as a collection of N elements or sites that can be in one out of S states, each identifying a different convective regime or cloud type. For 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 the element n at time t with a S-dimensional vector $\boldsymbol{\sigma}^{nt}$, whose components are the occupation numbers of the S states, that is $\sigma_s^{nt} = 1$ if the element n is in the 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

$$p_{ss'}^{nt}(dt) = R_{ss'}^{nt} dt \tag{3.1}$$

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

$$R_{ss}^{nt} = -\sum_{s' \neq s}^{S} R_{ss'}^{nt} \tag{3.2}$$

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

$$p_{ss}^{nt} = 1 - \sum_{s' \neq s}^{S} p_{ss'}^{nt} = 1 - \sum_{s' \neq s}^{S} R_{ss'}^{nt} dt = 1 + R_{ss}^{nt} dt \tag{3.3}$$

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 $\boldsymbol{p}^{nt}$, whose components are the absolute probabilities of finding the element n in state s at time t, evolves according to the master equation

$$\frac{d\boldsymbol{p}^{nt}}{dt} = \boldsymbol{R}^{nt} \boldsymbol{p}^{nt} \tag{3.4}$$

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 to activate convection: for example, we expect the probability of having deep precipitating convection to increase with larger CAPE and viceversa. 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 depend also on the state of its neighborhood, due to the fact that the existence of a cloud in a certain point of the lattice will activate processes influencing the probability of having other

clouds in the area nearby, either in a cooperative or competitive sense. Clustering effects are indeed observed in studies of cumulus clouds life cycle, as a consequence of mesoscale processes leading to positive near-neighbor feedbacks (e.g. Mapes (1993); Redelsperger et al. (2000); Tompkins (2001); Houze (2004); Moncrieff and Liu (2006) and references therein). Let us write in general the transition rate matrix for the element n as

$$\boldsymbol{R}^{nt}\left(\boldsymbol{x}^t, \boldsymbol{\sigma}^t\right) = \boldsymbol{F}^t\left(\boldsymbol{x}^t\right) + \boldsymbol{J}^{nt}\left(\boldsymbol{\sigma}^{n't} \in \Lambda^n\right) \tag{3.5}$$

where $\boldsymbol{F}^t\left(\boldsymbol{x}^t\right)$ is the same for each element of the lattice and represents the effect of the large scale conditions defined by the GCM state vector $\boldsymbol{x}^t$, 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 neighborhood Λ^n, whose range will depend on the nature of the processes involved.

Let us define now the cloud area fraction vector as (Khouider et al., 2010; Stechmann and Neelin, 2011)

$$\boldsymbol{\sigma}^t = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{\sigma}^{nt} \tag{3.6}$$

We would like to have an evolution equation for the process $\boldsymbol{\sigma}^t$ that does not involve the computation of all the individual processes $\boldsymbol{\sigma}^{nt}$. This can be achieved by taking a mean-field description of the system and subsequently applying the central limit theorem to the process $\boldsymbol{\sigma}^t$, assuming an expansion for large N (implying a certain degree of space scale separation) like the one of the van Kampen approach (van Kampen, 2007).

The mean-field approximation is a standard tool in statistical mechanics and population dynamics, and it 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 in neglecting the correlations between the individual processes

$$\langle \sigma_s^{nt}\sigma_{s'}^{n't}\rangle \approx \langle \sigma_s^{nt}\rangle\langle \sigma_{s'}^{n't}\rangle, \qquad \forall s, s', \quad n \neq n' \tag{3.7}$$

and in 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 $\boldsymbol{J}^t\left(\boldsymbol{\sigma}^t\right)$ which replaces each $\boldsymbol{J}^{nt}\left(\boldsymbol{\sigma}^{n't} \in \Lambda^n\right)$. In this way we obtain a mean-field transition matrix

$$\boldsymbol{R}^t\left(\boldsymbol{x}^t, \boldsymbol{\sigma}^t\right) = \boldsymbol{F}^t\left(\boldsymbol{x}^t\right) + \boldsymbol{J}^t\left(\boldsymbol{\sigma}^t\right) \tag{3.8}$$

valid for each element of the lattice. The mean-field approximation yields, of course, exact results (in the thermodynamic limit of infinite N) in the trivial case of a system of non interacting elements (where it is not an approximation at all), while in systems characterized by interactions leading to critical behaviours it gives poor results close to the critical points, because of the divergence of the decorrelation lengths that invalidates its first assumption. Its applicability will therefore depend on the specific form of the interaction terms.

Applying this approximation it is possible to derive a stochastic differential equation for the time evolution of the cloud area fraction $\boldsymbol{\sigma}^t$. 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 $d\boldsymbol{\sigma}^t$ from time t to time $t + dt$ are then uniquely determined. Since in mean-field approximation the individual processes can be treated as if they were independent random variables, thanks to the central limit theorem, the process $d\boldsymbol{\sigma}^t$ is normally distributed if N is sufficiently large

$$d\boldsymbol{\sigma}^t - \langle d\boldsymbol{\sigma}^t \rangle = \frac{1}{N}\sum_{n=1}^{N} d\boldsymbol{\sigma}^{nt} - \langle d\boldsymbol{\sigma}^t \rangle \to \frac{1}{\sqrt{N}}\mathcal{N}(\mathbf{0}, \boldsymbol{C}^t) \tag{3.9}$$

The process $d\boldsymbol{\sigma}^t$ is therefore completely described by its expectation value $\langle d\boldsymbol{\sigma}^t \rangle$ and its covariance matrix $\boldsymbol{C}^t$, and it can be written at each t as

$$d\boldsymbol{\sigma}^t = \langle d\boldsymbol{\sigma}^t \rangle + \epsilon_N \boldsymbol{\eta}^t \tag{3.10}$$

where $\epsilon_N = N^{-1/2}$ and $\boldsymbol{\eta}^t$ is a Gaussian random vector with zero expectation value and covariance matrix $\boldsymbol{C}^t$. Making again use of the mean-field approximation one can show that both $\langle d\boldsymbol{\sigma}^t \rangle$ and $\boldsymbol{C}^t$ scale with dt and can be expressed as functions of $\boldsymbol{R}^t$ and $\boldsymbol{\sigma}^t$.

Knowing that the state of the system at time t is $\boldsymbol{\sigma}^t$ (and knowing the configuration of the lattice, that is the value of each $\boldsymbol{\sigma}^{nt}$), the expectation value of $d\boldsymbol{\sigma}^t$ is componentwise given by (neglecting terms $O(dt^2)$)

$$\langle d{\sigma_s}^t \rangle_{\boldsymbol{\sigma}^t} = \frac{1}{N}\sum_{n=1}^{N} \langle d\sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t} = \frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{S} (p_{si}^{nt} - \delta_{si})\sigma_i^{nt} = \frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{S} R_{si}^{nt}\sigma_i^{nt} dt \tag{3.11}$$

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 we mean knowing the exact configuration of the entire lattice, although we use the symbol $\boldsymbol{\sigma}^t$ for simplicity of notation). In 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

$$\langle d\sigma_s{}^t \rangle_{\boldsymbol{\sigma}^t} = \sum_{i=1}^{S} R_{si}^t \sigma_i^t dt \tag{3.12}$$

or, in vectorial notation

$$\langle d\boldsymbol{\sigma}^t \rangle_{\boldsymbol{\sigma}^t} = \boldsymbol{R}^t \boldsymbol{\sigma}^t dt \tag{3.13}$$

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

$$\frac{1}{N} C_{ss'}^t = \langle d\sigma_s^t d\sigma_{s'}^t \rangle_{\boldsymbol{\sigma}^t} - \langle d\sigma_s^t \rangle_{\boldsymbol{\sigma}^t} \langle d\sigma_{s'}^t \rangle_{\boldsymbol{\sigma}^t} = \frac{1}{N^2} \sum_{n=1}^{N} \sum_{n'=1}^{N} [\langle d\sigma_s^{nt} d\sigma_{s'}^{n't} \rangle_{\boldsymbol{\sigma}^t} - \langle d\sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t} \langle d\sigma_{s'}^{n't} \rangle_{\boldsymbol{\sigma}^t}] \tag{3.14}$$

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

$$\langle d\sigma_s^{nt} d\sigma_{s'}^{n't} \rangle_{\boldsymbol{\sigma}^t} \approx \langle d\sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t} \langle d\sigma_{s'}^{n't} \rangle_{\boldsymbol{\sigma}^t} \tag{3.15}$$

so that we remain only with the terms for which $n = n'$

$$C_{ss'}^t = \frac{1}{N} \sum_{n=1}^{N} [\langle d\sigma_s^{nt} d\sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t} - \langle d\sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t} \langle d\sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t}] \tag{3.16}$$

Expliciting the increments and rearranging some terms the first term in the sum is given by

$$\langle d\sigma_s^{nt} d\sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t} = \langle \sigma_s^{nt+dt} \sigma_{s'}^{nt+dt} \rangle_{\boldsymbol{\sigma}^t} - \langle \sigma_s^{nt} \sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t} - \langle \sigma_s^{nt} d\sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t} - \langle \sigma_{s'}^{nt} d\sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t} \tag{3.17}$$

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, we have that the first two terms result

$$\langle \sigma_s^{nt+dt} \sigma_{s'}^{nt+dt} \rangle_{\boldsymbol{\sigma}^t} - \langle \sigma_s^{nt} \sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t} = \delta_{ss'} \langle \sigma_s^{nt+dt} \rangle_{\boldsymbol{\sigma}^t} - \delta_{ss'} \langle \sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t} = \delta_{ss'} \langle d\sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t} \tag{3.18}$$

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

$$\langle \sigma_s^{nt} d\sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t} + \langle \sigma_{s'}^{nt} d\sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t} = p_{s's}^{nt} \sigma_s^{nt} + p_{ss'}^{nt} \sigma_{s'}^{nt} \tag{3.19}$$

Therefore, making use of equation 3.11 and equation 3.1

$$\langle d\sigma_s^{nt} d\sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t} = -R_{s's}^{nt} \sigma_s^{nt} dt - R_{ss'}^{nt} \sigma_{s'}^{nt} dt + \delta_{ss'} \sum_{i=1}^{S} R_{si}^{nt} \sigma_i^{nt} dt \tag{3.20}$$

Making use of equation 3.11 the second term of equation 3.15 can be neglected, since

$$\langle d\sigma_s^{nt} \rangle_{\boldsymbol{\sigma}^t} \langle d\sigma_{s'}^{nt} \rangle_{\boldsymbol{\sigma}^t} = \left(\sum_{i=1}^{S} R_{si}^{nt} \sigma_i^{nt} dt \right) \left(\sum_{j=1}^{S} R_{sj}^{nt} \sigma_j^{nt} dt \right) = O\left(dt^2\right) \tag{3.21}$$

and the terms in equation 3.20 are $O(dt)$. Therefore we are left with

$$C^t_{ss'} = \frac{dt}{N}\sum_{n=1}^{N}\left(-R^{nt}_{s's}\sigma^{nt}_{s} - R^{nt}_{ss'}\sigma^{nt}_{s'} + \delta_{ss'}\sum_{i=1}^{S}R^{nt}_{si}\sigma^{nt}_{i}\right) \tag{3.22}$$

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

$$\boldsymbol{C}^t = \boldsymbol{D}^t dt \tag{3.23}$$

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

$$D^t_{ss'} = -R^t_{s's}\sigma^t_s - R^t_{ss'}\sigma^t_{s'} + \delta_{ss'}\sum_{i=1}^{S}R^t_{si}\sigma^t_i \tag{3.24}$$

We see therefore that both $\langle d\boldsymbol{\sigma}^t\rangle_{\boldsymbol{\sigma}^t}$ and $\boldsymbol{C}^t$ scale with dt and can be written as functions of $\boldsymbol{R}^t$ and $\boldsymbol{\sigma}^t$

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

$$\boldsymbol{\eta}^t = \boldsymbol{G}^t d\boldsymbol{W}^t \tag{3.25}$$

where $d\boldsymbol{W}^t = \boldsymbol{W}^{t+dt} - \boldsymbol{W}^t$ is the increment of a standard multivariate Wiener process $\boldsymbol{W}^t$, so that it is a multivariate Gaussian process with covariance matrix $\boldsymbol{I}dt$, and $\boldsymbol{G}^t$ is a suitable matrix transforming the covariance matrix of the process in $\boldsymbol{C}^t = \boldsymbol{D}^t dt$. 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 $d\boldsymbol{\sigma}^t$ as

$$d\boldsymbol{\sigma}^t = \boldsymbol{R}^t\boldsymbol{\sigma}^t dt + \epsilon_N \boldsymbol{G}^t d\boldsymbol{W}^t \tag{3.26}$$

Equation 3.26 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

$$\frac{d\boldsymbol{\sigma}^t}{dt} = \boldsymbol{R}^t\boldsymbol{\sigma}^t + \epsilon_N\boldsymbol{G}^t\boldsymbol{\xi}^t \tag{3.27}$$

where $\boldsymbol{\xi}^t = d\boldsymbol{W}^t/dt$ 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 3.27 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 3.27, whose solution is the probability distribution function ρ^t of the process $\boldsymbol{\sigma}^t$, is

$$\frac{\partial\rho^t}{\partial t} = -\sum_{s=1}^{S}\frac{\partial}{\partial\sigma_s}\left(\left[\boldsymbol{R}^t\boldsymbol{\sigma}^t\right]_s\rho^t\right) + \frac{\epsilon_N^2}{2}\sum_{s=1}^{S}\sum_{s'=1}^{S}\frac{\partial^2}{\partial\sigma_s\partial\sigma_{s'}}\left(\boldsymbol{D}^t_{ss'}\rho^t\right) \tag{3.28}$$

We have therefore derived an equation for the time evolution of the macrostate of the lattice model for large N and interactions suitable to be described in mean-field approximation. In a typical convective parameterization scheme only few cloud types are considered, typically non-precipitating shallow convection and precipitating deep convection. The computational burden of the numerical integration of equation 3.27 is therefore several order of magnitudes smaller than the direct simulation of the lattice model, which makes possible the inclusion in the convective parameterization of a real GCM.

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). A part 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 the one of Plant (2012)), and therefore 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 dependance of the transition rates by 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 SDEs for the cloud fractions. This is quite an attractive feature, since it makes possible to derive some general properties of the process from the form of equation 3.28, which may be useful in order to understand to some extent the impact that it could have on the large scale dynamics, and in order to perform experiments in a controlled and systematic way.

3.1.2 Minimal model: binary system of non-interacting elements

Let us consider the minimal case of a two states system, so that $\boldsymbol{\sigma}^t = (\sigma_1(t), \sigma_2(t))$, with fixed transition rates, without any dependance on external fields and in 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 convectively inactive (clear sky) and active (clouds). In the perspective of applications to a real convective parameterization, 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 anyway attractive in a first, explorative phase of the study of the impact of this kind of models on a convective parameterization, since it introduces only the effects coming from a demographic description of the cloud system, and, as it 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 be in general (dropping from now on the time dependence in the notation)

$$\boldsymbol{R} = \begin{pmatrix} -b & d \\ b & -d \end{pmatrix} \tag{3.29}$$

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 3.24 becomes

$$\boldsymbol{D} = D \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \tag{3.30}$$

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

$$\boldsymbol{E} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$$\boldsymbol{\Lambda} = D \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} \tag{3.31}$$

$$\boldsymbol{G} = \sqrt{D} \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}$$

The evolution equation results then to be

$$\frac{d}{dt} \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix} = \begin{pmatrix} -b & d \\ b & -d \end{pmatrix} \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix} + \epsilon_N \sqrt{D} \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \tag{3.32}$$

where ξ_1 and ξ_2 are independent Gaussian white noises. It is easy to see that equation 3.32 satisfies the condition

$$\frac{d}{dt}(\sigma_1 + \sigma_2) = 0 \tag{3.33}$$

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 then 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, that is

$$\frac{d\sigma}{dt} = (1-\sigma)b - \sigma d + \epsilon_N \sqrt{(1-\sigma)b + \sigma d}\,\xi \tag{3.34}$$

The system can therefore be modeled by a single stochastic differential equation with a linear deterministic drift term and a multiplicative stochastic forcing whose intensity depends on the size of the system. Note that the noise term is very similar to the one in Tome and de Oliveira (2009), who obtained their result applying the mean field approximation and the van Kampen expansion to a specific three-states spatial model.

In order to understand the properties of the solution of the model it is more convenient to express equation 3.34 in terms of the parameters (σ_0, τ) , where

$$\sigma_0 = \frac{b}{b+d} \tag{3.35}$$

$$\tau = \frac{1}{b+d} \tag{3.36}$$

In this way we have

$$\frac{d\sigma}{dt} = \frac{\sigma_0 - \sigma}{\tau} + \epsilon_N \sqrt{\frac{\sigma_0 + \sigma(1-2\sigma_0)}{\tau}}\xi \tag{3.37}$$

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 to equation 3.37 can be written as

$$\tau \frac{\partial \rho}{\partial t} = \frac{\partial \rho D_1'}{\partial \sigma} + \frac{\partial^2 \rho D_2'}{\partial \sigma^2} \tag{3.38}$$

Since the scaled coefficients D_1' and D_2' do not depend on τ, also the stationary solution ρ_s of equation 3.38, which corresponds to the equilibrium probability distribution of the process σ, 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

$$\rho_s \propto \frac{1}{D_2'} exp\left(\int_0^\sigma \frac{D_1'}{D_2'} d\sigma\right) \tag{3.39}$$

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

$$\rho_s \propto (\sigma_0 + (1-2\sigma_0)\sigma)^{\frac{4}{\epsilon_N^2}\frac{\sigma_0(1-\sigma_0)}{(1-2\sigma_0)^2}-1} \, exp\left(-\frac{2}{\epsilon_N^2}\frac{\sigma}{1-2\sigma_0}\right) \tag{3.40}$$

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 units by a rescaling of time, and it uniquely describes the memory of the process: the (not normalized) autocorrelation function of the process σ is

$$r(t) = \langle \sigma^{t_0} \sigma^{t_0+t} \rangle - \langle \sigma^{t_0} \rangle \langle \sigma^{t_0+t} \rangle \tag{3.41}$$

where t_0 can take any value. Taking the time derivative of 3.41 and using 3.34 we have

$$\frac{dr}{dt} = -\frac{r}{\tau} \tag{3.42}$$

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 3.34. The process σ is therefore memory-less 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.

3.1.3 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: $\sigma_0 = (0.001, 0.01, 0.05, 0.1)$, $\tau = (3, 6, 12, 24)$ hours and $N = (100, 225, 400, 1000)$. These values are compatible with applications 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 the evolution of the convective events of the order of few hours, and typical horizontal length scales of the individual cumulus

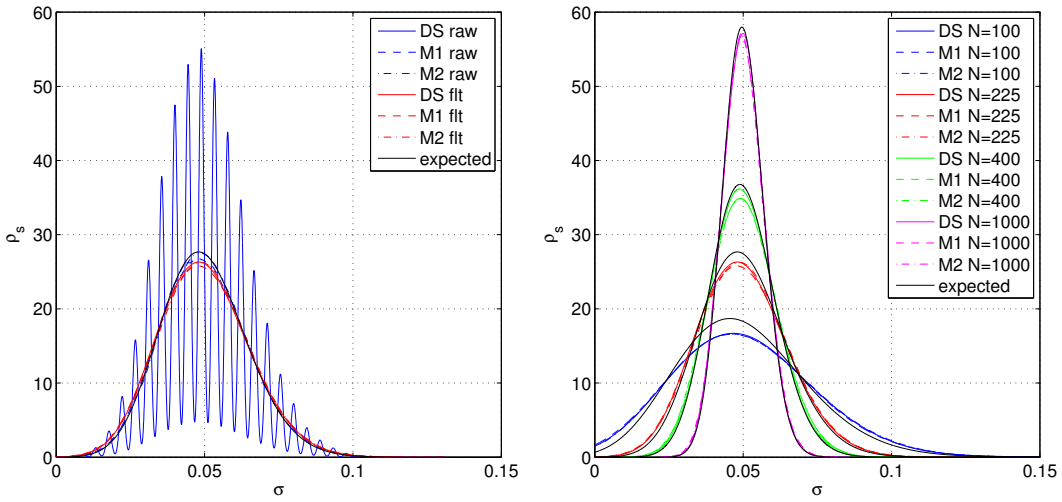

Figure 3.1: (left) Stationary distributions without and with filtering for $\sigma_0 = 0.05$, $\tau = 6$ hours and $N = 225$. (right) Stationary distributions (filtered) for different values of the noise with $\sigma_0 = 0.05$ and $\tau = 6$ hours.

clouds of the order of few km. Considering a coarse T42 resolution, equivalent to tropical grid boxes linear dimensions of about 300 km, the selected values of N correspond to linear sizes of the convective elements in the range of 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 $T = 3$ years with time step $\Delta t = 15$ minutes, starting from the initial condition $\sigma = \sigma_0$, so that no transient behavior 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 to avoid values larger than 1.

The left panel of figure 3.1 shows an important general difference between the lattice model and the SDE reduction. In the first case σ belongs to a discrete domain

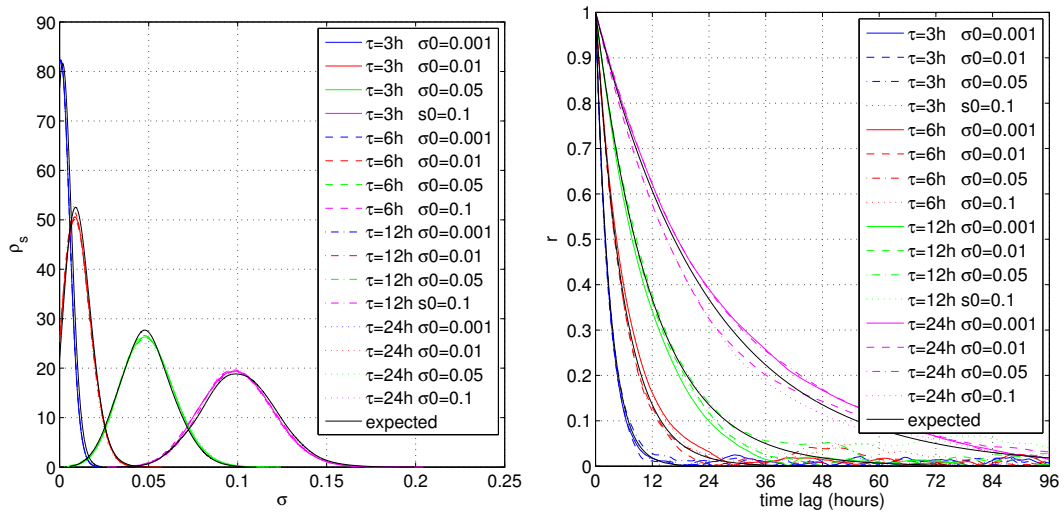

Figure 3.2: (left) Stationary distributions (filtered) for $N = 225$ and different values of σ_0 and τ. (right) Autocorrelation function of the process for $N = 225$ and different values of σ_0 and τ.

(the integer multiples of $1/N$), while in the second case it belongs to a continuos domain. This is evident from the left panel of figure 3.1, that shows in blue the stationary distributions obtained with the direct simulation of the lattice model DS and with the M1 and M2 algorithms for a case with $\sigma_0 = 0.05$, $\tau = 6$ hours and $N = 225$. This difference will be particularly strong in cases where the system hosts few active elements, that is for small values of σ_0 and N. Still, if the distributions are filtered by a factor $1/N$ (red) 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 (black). 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, reason being that this digitalization of the cloud fraction is indeed an artifact 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 the grid-box area). Instead, in a real cloud system each element will have a different size. This does not mean that the continuos version of the signal obtained with the SDE is closer to reality, but that at least the missing feature is not a real physical property of the system.

The right panel of figure 3.1 shows how the stationary distribution changes for different values of N, keeping $\sigma_0 = 0.05$ and $\tau = 3h$, with the 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), where all the distributions have been filtered with a window of width $1/N$. In black we have the expected behavior from the solution of the corresponding Fokker-Planck equation. As expected, increasing the size of the system the range of the fluctuations around σ_0 is reduced, and viceversa. When the distribution is broad enough to interact with the lower boundary, the agreement between the theory and the numerical results becomes worse, but is still acceptable in 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. We can therefore consider to be satisfied with the numerical performances 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 to be our best candidate for applications in a GCM.

Figure 3.2 shows the stationary distributions (left panel) and autocorrelation functions (right panel) obtained from the DS experiment and the theory (the results from M1 and M2 do not give additional informations 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 the autocorrelation functions for different values of σ_0. In this range of values changing σ_0 both shifts the center 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 parameterization 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 presence of interactions will depend on the degree of applicability of the mean-field approximation, and it 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 a non trivial statistics, which could already have an interesting impact when introduced in the convective parameterization of a GCM.

3.2 Coupling strategy

Once the geometry of the system and the nature of the transitions are defined, equation 3.27 determines the cloud fraction of each considered cloud type. The question is now how the cloud fractions should enter into the description of unresolved convection in a GCM. Since the aim of this paper is partly to define a strategy as general as possible to introduce the kind of models 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 parameterization scheme used in the host GCM. We also have to take into account the fact that many convective parameterization schemes do not explicitly include the cloud or updraft fraction in their formulations, and therefore it is not possible in general to directly substitute it 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 parameterization to conserve its skills in representing the bulk statistics of the convective activity, while affecting mainly higher order properties.

3.2.1 Stochastic extension of host deterministic parameterization

We consider the general case described in 3.1.1, 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. Let again $\boldsymbol{x}$ be the state vector of the resolved variables of a GCM (we do not show explicitly the time dependence from now on). We can represent its time evolution in general as $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \hat{\boldsymbol{\alpha}})$, where $\hat{\boldsymbol{\alpha}}$ is a vector containing the parameters of the parameterizations 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

$$\begin{cases} \dfrac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\alpha}(\boldsymbol{\sigma})) \\ \dfrac{d\boldsymbol{\sigma}}{dt} = \boldsymbol{R}(\boldsymbol{x}, \boldsymbol{\sigma})\boldsymbol{\sigma} + \epsilon_N \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{\sigma})\boldsymbol{\xi} \end{cases} \tag{3.43}$$

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), $d\boldsymbol{\sigma}/dt \to 0$. For these conditions the equation of the stochastic model reduces to $\boldsymbol{R}\boldsymbol{\sigma} = \boldsymbol{0}$. In case $\boldsymbol{R}$ does not depend on

$\boldsymbol{\sigma}$ then this is a linear system (where $\boldsymbol{x}$ 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 has always one and only one solution different from the trivial one. Recalling (3.4) 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 case $\boldsymbol{R}$ depends on $\boldsymbol{\sigma}$ the equation for $\hat{\boldsymbol{\sigma}}$ becomes non linear, thus possibly leading to more complicated stationary solutions, 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

$$\begin{cases} \dfrac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\alpha}(\hat{\boldsymbol{\sigma}})) \\ \boldsymbol{R}(\boldsymbol{x}, \hat{\boldsymbol{\sigma}})\hat{\boldsymbol{\sigma}} = \boldsymbol{0} \end{cases} \tag{3.44}$$

Since the original deterministic version of the parameterization 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 parameterization converges to the original version of the scheme. This can be obtained 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, already implemented and tested version of the host scheme.

As said, 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. On the contrary the relation $\boldsymbol{\alpha}(\hat{\boldsymbol{\sigma}}) = \hat{\boldsymbol{\alpha}}$ is well defined also when $\boldsymbol{R}$ depends on time through the dependence on $\boldsymbol{x}$, as we will simply have different values of $\hat{\boldsymbol{\alpha}}$ at different times.

Note that this coupling strategy essentially results into a more complex random-parameters approach. In more simple-minded random-parameters approaches (Lin and Neelin, 2000; Bright and Mullen, 2002; Bowler et al., 2008) a parameter is represented as 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 1) the range, distribution and autocorrelation function of the resulting process are not prescribed but determined by the nature of the transition rules, 2) 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, 3) there is potentially a coupling between the statistical properties of the resulting process and the state of the GCM.

We take now as an example the coupling of the minimal version of the stochastic model described in 3.1.2 to the BM and Kuo parameterization schemes.

3.2.2 Coupling to the BM scheme

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

$$\begin{cases} Q_1 = C_p \dfrac{T_c - T}{\tau_0} \\ Q_2 = -L \dfrac{q_c - q}{\tau_0} \end{cases} \tag{3.45}$$

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 hours for a deep convection case and 3-4 hours for a shallow convection case (the actual values depend on the resolution of the model).

For consistency with the formulation of the two-states 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 on a QE assumption, realizes in practice an exponential decay of a 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)), that is

$$\frac{dA}{dt} = -\frac{A}{\tau_0} \tag{3.46}$$

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 parameterization is to define a new relaxation time-scale

$$\tau_S = \frac{\sigma_0}{\sigma}\tau_0 \tag{3.47}$$

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 parameterization 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 viceversa. 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.

3.2.3 Coupling to the Kuo scheme

As we have discussed in the previous chapter, the Kuo scheme is essentially an adjustment scheme towards the pseudo adiabatic profile on time scales defined locally in time by the amount of moisture convergence occurring in the atmospheric column. The coupling to the minimal version of the model can therefore be easily designed as for the BM scheme, with some additional considerations. In this case the exponential decay of the equivalent of the cloud work function is done as

$$\frac{dA}{dt} = -\frac{A}{\tau_{0T}} \tag{3.48}$$

where we recall the definition of the relaxation time-scales for temperature and moisture τ_{0T} and τ_{0q}

$$\begin{cases} \tau_{0T} = \dfrac{\int_0^{+\infty} C_p(T_c - T)dz}{(1-\beta_0)F_q dt} \\ \tau_{0q} = \dfrac{\int_0^{+\infty} L(q_c - q)dz}{\beta_0 F_q dt} \end{cases} \tag{3.49}$$

It is easy to see that, in order to have an effective time scale for A

$$\tau_T = \frac{\sigma_0}{\sigma}\tau_{0T} \tag{3.50}$$

maintaing at the same time the constraints of energy and water conservation, it is sufficient to define

$$\begin{cases} \tau_T = \dfrac{1-\beta_0}{1-\beta}\tau_{0T} = \dfrac{\int_0^{+\infty} C_p(T_c - T)dz}{(1-\beta)F_q dt} \\ \tau_q = \dfrac{\beta_0}{\beta}\tau_{0q} = \dfrac{\int_0^{+\infty} L(q_c - q)dz}{\beta F_q dt} \end{cases} \tag{3.51}$$

defining the effective Kuo parameter

$$\beta = 1 - \frac{\sigma}{\sigma_0}(1-\beta_0) \tag{3.52}$$

Formally therefore we maintain the same structure of the equations for both Q_1 and Q_2, only we substitute the term β_0 how it would be calculated at each time-step by the deterministic parameterization with its modified version β. Note that, while $\beta_0 \in [0,1]$, now $\beta \in (-\infty, 1]$. Remembering that β is the fraction of moisture convergence that is stored in the atmospheric column while the rest is turned into precipitation, this means that, while the modified version still (correctly) does not create moisture out of nowhere (as it would be for $\beta > 1$), now it is possible to precipitate more than what is converging into the column (when $\beta < 0$), actually extracting moisture from the environment.

3.3 Summary and plan of the experiments

In this chapter we have described our proposal for a stochastic parameterization of a cloud system in a GCM. Summarizing, we have:

- proposed the usage of a sub-grid stochastic-lattice gas model for the description of a cloud system in a GCM grid-box;
- derived a general method to reduce the full stochastic lattice-gas model to a system of few SDEs;
- tested numerically the reduction method in few cases suitable for applications in real GCMs;
- proposed a general strategy to couple the stochastic model to a parameterization scheme, so that the coupled stochastic scheme converges to the original deterministic version of the host scheme in the asymptotic limit of space and time scale separation between the large and small scale dynamics;

- described in detail how to perform the coupling to the Betts-Miller and the Kuo schemes.

In the next chapters we will describe the results of the experiments performed coupling the minimal version of the stochastic model to PlaSim with both the BM and Kuo parameterizations:

- in chapter 4 we characterize the climate produced by PlaSim with the standard deterministic versions of the schemes, that will be our reference state;
- in chapter 5 we perform experiments coupling PlaSim to the minimal version of the stochastic model with fixed transition rates, realizing therefore only a one-way coupling;
- in chapter 6 we perform experiments coupling PlaSim to the minimal version of the stochastic model with transition rates dependent on the value of the relative humidity of the atmospheric column, realizing in this way a full two-way coupling;

Finally in chapter 7 we summarize the content of the thesis and we draw our conclusions.

Chapter 4

Experiments with deterministic parametrizations

4.1 Experimental settings

In this chapter we analyze the behavior of the system in the deterministic case, in order to have a reference to which compare the results of the experiments with the stochastic model. We test different configurations of the convective parameterization, comparing the standard ones with modified versions more suited to be coupled in a simple way to the stochastic model. This preliminary analysis is needed in order to understand how the system is affected by these modifications of the schemes, before the introduction of the stochastic model. The analysis is focused on the tropical region, where convection is the dominant driver of the atmospheric dynamics, using a standard setup for testing convective parameterization schemes in AGCMs (Neale and Hoskins, 2001a). In the following we describe the experimental settings we have chosen, the characteristics of the AGCM used for the experiments, and the modification introduced in the parameterization schemes. Then we show the results of the simulations, highlighting the changes due to these modifications. Finally we present the conclusions and we discuss the issues related to the experiments with the stochastic model, that are analyzed in chapters 5 and 6.

The numerical model applied in 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 parameterization techniques (Seiffert et al., 2006). The primitive equations are solved by the spectral transform method (Eliasen et al., 1970; Orszag, 1970).

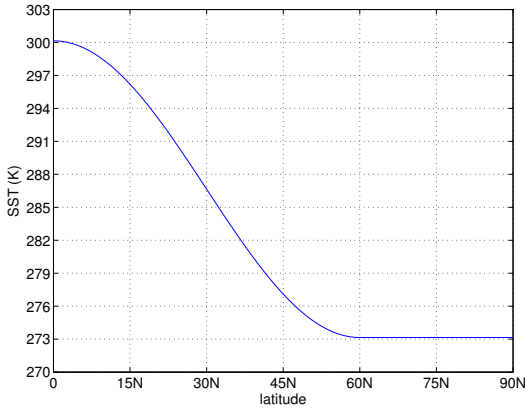

Figure 4.1: Profile of imposed SST used in the experiments.

Parameterizations include long and shortwave 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; Roeckner et al., 1992). Stratiform precipitation is generated in supersaturated states. In the standard setup the Kuo scheme (Kuo, 1965, 1974) is used for deep moist convection while shallow cumulus convection is parameterized by means of vertical diffusion. Alternatively the Betts-Miller (Betts and Miller, 1986) and the Manabe (Manabe et al., 1965) schemes are available, even if the model has never been properly tested with neither of the two.

As experimental settings we have taken the standard setup proposed by Neale and Hoskins (2001a). The model is run in aqua-planet conditions with T42 horizontal resolution and 10 σ-levels in the vertical, with a timestep of 15 minutes. The SST of the model is fixed following the control distribution in Neale and Hoskins (2001a):

$$T_s(\lambda,\phi) = \begin{cases} 27\left(1 - sin^2(\frac{3\phi}{2})\right) & : |\phi| \leq \frac{\pi}{3} \\ 0 & : |\phi| > \frac{\pi}{3} \end{cases} \tag{4.1}$$

where λ and ϕ represent longitude and latitude respectively. The model is run under perpetual equinoctial conditions and without 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 is driven by zonally symmetric boundary conditions.

Experiment	Scheme	Shallow convection	Version
BM-Sh-Det	BM	yes	deterministic
BM-Det	BM	no	deterministic
KUO-Sh-Det	Kuo	yes	deterministic
KUO-Det	Kuo	no	deterministic

Table 4.1: List of experiments for the deterministic test.

This setup has already been explored with a previous version of PlaSim by Dahms et al. (2011), using the standard Kuo parameterization. In each experiment the model is run for 26 years of integration and the analysis is limited to the last 25 years in order to account for the spin-up.

We have performed four experiments, running the model with the BM parameterization with (BM-Sh-Det) and without (BM-Det) shallow convection, and with the Kuo parameterization with (KUO-Sh-Det) and without (KUO-Det) shallow convection. In the case of Kuo, that is the default parameterization scheme in use in PlaSim, the standard setup of the parameterizations includes shallow convection, while in the case of BM the standard setup does not. As said, the BM scheme has never been properly tested in PlaSim. In any case, since for sake of simplicity we have decided to consider a two-state stochastic model to be coupled to the convective parameterization, switching off the shallow convection, we need to know how the climate looks like in absence of shallow convection also in the deterministic case, with particular focus on the tropical variability. 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 for the simplest possible conditions, and we are not concerned about realism or specific aspects of tropical dynamics, which remain subjects of later works.

Some of the quantities presented in this and in the next chapters are taken from the standard daily output of PlaSim. When a higher temporal resolution was needed, a shorter run of 3 months have been performed, saving the time-step output. Other quantities characteristic of convection but not considered in the standard output have been computed with an in-line diagnostics that has been developed during the PhD. All the quantities are computed at each time-step and accumulated during the run in order to produce an output with the same temporal resolution as the output of PlaSim set by the user.

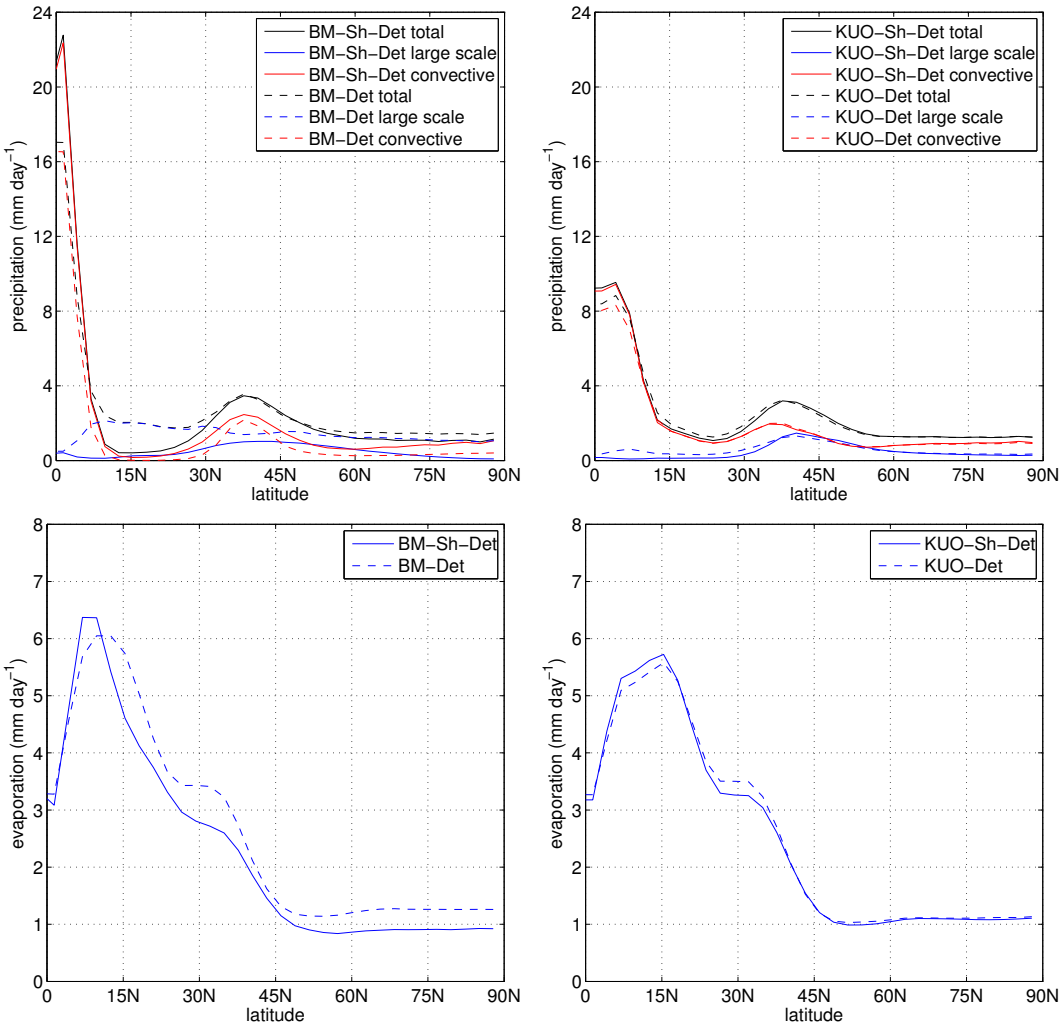

Figure 4.2: Zonal mean of precipitation (top) and evaporation (bottom) for the BM (left) and Kuo (right) scheme. Full and dotted lines correspond to the case with and without shallow convection respectively.

4.2 Results

Figures 4.2, 4.3 and 4.4 show for our set of experiments the climatological zonal mean of the most common quantities characterizing convective activity. Precipitation, evaporation and zonal component of horizontal wind at 850 hPa are taken from the standard output of PlaSim. Precipitable water (actual and at saturation), relative humidity and CAPE are computed with the in-line diagnostics. With saturation precipitable water we mean the precipitable water that the atmospheric

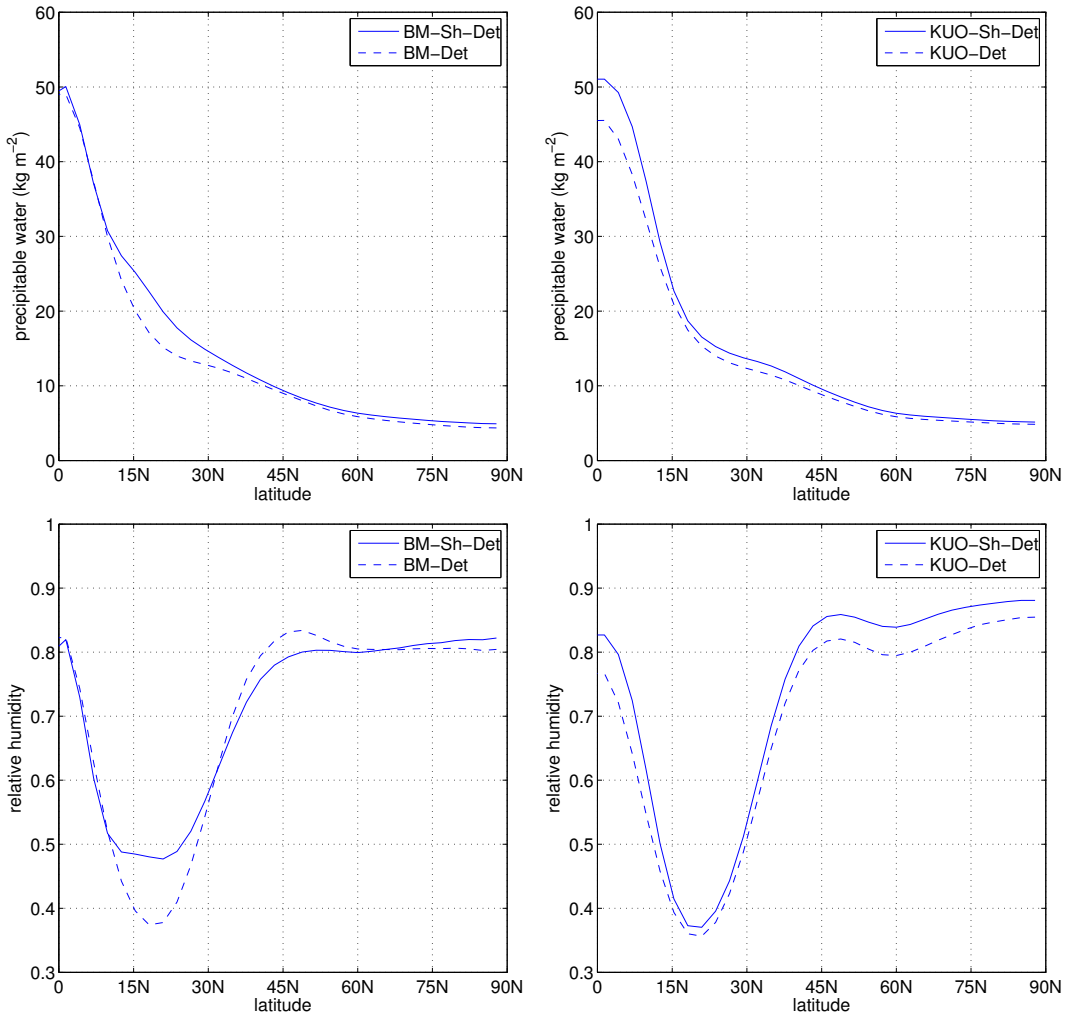

Figure 4.3: Zonal mean of precipitable water (top) and relative humidity (bottom) for the BM (left) and Kuo (right) scheme. Full and dotted lines correspond to the case with and without shallow convection respectively.

column would have if the air were at each point in conditions of saturation. The relative humidity of the atmospheric column is computed dividing the actual precipitable water by the saturation precipitation water. CAPE is computed following the standard definition from the difference between the actual temperature profile and the pseudo-adiabatic profile computed starting from the surface.

It is interesting to note that the two schemes produce a rather similar climatology for all the considered quantities but for precipitation. As we can see from figure 4.2, precipitation in the deep Tropics between 5°S and 5°N is much higher in the case of the BM scheme than in the case of the Kuo scheme. The BM scheme present a clear single maximum at the equator, while the Kuo scheme has an almost flat

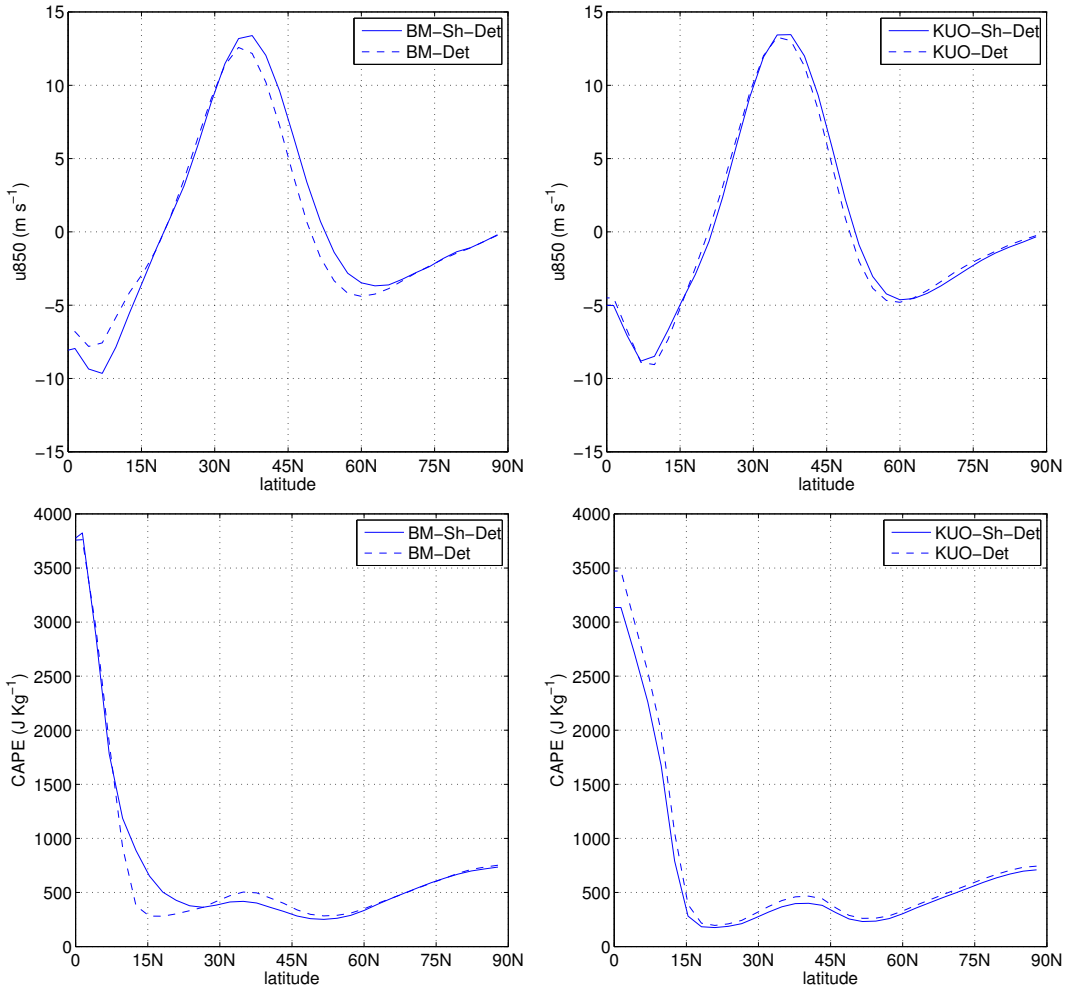

Figure 4.4: Zonal mean of 850 hPa zonal wind (top) and CAPE (bottom) for the BM (left) and Kuo (right) scheme. Full and dotted lines correspond to the case with and without shallow convection respectively.

profile in the deep Tropics, with actually a local maximum off the equator, at about 5°N. This hints at a possible double ITCZ structure not properly captured because of the relatively coarse resolution of the model. Both single and double ITCZ structures have already been found with PlaSim varying the SST profile, using the Kuo scheme with shallow convection and in the same aqua-planet equinoctial conditions, although with an older version of the model (Dahms et al., 2011). As a result, the climatological mean of precipitation found by Dahms et al. (2011) is rather different from what we obtain in 4.2 for the same setup, so that the results are not consistently comparable.

We see that in the case of Kuo switching shallow convection off has a very little

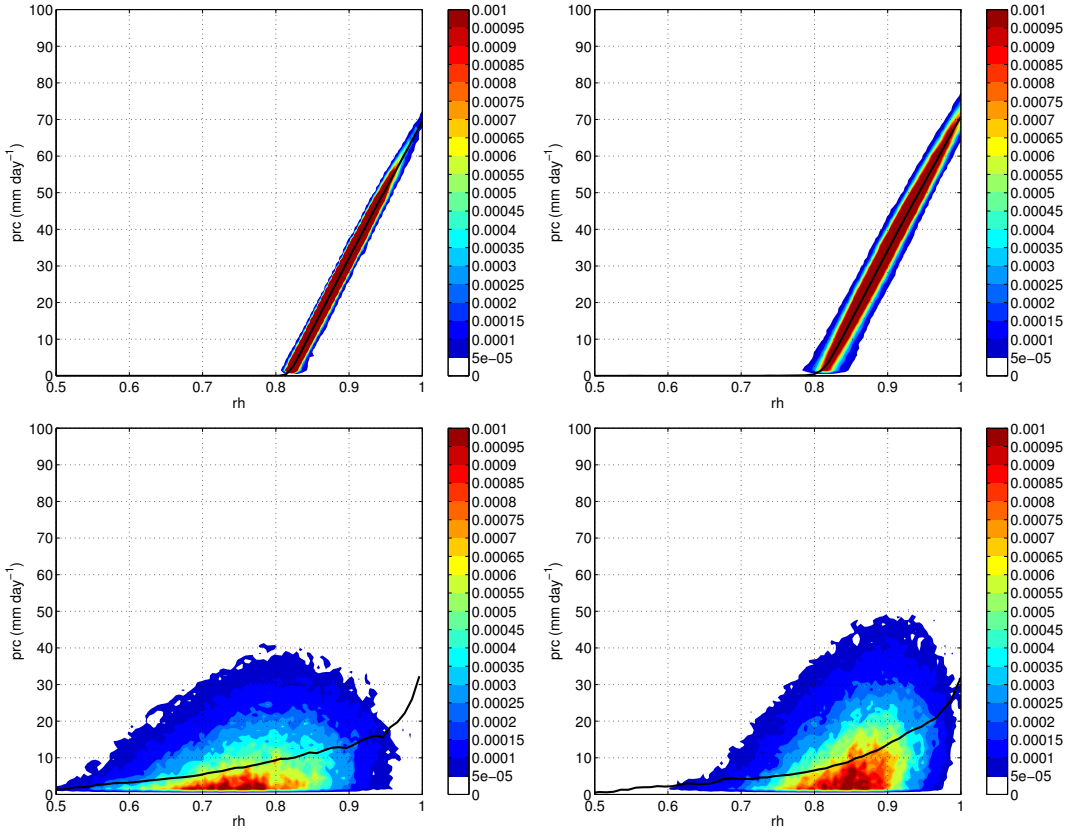

Figure 4.5: Bidimensional pdf of precipitating events in the region 10°S-10°N, as function of convective precipitation and relative humidity. In black we show the average value of precipitation conditional on the value of relative humidity. Computed from 3 months of 15 minutes data for the deterministic BM (top) and Kuo (bottom) schemes with (left) and without (right) shallow convection.

impact on the climatology. The climate becomes slightly drier, precipitation in the tropical region is slightly reduced and convective instability (measured by CAPE) slightly increased, but overall the changes are rather minor. In the case of BM the impact of removing shallow convection is larger. Precipitation in the Tropics is reduced substantially. In the subtropics the atmosphere becomes much drier and convective precipitation is reduced to zero, while large scale precipitation increases substantially everywhere but at the equator.

Figure 4.5 shows the bidimensional pdf of precipitating events in the region 10°S-10°N as function of relative humidity and intensity of precipitation. The pdf has been computed using 3 months of time-step (15 minutes) data. We can see that the two schemes realize completely different relationships between relative humidity and precipitation. In the case of the BM scheme precipitation is always zero when

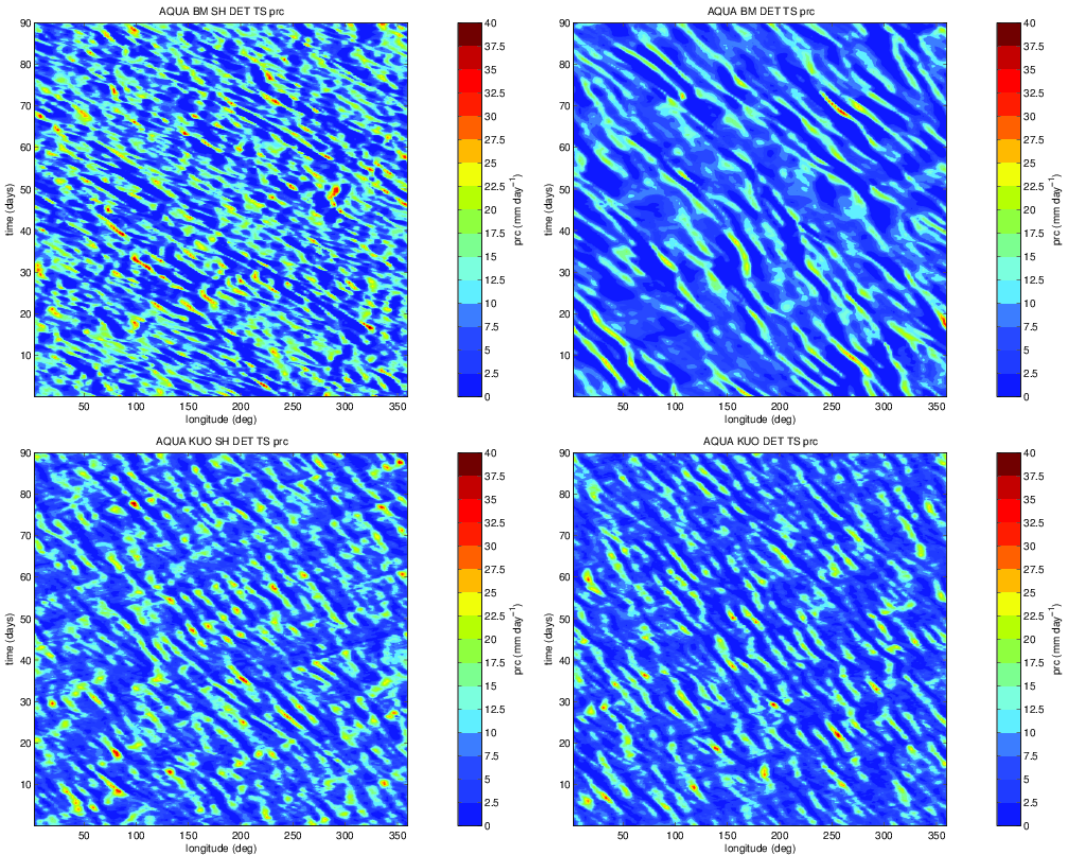

Figure 4.6: Hovmöller diagram of convective precipitation averaged in the band 10°S-10°N. Computed from 3 months of 15 minutes data for the BM (top) and Kuo (bottom) schemes with (left) and without (right) shallow convection.

the relative humidity is lower than 0.8, and then increases linearly for values higher than 0.8. The behavior is the same with and without shallow convection. The BM scheme is characterized therefore by a very strong connection between precipitation and the moisture field. In the case of the Kuo scheme such strong relationship is absent. Precipitation is still increasing with relative humidity, as one can see from the average value of precipitation conditional on the value of relative humidity (black line in figure 4.5), but the distribution in this case is rather broad, allowing a much wider range of values for the relative humidity.

Figure 4.6 shows the Hovmöller diagram of the time-step convective precipitation averaged in the region 10°S-10°N. Again it is clear that removing the shallow convection has basically no impact in the case of Kuo, while in the case of Betts-Miller the changes are more pronounced, with the disappearing of many small-scale structures. Also the speed of the precipitating systems is different, although not dramatically, with faster westward propagating systems in absence of shallow convection.

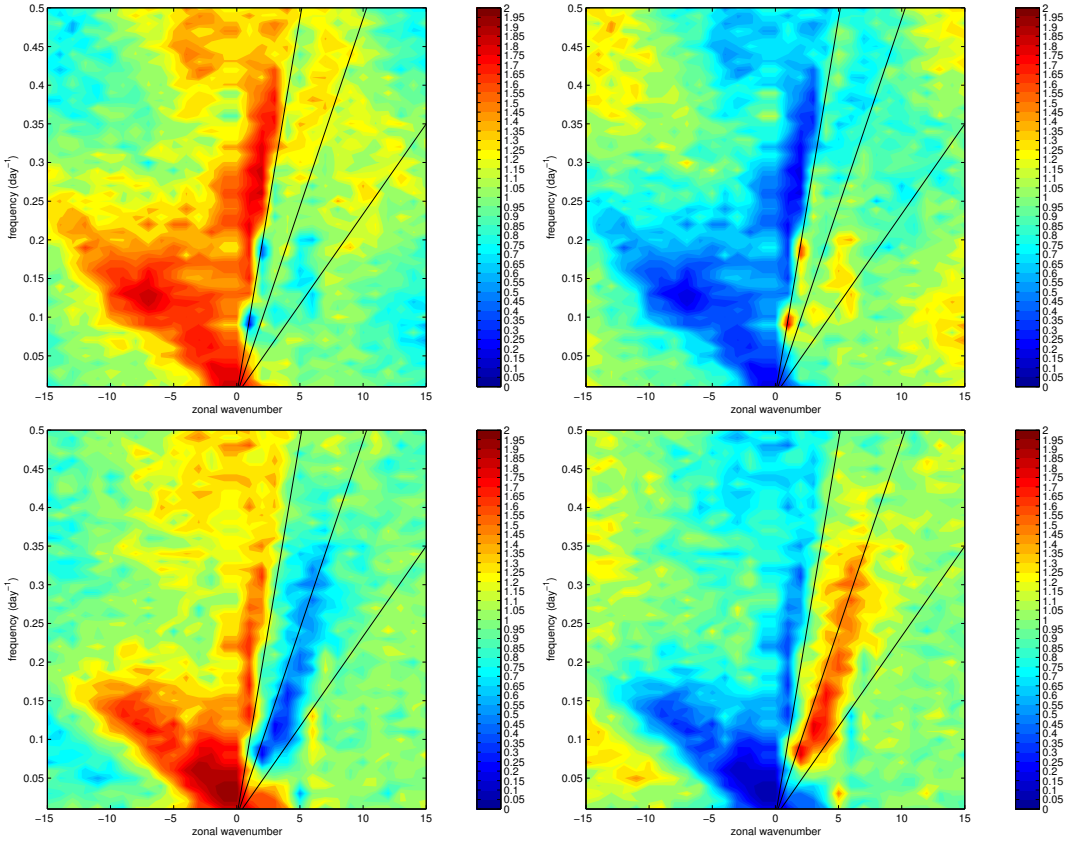

Figure 4.7: Antisymmetric (left) and symmetric (right) components of the WK spectrum of precipitation for the BM scheme with (top) and without (bottom) shallow convection. Computed from 25 years of daily data averaged in the band 10°S-10°N.

Figures 4.7, 4.8, 4.9 and 4.10 show the Wheeler-Kiladis (Wheeler and Kiladis, 1999) spectra for disturbances antisymmetric and symmetric with respect to the equator, for convective precipitation and the zonal component of the horizontal velocity at 850 hPa for both schemes, with and without shallow convection. The black lines correspond to the dispersion curves of the Kelvin waves for typical values of the phase velocity. The signature of the Kelvin waves is clear in the spectra of the symmetric part for both the variables and for all the experiments, but the signal is very weak in the case of the BM scheme with shallow convection. An evident difference between the two schemes is that in the case of Kuo we have very strong signature of westward propagating antysimmetric disturbances with wavenumber larger than 5 and period between 3 and 5 days (particularly clear in the spectra of the 850 hPa zonal wind), that is totally absent in the case of the BM scheme.

In general again the Kuo scheme shows basically no sensitivity to the removal of

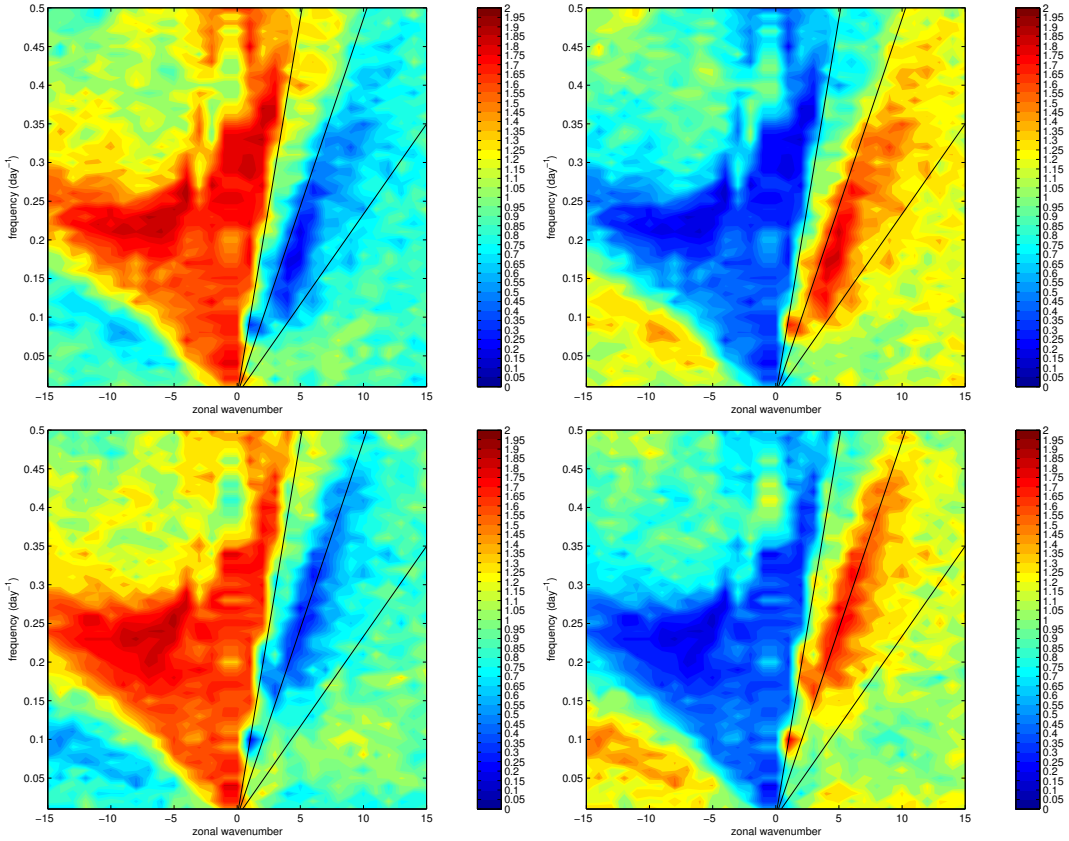

Figure 4.8: Antisymmetric (left) and symmetric (right) components of the WK spectrum of precipitation for the Kuo scheme with (top) and without (bottom) shallow convection. Computed from 25 years of daily data averaged in the band 10°S-10°N.

shallow convection, while the BM scheme seems to be more sensitive, although in terms of representation of tropical wave dynamics it seems to perform better without shallow convection. As we have said, PlaSim has received little testing with the BM scheme, so that some inconsistencies in the performances of the scheme are expected. Overall the model gives a reasonable representation of the tropical convective activity with both schemes in absence of shallow convection, so that we can safely perform our experiments in this setup.

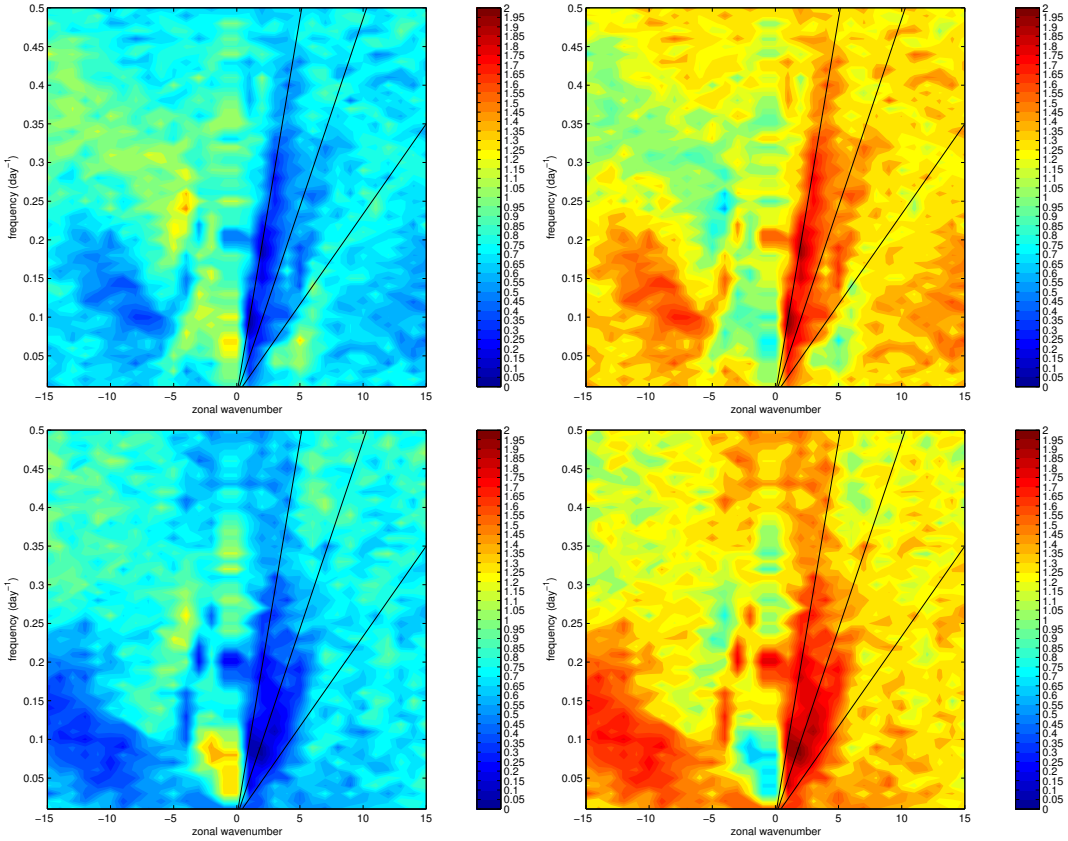

Figure 4.9: Antisymmetric (left) and symmetric (right) components of the WK spectrum of 850 hPas zonal wind for the BM scheme with (top) and without (bottom) shallow convection. Computed from 25 years of daily data averaged in the band 10°S-10°N.

4.3 Discussion and conclusions

In this chapter we have analyzed the climate produced by PlaSim with the original deterministic versions of both the BM and Kuo parameterization schemes. In particular we have checked wether the removal of the shallow convection scheme resulted in a reasonable climate or not, since for our experiments with the stochastic model we need to use this setup. The answer is positive, since the removal of shallow convection does not alter much the most important properties of the representation of convective activity in neither the two schemes. We have shown the differences in the properties of the BM and Kuo schemes in terms of statistics of the most important quantities related to the representation of convection. We have shown how the two schemes differ substantially in the link they realize between precipitation

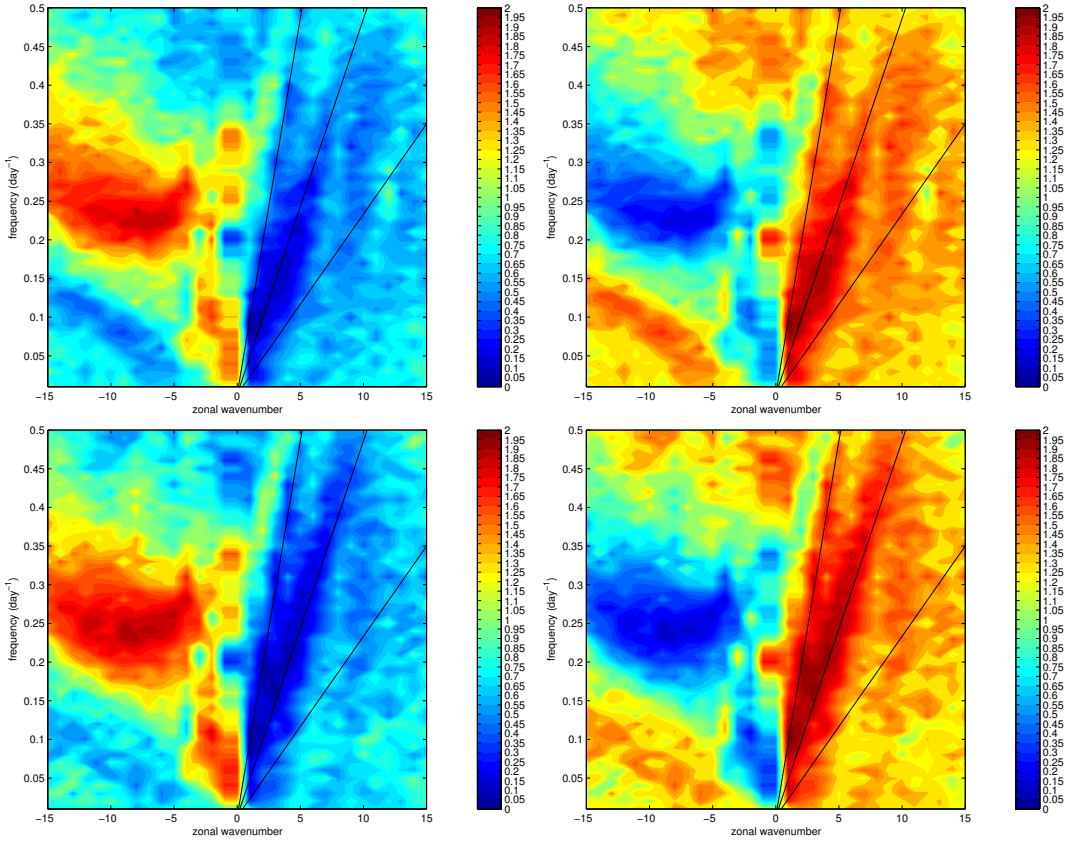

Figure 4.10: Antisymmetric (left) and symmetric (right) components of the WK spectrum of 850 hPas zonal wind for the Kuo scheme with (top) and without (bottom) shallow convection. Computed from 25 years of daily data averaged in the band 10°S-10°N.

and relative humidity, that will be a key element of the analysis shown in chapter 6. Finally, we have shown how the representation of the tropical wave dynamics is qualitatively consistent with what is normally observed or obtained with GCM simulations. Characterized the deterministic reference state, we are now ready to perform the experiments coupling with the stochastic model.

Chapter 5

Experiments with fixed transition rates

5.1 Experimental settings

In this chapter we perform a first set of experiments coupling the minimal version of the stochastic model described in chapter 3 to the aqua-planet version of the Planet Simulator. As discussed in the previous chapters, we have switched off the shallow convection in the model in order to perform the coupling with the two-state model. As a first attempt, we consider the very simple case of fixed transition rates. This in practice configures only a one-way coupling, since the state of the GCM does not feed back the stochastic model through a change in the transition rates. An example of a case of transition rates dependent on the state of the GCM, and therefore of full two-way coupling, is subject of Chapter 6. In the present case the stochastic model adds to the host parameterization only the effect coming from a demographic description of the cloud system.

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

$$\frac{d\gamma}{dt} \approx \frac{1-\gamma}{\tau} + \epsilon'_N \sqrt{\frac{1+\gamma}{\tau}}\xi \tag{5.1}$$

where $\epsilon'_N = \sigma_0^{-1/2}\epsilon_N$. This means that in the regime $\sigma_0 << 1$ (that is 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 $\hat{\sigma}_0 = 0.05$ in all the experiments. The fixed relaxation time scale takes the values $\hat{\tau} = (3, 6, 12, 24)$ hours, and the number of convective elements the values $N = (100, 225, 400, 1000)$. The analysis of the impact of the introduction of the stochastic parameterization is focused here only on the statistical properties of the convective precipitation, that is the quantity directly modified by the stochastic term. We will consider changes in the climatology of convective precipitation, that is mean and standard deviation as function of latitude, as well as changes in the daily extremes. Before showing the results we recall the basics of extreme values statistics and we describe how we performed the statistical analysis.

A particularly interesting aspect of stochastic parameterizations is the impact that they could have on the statistics of the extremes (Stechmann and Neelin, 2011) of a GCM. Deterministic parameterizations currently in use in state-of-the-art GCMs are more or less able to reproduce the climatology of the system, that is its bulk statistics, although with different levels of geographical detail and performances, depending on the complexity of the scheme and of the GCM itself. To which extent 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 events analysis in a geophysical context can be found in (Ghil et al., 2011; Sura, 2011).

The most common approach to extreme value analysis is the so called block-maxima approach. It consists in dividing a time serie 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 whose cumulative distribution function reads

$$F(x; \mu, \sigma, \xi) = exp\left\{ -\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi} \right\} \tag{5.2}$$

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 sub-families. When ξ=0 the distribution is of the Gumbel or type I kind, when ξ >0 the distribution is of the Frechét or type II kind, 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

Experiment	Scheme	Version	$\hat{\sigma}_0$	$\hat{\tau}$	N
BM-Det	BM	deterministic	-	-	-
BM-Fix-T3-N100	BM	stoch., fixed	0.05	3	100
BM-Fix-T6-N100	BM	stoch., fixed	0.05	6	100
BM-Fix-T12-N100	BM	stoch., fixed	0.05	12	100
BM-Fix-T24-N100	BM	stoch., fixed	0.05	24	100
BM-Fix-T3-N225	BM	stoch., fixed	0.05	3	225
BM-Fix-T6-N225	BM	stoch., fixed	0.05	6	225
BM-Fix-T12-N225	BM	stoch., fixed	0.05	12	225
BM-Fix-T24-N225	BM	stoch., fixed	0.05	24	225
BM-Fix-T3-N400	BM	stoch., fixed	0.05	3	400
BM-Fix-T6-N400	BM	stoch., fixed	0.05	6	400
BM-Fix-T12-N400	BM	stoch., fixed	0.05	12	400
BM-Fix-T24-N400	BM	stoch., fixed	0.05	24	400
BM-Fix-T3-N1000	BM	stoch., fixed	0.05	3	1000
BM-Fix-T6-N1000	BM	stoch., fixed	0.05	6	1000
BM-Fix-T12-N1000	BM	stoch., fixed	0.05	12	1000
BM-Fix-T24-N1000	BM	stoch., fixed	0.05	24	1000
KUO-Det	Kuo	deterministic	-	-	-
KUO-Fix-T3-N100	Kuo	stoch., fixed	0.05	3	100
KUO-Fix-T6-N100	Kuo	stoch., fixed	0.05	6	100
KUO-Fix-T12-N100	Kuo	stoch., fixed	0.05	12	100
KUO-Fix-T24-N100	Kuo	stoch., fixed	0.05	24	100
KUO-Fix-T3-N225	Kuo	stoch., fixed	0.05	3	225
KUO-Fix-T6-N225	Kuo	stoch., fixed	0.05	6	225
KUO-Fix-T12-N225	Kuo	stoch., fixed	0.05	12	225
KUO-Fix-T24-N225	Kuo	stoch., fixed	0.05	24	225
KUO-Fix-T3-N400	Kuo	stoch., fixed	0.05	3	400
KUO-Fix-T6-N400	Kuo	stoch., fixed	0.05	6	400
KUO-Fix-T12-N400	Kuo	stoch., fixed	0.05	12	400
KUO-Fix-T24-N400	Kuo	stoch., fixed	0.05	24	400
KUO-Fix-T3-N1000	Kuo	stoch., fixed	0.05	3	1000
KUO-Fix-T6-N1000	Kuo	stoch., fixed	0.05	6	1000
KUO-Fix-T12-N1000	Kuo	stoch., fixed	0.05	12	1000
KUO-Fix-T24-N1000	Kuo	stoch., fixed	0.05	24	1000

Table 5.1: List of experiments with fixed transition rates.

and the typical range of variability of the extreme events (in a loose sense they are a sort 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 the convergence of the empirical distribution of extremes obtained with the block-maxima 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 "real" distribution requires a very large amount of data, and in any case the convergence properties change 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 setup. In this conditions 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 of them as independent realizations of the same process, and put them together in oder 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 every fourth grid point is sufficient to have 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 analysis. In this way our sample of data consists of 576000 daily values. Defining a block length of 720 days, we have 800 maxima to build our statistics, which should be enough in order to perform a robust analysis.

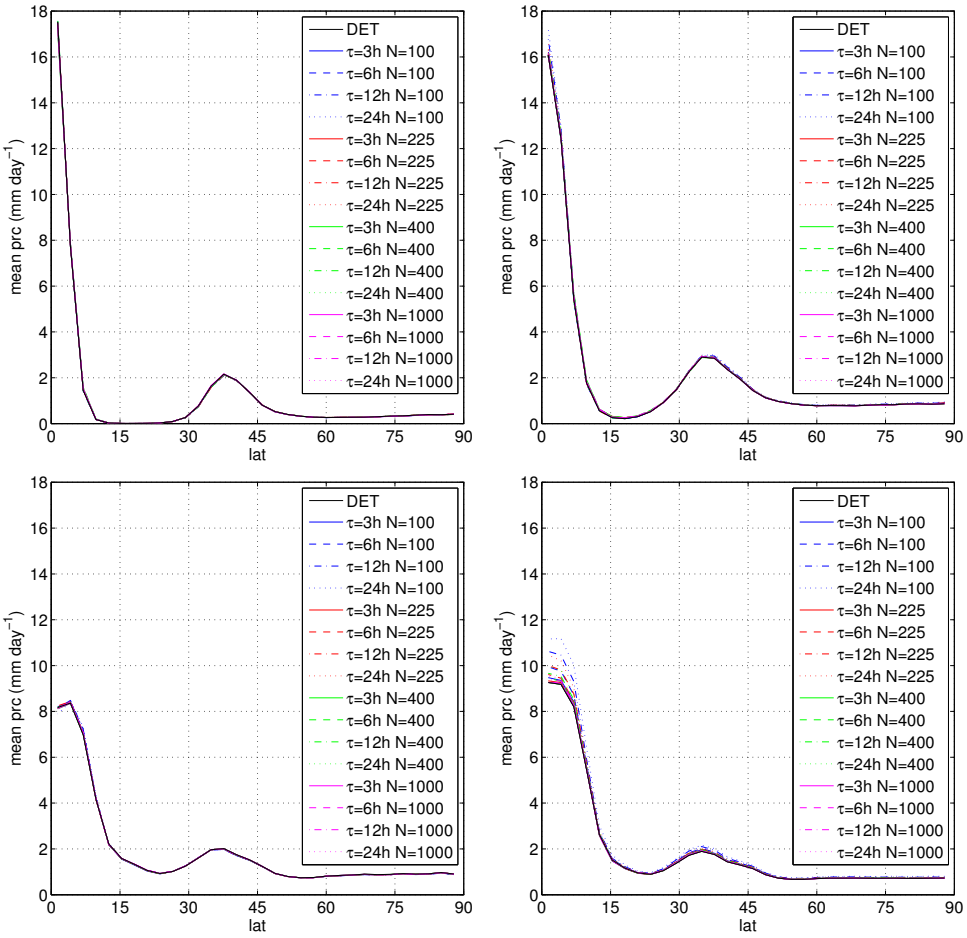

Figure 5.1: Zonal mean (left) and standard deviation (right) of daily convective precipitation, computed over 25 years of simulation after 5 years of spin-up time, for the BM (top) and Kuo (bottom) schemes. The curves refer to the deterministic (black) and stochastic (colors) cases for different values of τ and N.

5.2 Results

Figure 5.1 shows the zonal mean and standard deviation of daily convective precipitation over 25 years after 5 years of spin-up, for the standard deterministic run and for the experiments with the stochastic model, for both the BM and Kuo parameterization schemes. Because of the hemispheric and zonal symmetry of the dynamics, only the zonal quantities in the northern hemisphere are shown. When adding a stochastic term to a nonlinear system, one can in general expect a change

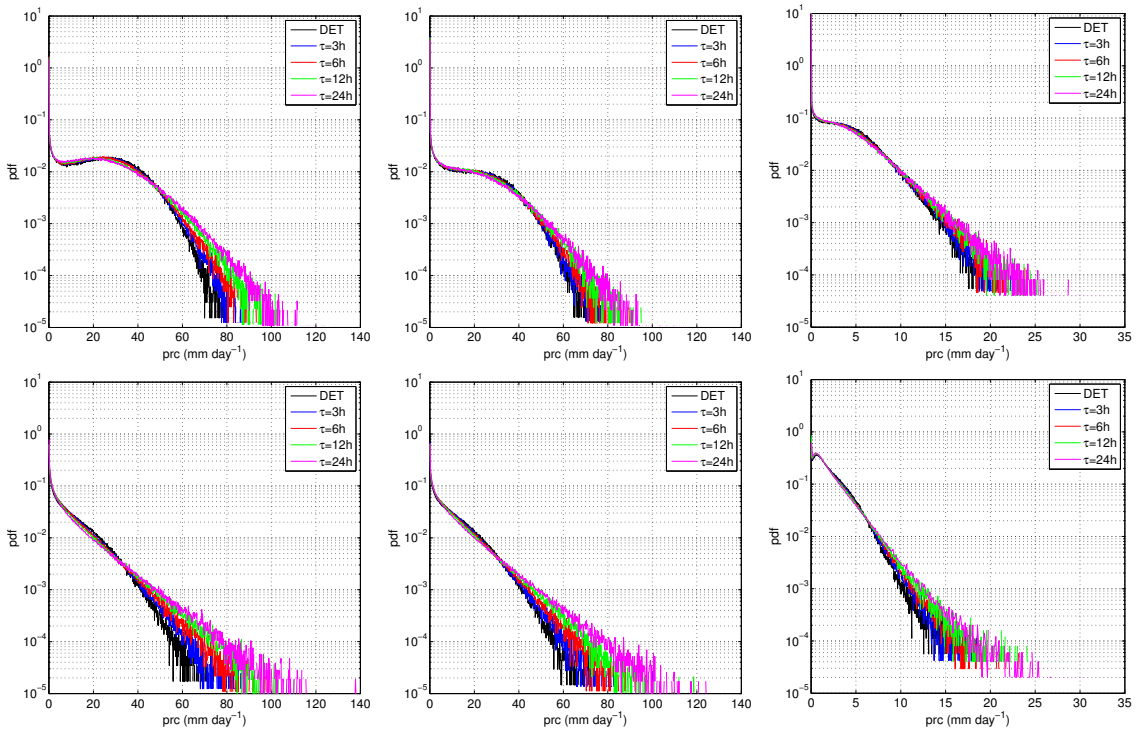

Figure 5.2: Pdf of convective precipitation at 1.5°N (left), 5.0°N (center) and 35°N (right) for the BM (top) and Kuo (bottom) schemes. The curves refer to the deterministic (black) and stochastic (colors) cases for $N = 100$ and different values of τ.

in the mean state of the system. In our case the introduction of the stochastic model does not affect the climatology of daily convective precipitation, that is the quantity directly modified by the stochastic model (and therefore the one in which one expects to see the largest impact). We can see that the mean value of convective precipitation remains exactly the same for all the experiments, at each latitude. Surprisingly, the standard deviation is also exactly the same in the case of the BM parameterization, while for the Kuo parameterization there is a slight increase concentrated in the deep Tropics below $10°N$. Therefore, the climatology of the original deterministic model is preserved by the inclusion of the stochastic model.

The impact of the stochastic parameterization is instead concentrated only on the higher moments of precipitation. Figure 5.2 shows the probability distribution function of the daily convective precipitation for three characteristic latitudes: the closest grid point to the equator (ca. 1.4°), where convective precipitation is at its maximum, . The curves refer to the deterministic run and the stochastic runs for

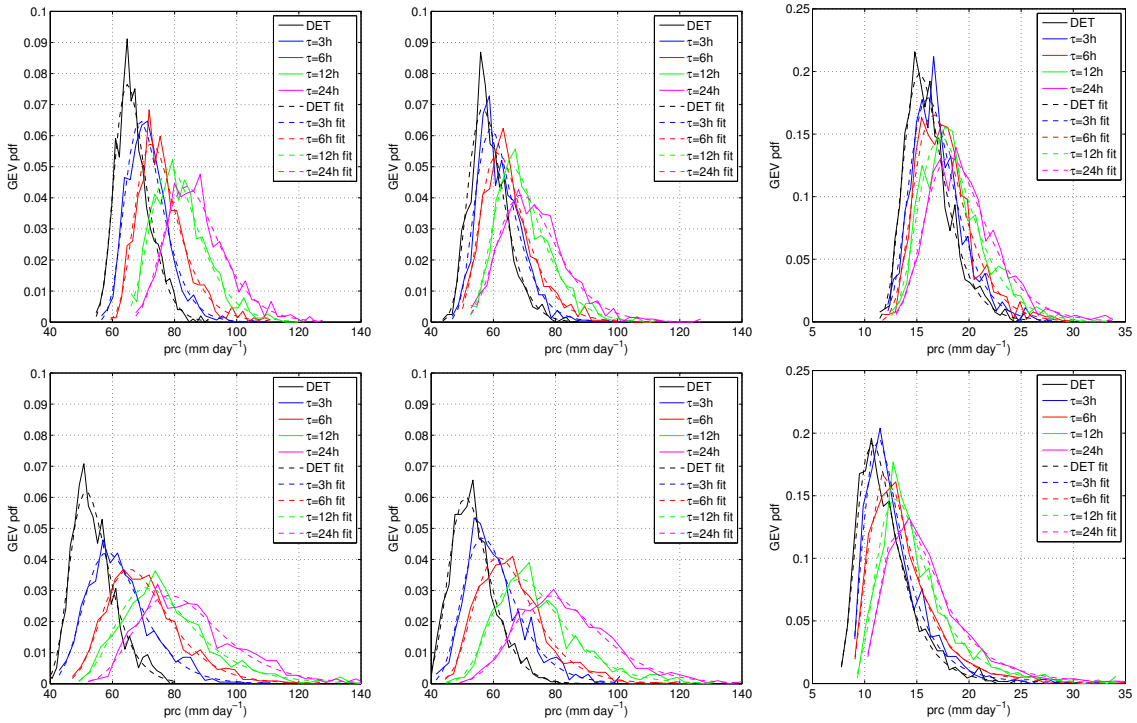

Figure 5.3: GEV pdf of convective precipitation at 1.5°N (left), 5.0°N (center) and 35°N (right) for the BM (top) and Kuo (bottom) schemes. The curves refer to the deterministic (black) and stochastic (colors) cases for $N = 100$ and different values of τ.

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 5.3 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 distribution of extreme values of convective precipitation become broader and shifted towards higher values. We can also see that the range of the GEV distributions coincides with the range over which the pdfs of Figure 5.2 differ substantially: it seems therefore that for some reason only the extreme values (in the proper statistical sense) are affected by the introduction of the stochastic model.

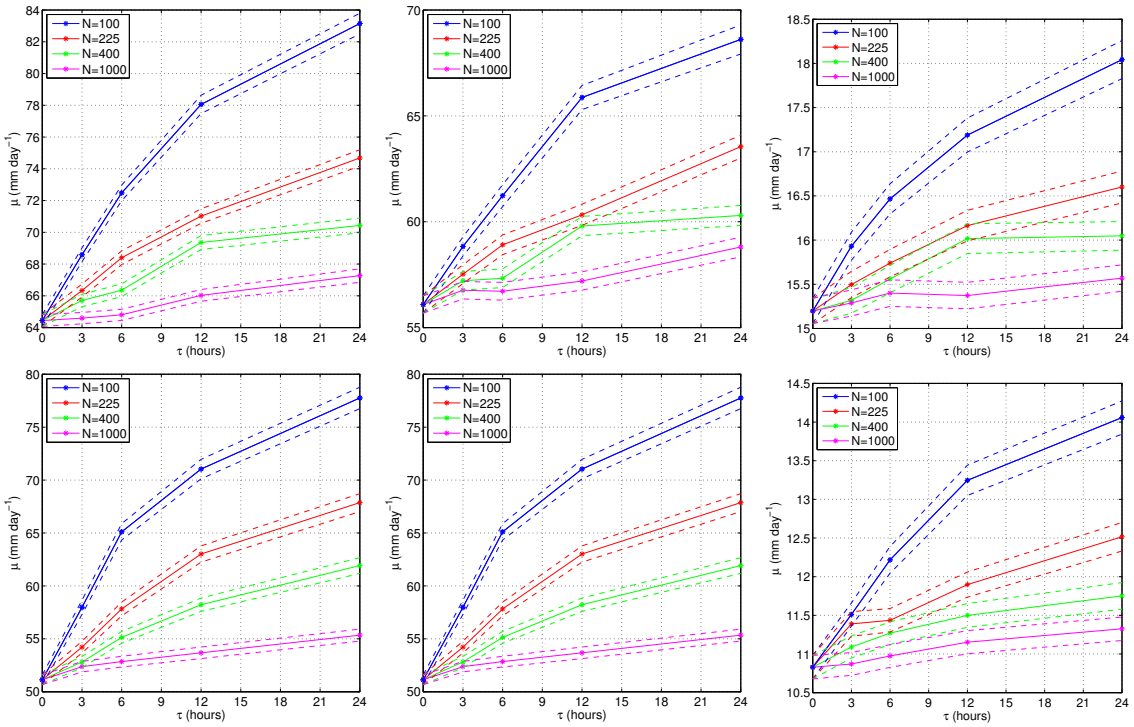

Figure 5.4: Location parameter of GEV pdf of convective precipitation at 1.5°N (left), 5.0°N (center) and 35°N (right) for the BM (top) and Kuo (bottom) schemes plotted as a function of τ. Each branch correspond to a different value of N.

In order to make the analysis more quantitative, Figures 5.4 and 5.5 show the estimates of the location and scale parameters respectively 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 at all by changes in the parameters (not shown), so that the nature of the GEV is not affected. Figures 5.4 and 5.5 represent a possible parameterization of extremes of daily convective precipitation through the parameters of our stochastic model.

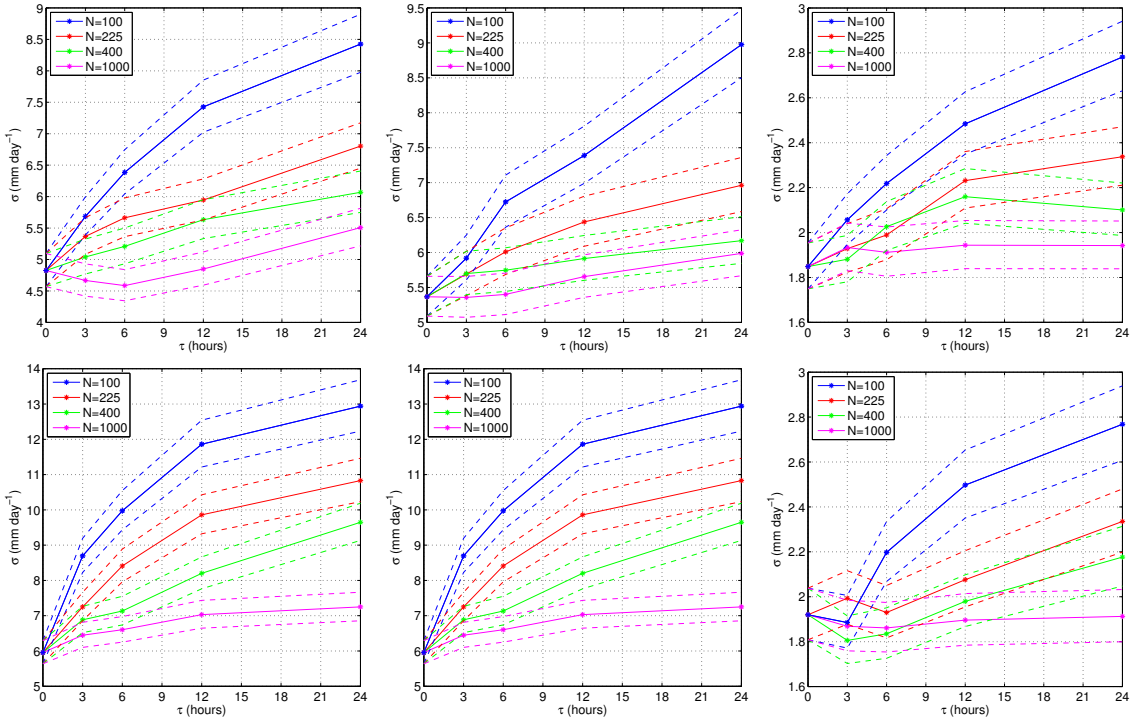

Figure 5.5: Scale parameter of GEV pdf of convective precipitation at 1.5°N (left), 5.0°N (center) and 35°N (right) for the BM (top) and Kuo (bottom) schemes plotted as a function of τ. Each branch correspond to a different value of N.

5.3 Discussion and conclusions

The analysis has focused on convective precipitation, that is the quantity directly modified by the stochastic term. In these settings the stochastic extension of the parameterization 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 the 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 extremes statistics using EVT. The location and scale parameters of the GEV distribution of tropical daily convective precipitation result to be highly sensitive to both the noise intensity and the autocorrelation time of the stochastic forcing. They increase seemingly 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 indeed 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, that is approaching space and time scale separation respectively;

- while the increase of the typical value and range of the extremes of daily convective precipitation with increasing amplitude of the fluctuations of the stochastic process is somehow expected, why these should increase with larger autocorrelation times of the noise is less clear. Lin and Neelin (2000, 2002, 2003) already showed sensitivity of tropical variability to the autocorrelation time of a stochastic forcing;

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

Chapter 6

Experiments with non-fixed transition rates

6.1 Experimental settings

In this chapter we perform experiments introducing a dependence of the transition rates on the state of the large scale model, in order to realize an effective two way coupling. As said, previous works have used quantities like CAPE, CIN and measures of the dryness of the atmospheric column as coupling quantities (Majda and Khouider (2002), Khouider et al. (2003), Khouider et al. (2010), Frenkel et al. (2012)). Inspired by a number of works on the onset of precipitating convection in the Tropics (Peters et al., 2002; Peters and Neelin, 2006, 2009; Peters et al., 2009; Neelin et al., 2009; Peters et al., 2010), recently Stechmann and Neelin (2011) have proposed in a similar context 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 a measure of the moisture content of the atmospheric column, even if in its most recent version it is based on results pointing at a novel interpretation of the properties of convection and of the interactions with its environment based on the SOC framework, is not new from the practical point of view: 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 constrain in different ways the release of latent heat due to convection) has a substantial impact on the intraseasonal variability of the model.

However, using precipitable water as such measure would be impractical for applications to global simulations, since it has been clearly shown that the critical

value of the precipitable water depends on the location (Peters et al., 2002; Peters and Neelin, 2006, 2009; Peters et al., 2009; Neelin et al., 2009; Peters et al., 2010), and therefore it would be impossible to prescribe a fixed value on the whole planet. In simple words: to prescribe globally a value reasonable for tropical regions, say 60 Kg/m^2, would imply never allowing convective precipitation to occur at midlatitudes. Also, the investigation of this property of the onset of precipitating convection is relatively recent, and from first attempts it is not clear how to identify an approximate dependence of the critical value of precipitable water to other fields, like surface temperature, that could be used in order to represent the regional variability of the parameter (Peters et al., 2002; Peters and Neelin, 2006, 2009; Peters et al., 2009; Neelin et al., 2009; Peters et al., 2010).

On the contrary, to introduce dependence on a critical value (constant over the globe) of the relative humidity is a common practice with some parameterization schemes, and its impact on the tropical dynamics has been already investigated in the past. Therefore, even if it has been shown that the critical value of precipitable water is not a constant fraction of the saturation precipitable water of the atmospheric column (Peters et al., 2002; Peters and Neelin, 2006, 2009; Peters et al., 2009; Neelin et al., 2009; Peters et al., 2010), it is probably a better choice for a first investigation to make the transition rates dependent on relative humidity, until the nature of the critical value of the precipitable water is clarified and its value parameterized at least qualitatively as a function of other fields.

In their model Stechmann and Neelin (2011) defined the birth and death rates of cumulus clouds as sigmoid functions respectively increasing and decreasing for high values of the the control parameter. The functions included several parameters whose values were taken without particular justifications, since the aim was to give a qualitative demonstration of the properties of the model. In order to keep the analysis as simple and systematic as possible, we define the transition rates in a slightly different way. We remember that the effective parameters of the two-states stochastic model, the equilibrium cloud fraction σ_0 and the relaxation time scale τ, depend on the birth and death rates as

$$\begin{cases} \sigma_0 = \dfrac{b}{b+d} \\ \tau = \dfrac{1}{b+d} \end{cases} \tag{6.1}$$

The simplest way of introducing the dependence of b and d on a control parameter is to require that only σ_0 is affected, while τ remains constant. Therefore, we prescribe the following expression for the equilibrium cloud fraction σ_0 and relaxation time

scale τ as function of the relative humidity of the atmospheric column rh

$$\begin{cases} \sigma_0 = \hat{\sigma}_0 \Gamma(rh) \\ \tau = \hat{\tau} \end{cases} \tag{6.2}$$

where $\hat{\sigma}_0$ and $\hat{\tau}$ are constants and $\Gamma(rh)$ is an activation function with values between 0 and 1. We define therefore the birth and death rates inverting 6.1 and imposing 6.2, obtaining

$$\begin{cases} b = \dfrac{\sigma_0}{\tau} = \dfrac{\hat{\sigma}_0 \Gamma(rh)}{\hat{\tau}} \\ d = \dfrac{1-\sigma_0}{\tau} = \dfrac{1-\hat{\sigma}_0 \Gamma(rh)}{\hat{\tau}} \end{cases} \tag{6.3}$$

Note that, since in general we suppose $\hat{\sigma}_0$ to be rather small (order of few percents, like in chapter 5), d will show little variations with different values of the control parameter. Therefore, this approach is basically equivalent to considering only the activation of convection to be dependent on relative humidity, while the termination of convection is supposed to act on the (almost) fixed timescale $\hat{\tau}$.

We consider for our experiments three different activation functions, corresponding to the three classes of experiments $H08$, $H09$ and $LIN085$. In the experiments labelled with $H08$ and $H09$ the activation function is an Heaviside function $\Gamma(rh) = H(rh - rh_c)$ centered respectively about the critical values of relative humidity $rh_c = 0.8$ and $rh_c = 0.9$. Therefore in this experiments for $rh < rh_c$ there will be no growth of new clouds, but only decaying of the existent clouds with rate $d = 1/\hat{\tau}$, while for $rh \geq rh_c$ new clouds will grow at rate $\hat{\sigma}_0/\hat{\tau}$ and existing clouds will decay with rate $d = (1-\hat{\sigma}_0)/\hat{\tau} \approx 1/\hat{\tau}$. In the experiments labelled with Lin we consider an activation function that is zero for $rh < 0.8$, increases linearly in the interval $0.8 \leq rh < 0.9$, and is equal to one for $rh \geq 0.9$. This is interpreted as having a critical value rh_c=0.85 and a relaxed window of activation of width 0.1.

For each of this experiments we consider different values of the parameters of the stochastic model. Similarly to what is done in Chapter 5 we keep always $\hat{\sigma}_0 = 0.05$. For each class $H08$, $H09$ and Lin we perform 4 experiments , with $\hat{\tau} = (6, 24)$ hours and $N = (100, 1000)$, in order to understand the different impact of memory and noise. Again we perform the experiments for both the BM and Kuo schemes, for a total of 24 experiments. See table 6.1 for the full list of experiments.

In these experiments the value of the relative humidity of the atmospheric column will determine the equilibrium cloud fraction towards which the stochastic cloud fraction will be (locally in time) relaxed, while the time scale of the relaxation will remain constant. A different definition would have mixed the effect on σ_0 and τ, complicating the interpretation of the impact of the stochastic model.

Experiment	Scheme	Version	rh_c	$\hat{\sigma}_0$	$\hat{\tau}$	N
BM-Det	BM	deterministic	-	-	-	-
BM-H08-T6-N100	BM	stoch., step	0.8	0.05	6	100
BM-H08-T24-N100	BM	stoch., step	0.8	0.05	24	100
BM-H08-T6-N1000	BM	stoch., step	0.8	0.05	6	1000
BM-H08-T24-N1000	BM	stoch., step	0.8	0.05	24	1000
BM-H09-T6-N100	BM	stoch., step	0.9	0.05	6	100
BM-H09-T24-N100	BM	stoch., step	0.9	0.05	24	100
BM-H09-T6-N1000	BM	stoch., step	0.9	0.05	6	1000
BM-H09-T24-N1000	BM	stoch., step	0.9	0.05	24	1000
BM-Lin-T6-N100	BM	stoch., linear	0.85	0.05	6	100
BM-Lin-T24-N100	BM	stoch., linear	0.85	0.05	24	100
BM-Lin-T6-N1000	BM	stoch., linear	0.85	0.05	6	1000
BM-Lin-T24-N1000	BM	stoch., linear	0.85	0.05	24	1000
KUO-Det	Kuo	deterministic	-	-	-	-
KUO-H08-T6-N100	Kuo	stoch., step	0.8	0.05	6	100
KUO-H08-T24-N100	Kuo	stoch., step	0.8	0.05	24	100
KUO-H08-T6-N1000	Kuo	stoch., step	0.8	0.05	6	1000
KUO-H08-T24-N1000	Kuo	stoch., step	0.8	0.05	24	1000
KUO-H09-T6-N100	Kuo	stoch., step	0.9	0.05	6	100
KUO-H09-T24-N100	Kuo	stoch., step	0.9	0.05	24	100
KUO-H09-T6-N1000	Kuo	stoch., step	0.9	0.05	6	1000
KUO-H09-T24-N1000	Kuo	stoch., step	0.9	0.05	24	1000
KUO-Lin-T6-N100	Kuo	stoch., linear	0.85	0.05	6	100
KUO-Lin-T24-N100	Kuo	stoch., linear	0.85	0.05	24	100
KUO-Lin-T6-N1000	Kuo	stoch., linear	0.85	0.05	6	1000
KUO-Lin-T24-N1000	Kuo	stoch., linear	0.85	0.05	24	1000

Table 6.1: List of experiments with non-fixed transition rates.

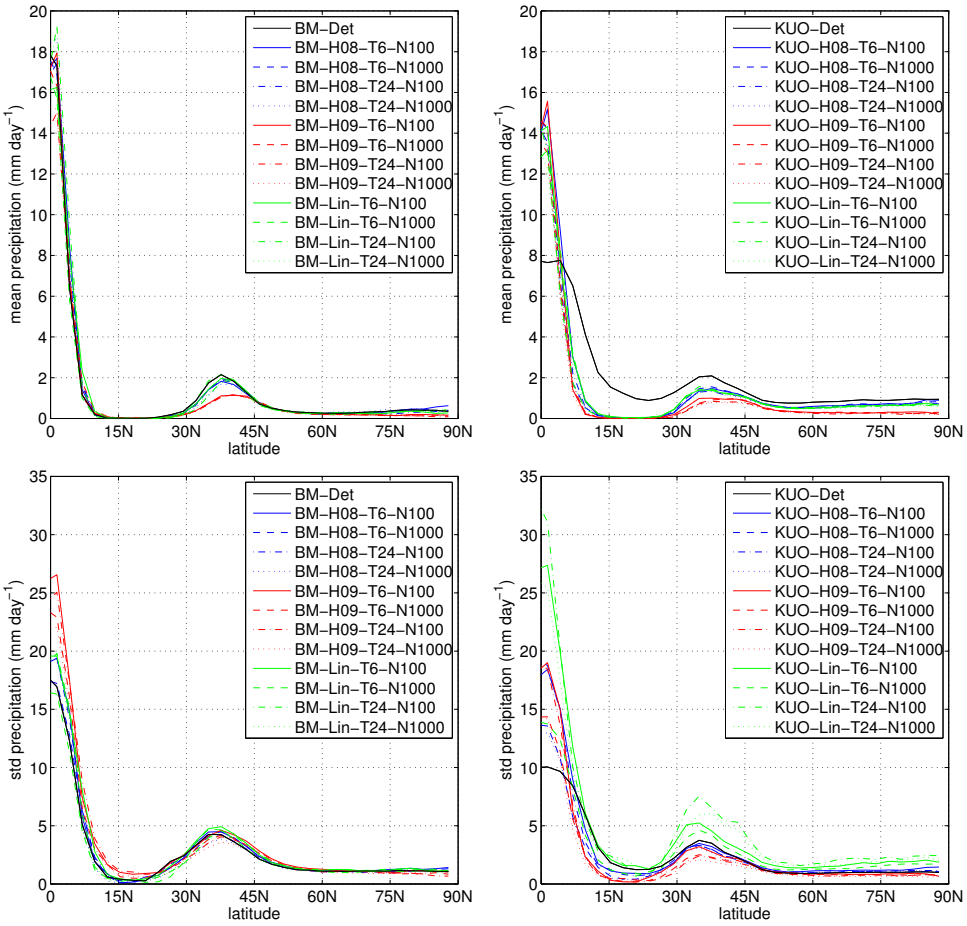

Figure 6.1: Zonal mean (top) and standard deviation (bottom) over 25 years of precipitation for all the experiments with the BM (left) and Kuo (right) scheme.

6.2 Results

Figures 6.1, 6.2, 6.3, 6.4 show the zonal mean and standard deviation of characteristic quantities of the system over 25 years of integration for all the considered experiments. We focus on the quantities most significative for the present setup: convective precipitation, zonal wind at 850 hPa, relative humidity, and the cloud fraction generated by the stochastic model. In this case, as expected, the introduction of the stochastic model has an impact on the mean state of the system, although in very different ways for the two schemes.

For the BM scheme, we can see that experiments *H*08 and *Lin* impacts only slightly the value of the mean precipitation in the deep Tropics (basically only the first grid point at 1.4 °N). On the contrary, experiments H09 show also a strong

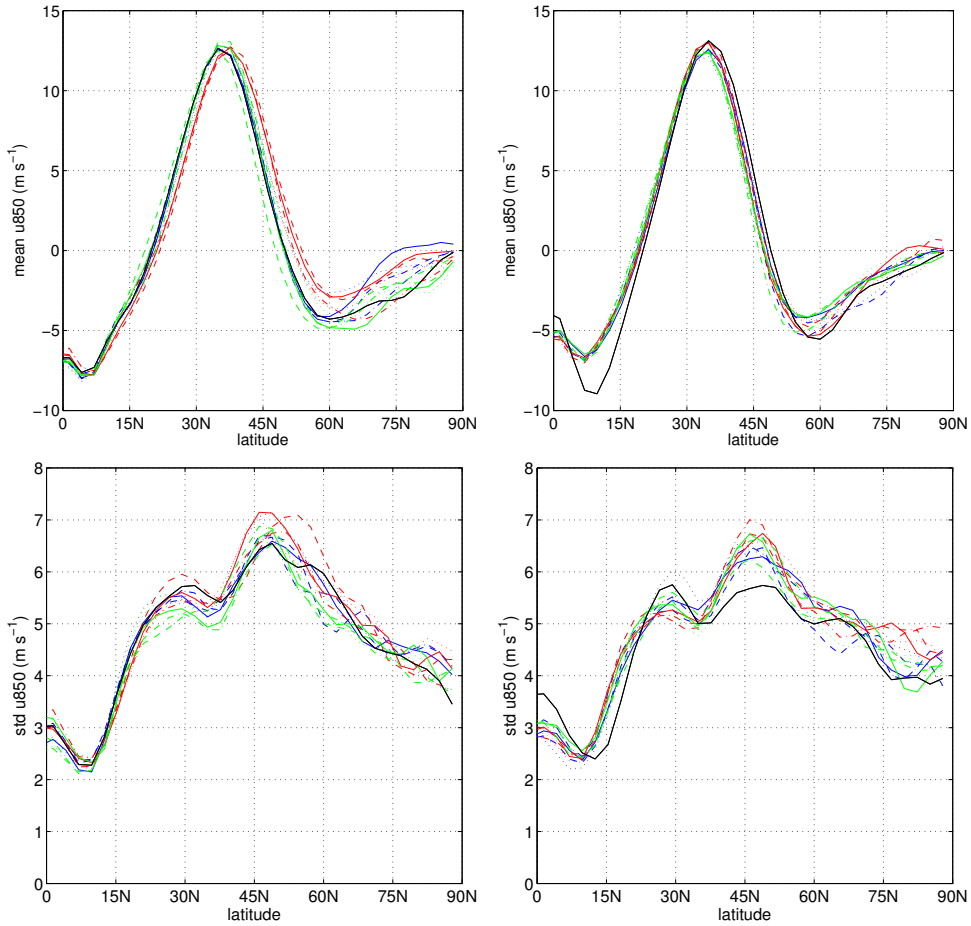

Figure 6.2: Zonal mean (top) and standard deviation (bottom) over 25 years of 850 hPa zonal wind for all the experiments with the BM (left) and Kuo (right) scheme. The legend has been omitted for better readability, refer to figure 6.1

decrease of convective precipitation at the midlatitudes, a slight poleward shift of the profile of the 850 hPa zonal wind, and a moistening of the subtropical regions. The cloud fraction generated by the stochastic model features very small values with respect to the limit value $\hat{\sigma}_0$=0.05, with distinct maxima in the tropical regions and at the midlatitudes. The cloud fraction is everywhere smaller for the more constrained experiments. For the long term mean the sensitivity to the parameters of the stochastic model, the relaxation time scale and the intensity of the noise, is not significant. Only the standard deviation of convective precipitation in the deep Tropics shows a distinct sensitivity to the noise intensity, with (reasonably) a larger increase for the experiments with higher intensity of the noise, while again the relaxation time scale seems to play no role.

For the Kuo scheme the results are rather different. The impact on the zonal

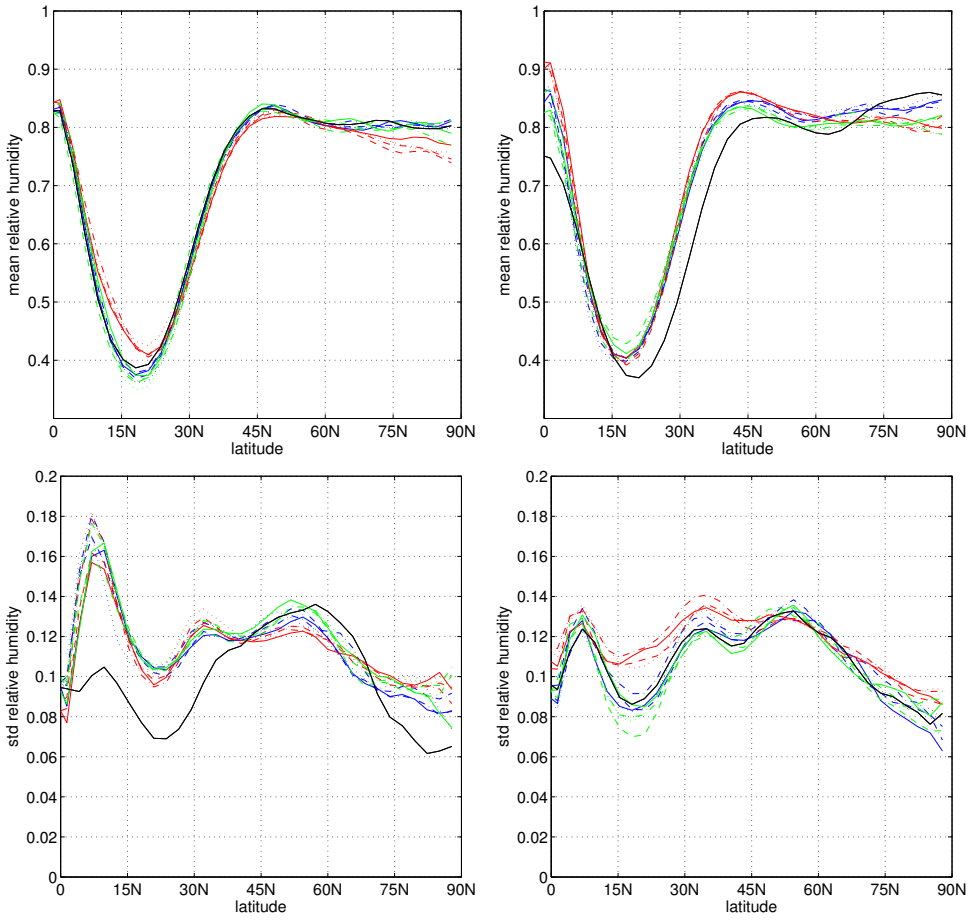

Figure 6.3: Zonal mean (top) and standard deviation (bottom) over 25 years of relative humidity for all the experiments with the BM (left) and Kuo (right) scheme. The legend has been omitted for better readability, refer to figure 6.1.

mean of precipitation is extremely pronounced for all the experiments, with almost a doubling of precipitation near the equator and a strong decrease outside the deep Tropics, with average precipitation vanishing in the subtropical regions. In general the latitudinal structure of the long term mean of all the quantities becomes qualitatively like the one obtained in the BM experiments. Again the sensitivity to the parameters of the stochastic model is limited to the standard deviation of convective precipitation, where a large difference exists in the response obtained in particular with the *Lin* experiments, showing a very large increase of variability. Finally, it is worth to note that for the Kuo scheme the average cloud fraction in the case of the *Lin* experiment behaves differently with respect to the other experiments in the Tropics and at the midlatitudes, with systematically smaller value in the Tropics and relatively higher values at the midlatitudes.

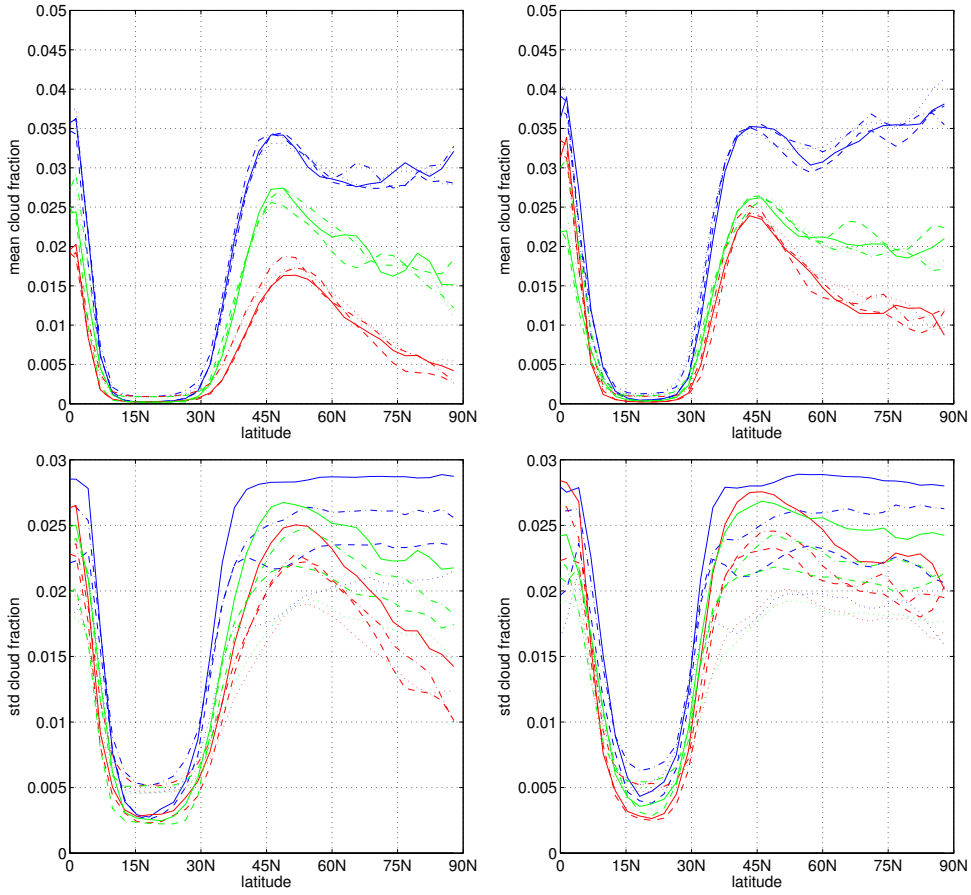

Figure 6.4: Zonal mean (top) and standard deviation (bottom) over 25 years of the cloud fraction for all the experiments with the BM (left) and Kuo (right) scheme. The legend has been omitted for better readability, refer to figure 6.1

Overall, two basic conclusions can be drawn from these experiments. Firstly, introducing a lower threshold of relative humidity below which convection is not allowed to occur pushes the Kuo parameterization to produce a climate that qualitatively is more similar to the one obtained with the BM parameterization, that already includes naturally this property, as we have seen in chapter 4. In the deterministic BM scheme the natural critical value is 0.8, therefore the $H08$ experiments show basically no differences on the first two moments of precipitation (consistently with what obtained in chapter 5), the *Lin* experiments small changes in the variability, and the $H09$ experiments minor (but significant) changes in both mean and standard deviation in the more sensitive regions (Tropics and midlatitudes). Secondly, the parameters of the stochastic model have a very little impact in determining the properties of the first two moments of the statistics of the system, with only the noise intensity showing some influence on the variability of precipitation.

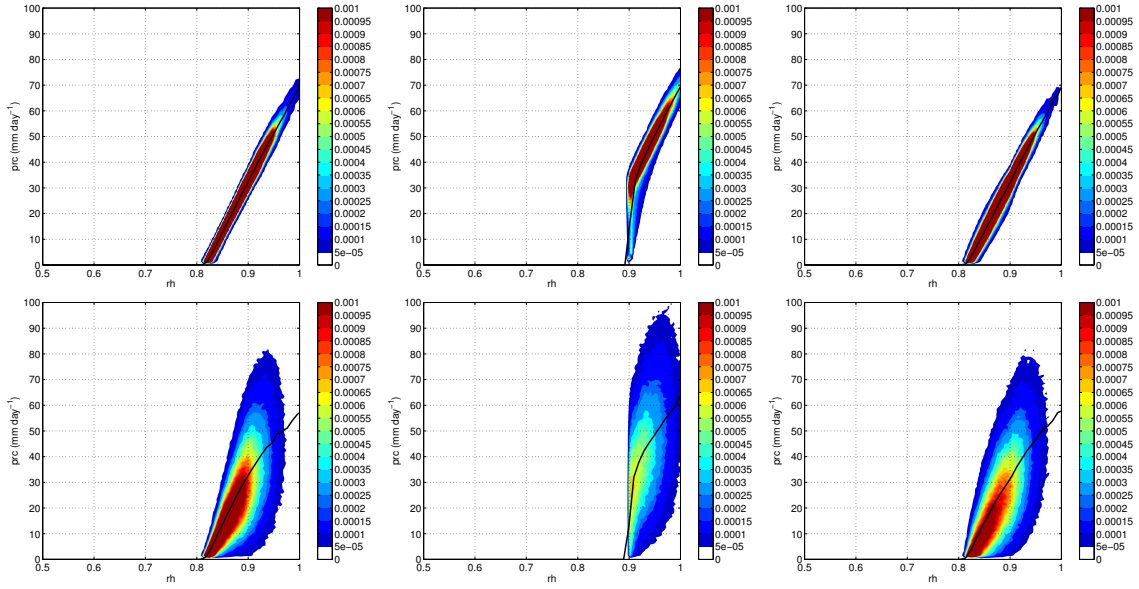

Figure 6.5: Bidimensional pdf of precipitating events in the region 10°S-10°N, as function of convective precipitation and relative humidity. In black we show the average value of precipitation conditional on the value of relative humidity. Computed from 1 year of 15 minutes data for the BM scheme in the $H08$ (first column), $H09$ (second column) and Lin (third column) experiments for $\hat{\tau}$=6 hours, N=1000 (first row) and N=100 (second row).

In order to understand to some extent the reason of this behavior, we show in figures 6.5 and 6.6 the bidimensional pdf of precipitating events in the region 10°S-10°N, as function of convective precipitation and relative humidity, for the BM and Kuo schemes respectively. The relaxation time scale also in this case has almost no role in shaping the distributions, therefore we limit ourselves to show for each class of experiments only the case $\hat{\tau}$=6 hours, for the two possible values of the noise.

From figure 6.5 we clearly see that for a small value of the noise the distribution remains exactly as in the deterministic case for both the $H08$ and Lin experiments. This is expected from what we have discussed before, since the BM parameterization already includes a natural threshold value of relative humidity of 0.8, so that the stochastic model does not change anything in determining the occurrence and amount of convective precipitation (since the noise is small, and the relaxation time scale seems not to have any role in this). For experiment $H09$ the distribution is basically the same as for the natural case but cut at the imposed threshold of 0.9. For the higher value of the noise the pdf broadens considerably, but the average value of precipitation conditional on the value of the relative humidity keeps basically the same linear relationship as in the original deterministic case. Therefore, a part from

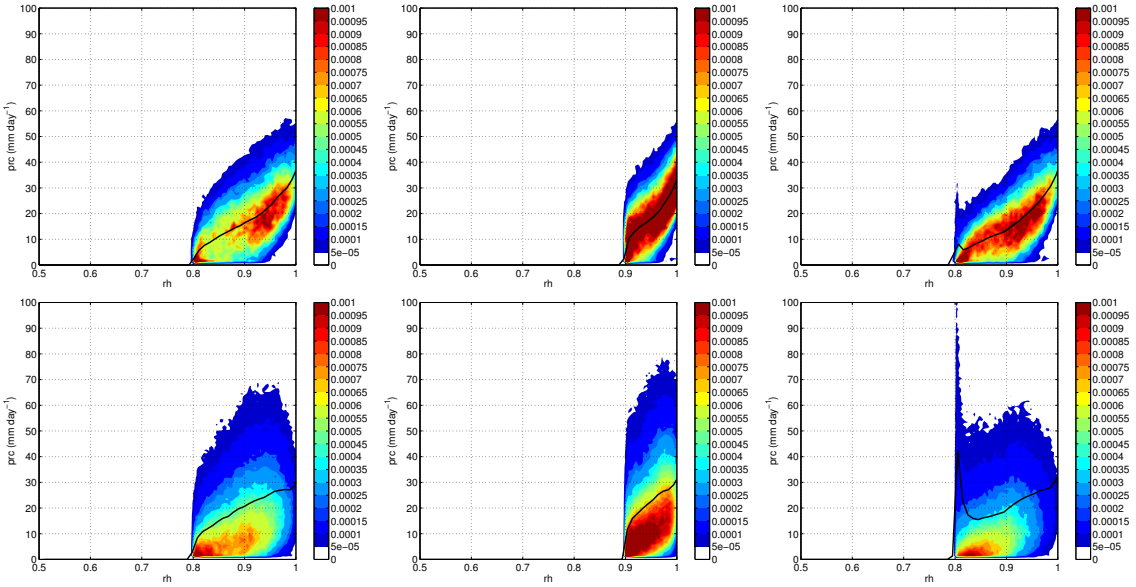

Figure 6.6: Bidimensional pdf of precipitating events in the region 10°S-10°N, as function of convective precipitation and relative humidity. In black we show the average value of precipitation conditional on the value of relative humidity. Computed from 3 months of 15 minutes data for the Kuo scheme in the $H08$ (first column), $H09$ (second column) and Lin (third column) experiments for $\hat{\tau}$=6 hours, N=1000 (first row) and N=100 (second row).

a limited increase of the variability, the introduction of the stochastic model has basically no impact on the bulk statistics of the system.

From figure 6.6 instead we see that in the case of Kuo the distribution assumes a rather different shape from the deterministic case. The stochastic model forces the parameterization to realize a linear relationship between precipitation and relative humidity, starting from the critical value. This enters into conflict with the kind of relationship naturally generated by the Kuo scheme, distorting the distribution into a compromise between the two behaviors. This modification, although in principle acceptable and at the basis of the strong impact that the introduction of the stochastic model has on the mean state in particular of convective precipitation, can lead to inconsistent behaviors that are difficult to justify. In the $H08$ experiment for the lowest value of the noise the distribution becomes bimodal, with many weak precipitating events clustered just after the critical value of the relative humidity, and a second maximum of the distribution for very high values of relative humidity and precipitation. The most worrying behavior occurs for the Lin experiment. In this case, in particular for the highest value of the noise, we have very high values of precipitation just after the critical value. As a consequence, the average value of

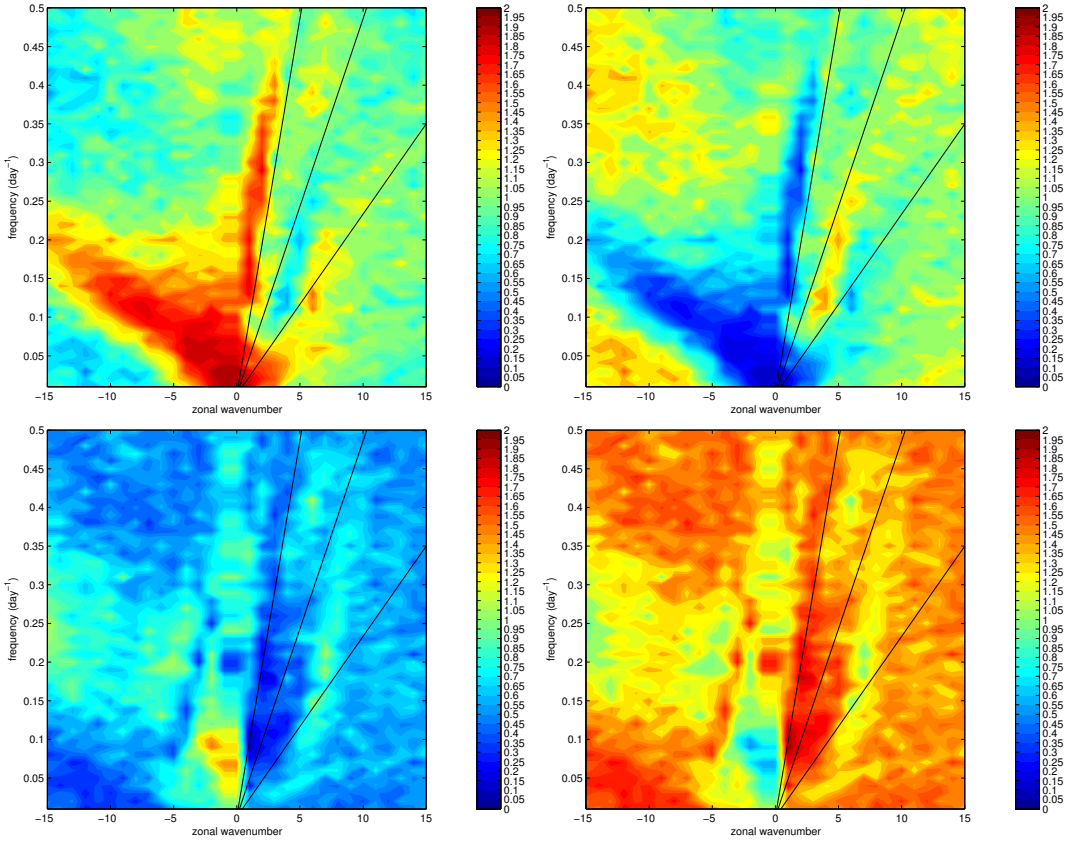

Figure 6.7: Antisymmetric (left) and symmetric (right) parts of the WK spectra of precipitation (top) and 850 hPa zonal wind (bottom) averaged in the region 10°S-10°N for the BM-H09-T6-N100 experiment. The spectra are computed as described in chapter 4.

precipitation conditional on the relative humidity is no more monotone, but has a sudden peak at the critical value, then decreases, and then increases again. This quite unnatural behavior is caused by the fact that in the *Lin* experiments for values of the relative humidity just above the critical value the equilibrium cloud fraction σ_0 assumes very small values. Therefore, even a small positive fluctuation of σ causes the ratio σ/σ_0, and therefore precipitation, to reach extremely high values. This form of numerical instability is of course extremely problematic, and will be present in principle every time we will define a stochastic model for which the equilibrium cloud fraction can assume very small values. The problem does not occur in the $H08$ and $H09$ experiments, since the equilibrium cloud fraction in these cases jumps discontinuously from 0 to $\hat{\sigma}_0$ at the critical value. This problem can be serious for applications meant to improve the representation of deep precipitating convection in more realistic cases.

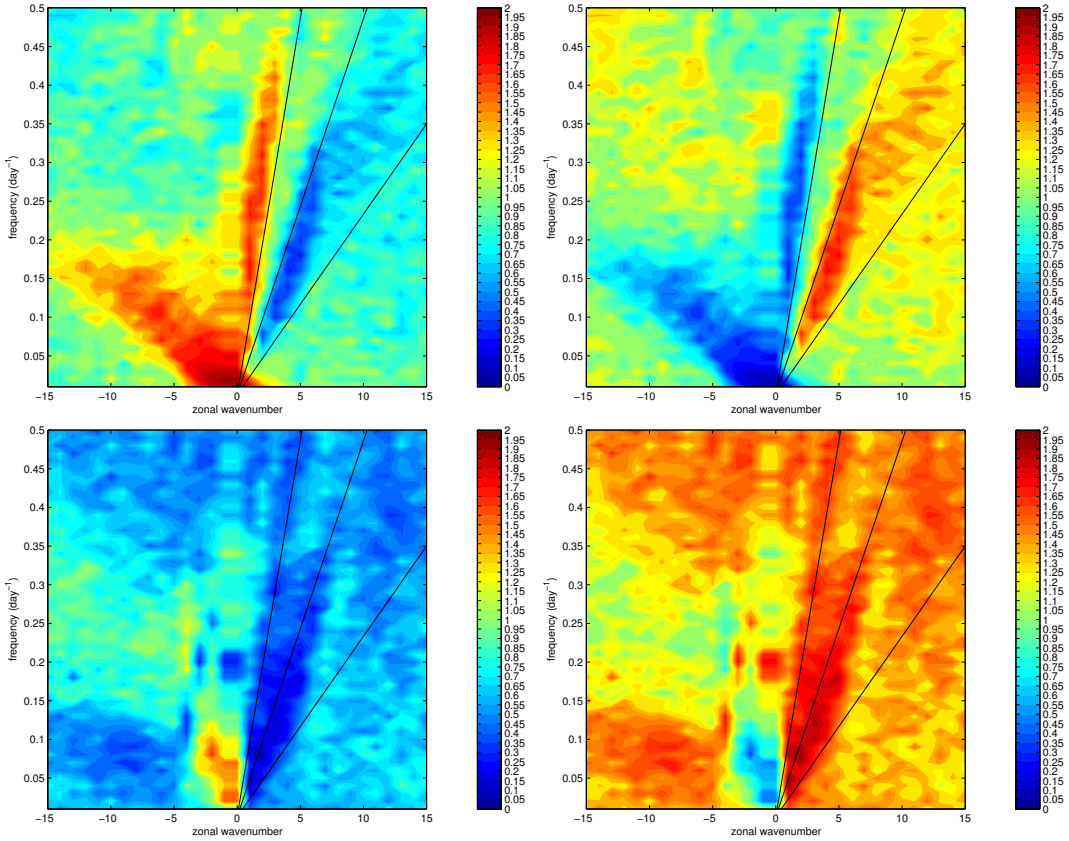

Figure 6.8: Antisymmetric (left) and symmetric (right) parts of the WK spectra of precipitation (top) and 850 hPa zonal wind (bottom) averaged in the region 10°S-10°N for the KUO-H09-T6-N100 experiment. The spectra are computed as described in chapter 4.

In order to see the impact that this has on the tropical wave dynamics we show in figures 6.7 and 6.8 the WK spectra of precipitation and 850 hPa zonal wind for the BM-H09-T6-N100 and KUO-H09-T6-N100 experiments respectively. All the other experiments behave basically in the same way. We can see that, again, the introduction of the stochastic model forces the Kuo scheme to behave similarly to the BM scheme. In particular the strong signal in the antisymmetric part of the spectrum for the u850 hPa zonal wind observed in chapter 4 is completely cancelled by the introduction of the stochastic model. It seems therefore that the relationship between precipitation and relative humidity, and in particular the presence of a threshold value of relative humidity constraining the activation of convection, is crucial in shaping the behavior of the tropical dynamics of the system.

6.3 Discussion and conclusions

In this chapter we have performed experiments coupling PlaSim to the minimal version of the stochastic model defining transition rates dependent on the value of the relative humidity of the atmospheric column, in three different ways. We have shown that in this case the mean state is affected by the introduction of the climate model, although differently in the two schemes. The difference in the response to the introduction of the stochastic model is caused by the different original relationship between precipitation and relative humidity realized by the two schemes in the deterministic version. In the case of the BM scheme, that realizes by itself a strong connection between precipitation and relative humidity (in particular already featuring a critical value of relative humidity below which precipitating convection does not occur), the introduction of the stochastic model has a small impact. In the case of the Kuo scheme on the contrary the impact is substantial, with the Kuo scheme being forced to replicate the behavior of the BM scheme. This effect, besides being promising in the fact that indeed it is possible to introduce specific properties of the activation of convection coupling the stochastic model to a parameterization scheme with different properties, can cause inconsistencies in case the properties in this way introduced enter into conflict with the original properties of the host scheme. Therefore, the design of the properties of the stochastic model should be done in an extremely careful way considering the nature of the host parameterization scheme in use.

Chapter 7

Conclusions

7.1 Summary and discussion

In this thesis we have developed a rather general framework to include sub-grid stochastic lattice-gas models for the population dynamics of an ensemble of convective elements (clouds) into a generic host deterministic parameterization scheme. The proposed formalism is on the line of models previously presented in the literature (Majda and Khouider, 2002; Khouider et al., 2010; Stechmann and Neelin, 2011; Frenkel et al., 2012), and partially on the line of (Plant, 2012), and could be used in order to bridge those approaches to the world of operational GCMs. Alongside the theoretical derivation, we have performed numerical experiments coupling a minimal version of the stochastic model to the Planet Simulator with two different parameterization schemes, while previous studies were limited to the coupling to extremely idealized models of tropical (Majda and Khouider, 2002; Khouider et al., 2010; Frenkel et al., 2012) or local (Stechmann and Neelin, 2011; Plant, 2012) convective activity.

In the first part of the thesis, after giving in **Chapter 2** a brief overview on the problem of the parameterization of unresolved atmospheric convection, we have presented and developed in **Chapter 3** our proposal. In order to make applications to real GCMs numerically treatable, we have derived a reduced model for the time evolution of the macrostate (cloud fraction) of the lattice model in mean-field approximation. This reduction method is not present in the literature in this form to our best knowledge, although similar results have been obtained for specific models (Tome and de Oliveira, 2009) in the context of the formalism of the van Kampen system size expansion (van Kampen, 2007). The van Kampen system size expansion has already been proposed in the context of stochastic parameterization of convection with mass-flux schemes (Plant, 2012), but the approach presented here

is formulated in a way that makes the application to systems featuring weak spatial interactions more intuitive, can be applied to a broader class of parameterization schemes, and results in a stochastic model with partially different properties.

We have studied in some detail the properties of the minimal version of the stochastic model, a binary system with fixed transition rates. In this formulation the stochastic model reduces to a single SDE for the cumulus cloud fraction 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 which depend on the transition (birth and death) rates, that are the equilibrium cloud fraction and the relaxation time scale. The analysis of the Fokker-Planck equation associated with the SDE shows that the stationary distribution of the model depends only on the intensity of the noise and on the equilibrium cloud fraction, and is independent of the relaxation time scale. On the contrary 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 behavior of the model when applications with a GCM are considered. The complexity of the model can be easily increased by adding 1) multiple convective states, 2) dependence of the transition rates on some large-scale variable, 3) local interactions between convective elements suited to be represented in mean-field approximation.

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 in a range of values compatible with applications 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, also 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 transition rates dependent on external fields. In case of systems featuring local interactions and, therefore, state dependent transition rates, the validity and accuracy of the reduction method depends on the nature and strength of the interactions, and has to be tested case by case. Overall the reduction method seems to be promising for the kind of applications we have in mind.

We have defined a coupling strategy in order to include the stochastic model into a pre-existing, host deterministic parameterization scheme. The state of the stochastic model modifies a relevant parameter of the parameterization in such a way that when perfect space and time scale separation is achieved (infinite number of convective elements and fast transition rates with respect to the rate of change of the external fields) we retrieve the usual value used in the deterministic version of the parameterization. In this way we define a robust coupling which introduces first order corrections due to the finite size and the time evolution of the ensemble of convective events, around the zeroth order description given by the original deterministic version of the parameterization. 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. Eventually, we have given practical examples of how this coupling strategy is implemented when coupling the minimal version of the stochastic model analyzed above to both the BM and Kuo schemes.

We have then performed numerical experiments with an aqua-planet version of the Planet Simulator, an intermediate complexity AGCM with a full set of physical parameterizations. From the methodological point of view the inclusion and testing of the stochastic model with the Planet Simulator is not different from what would be needed to include and test the stochastic model in a highly complex, state of the art GCM, with the advantage of having a higher level of portability and better computational performances. The experiments have been performed with a fixed zonally symmetric distribution of the SST without seasonal and daily cycle, 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, following a standard set up for the testing of convective parameterization schemes and for fundamental studies of tropical meteorology. First of all in **Chapter 4** we have characterized the climate resulting from the standard deterministic parameterization in order to have a reference state to compare the results obtained with the stochastic model with. Then we have performed experiments with the stochastic model set up in its minimal version, a binary system without local interactions, coupled to the BM and Kuo convective schemes as described above.

In **Chapter 5** we have performed experiments considering fixed transition rates for the stochastic model. As we have observed, this formulation of the model introduces into the GCM only the effects coming from considering a demographic description of the cloud system and realizes only a one-way coupling between the small (lattice model) and large scale (GCM) dynamics. We have performed an ex-

tensive 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 the 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, that is the quantity directly modified by the stochastic term. In these settings the stochastic extension of the parameterization conserves the climatology of its deterministic limit, thus confirming that the coupling has been defined in a robust way. The analysis of the distributions of the daily convective precipitation in the tropical areas and at the midlatitudes shows that the inclusion of the stochastic term impacts only the upper tails of the distributions, without affecting the bulk statistics. We have performed a detailed analysis of the changes in the statistics of extremes using EVT. The location and scale parameters of the GEV distribution of both tropical and midlatitude daily convective precipitation result to be highly sensitive to both the noise intensity and the autocorrelation time of the stochastic forcing. They increase seemingly logarithmically with larger noise intensity and larger autocorrelation time. This means increased and more variable typical values for the daily extremes of convective precipitation. In the limit of vanishing noise intensity and autocorrelation time the values of the location and scale parameters converge to the values of the deterministic case, again confirming the robustness of the coupling. The shape parameter is insensitive to changes in any parameter of the stochastic model. While the increase of the typical value and range of the extremes of daily convective precipitation with increasing amplitude of the fluctuations of the stochastic process is somehow expected, why these should increase with larger autocorrelation times of the noise is less clear. Sensitivity of tropical variability to the autocorrelation time of a stochastic forcing was already showed by Lin and Neelin (2000, 2002, 2003), although in a very different kind of analysis.

These results constitute also an instructive example of the fact that a parameterization calibrated on the climatology of a process is not necessarily a good parameterization for studying the extremes statistics of that process. We have given a practical example of a parameterization 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 parameterization is stochastic and has been derived in order to represent specific

features of atmospheric convection, but the principle holds in general.

In **Chapter 6** we have performed experiments considering transition rates dependent on a large-scale parameter. Consistently with observations of the onset of precipitating convection, we have considered the birth rate of cumulus cloud to be dependent on the value of the relative humidity of the atmospheric column, while keeping the death rate fixed. We have considered three cases: 1) birth rate equal to zero for relative humidity lower than 0.8, and equal to a constant above, 2) birth rate equal to zero for relative humidity lower than 0.9, and equal to a constant above, 3) birth rate equal to zero for relative humidity lower than 0.8, linearly increasing between 0.8 and 0.9, and equal to a constant above. The value of the constant is the same for the three cases and is chosen so that the equilibrium cloud fraction is the same as in the experiments performed in the previous chapters.

The results of the simulations show that in this case the mean state is affected by the introduction of the climate model, although by different extents in the two schemes. The origin of the difference in the response to the introduction of the stochastic model relies in the original relationship between precipitation (as integrated measure of the convective activity) and relative humidity realized by each scheme in the deterministic version. In the case of the BM scheme, that already includes a strong link between precipitation and relative humidity, and already features a critical value of relative humidity below which precipitating convection does not occur, the introduction of the stochastic model has a very limited impact, since it adds basically nothing new to the parameterization. In the case of the Kuo scheme on the contrary the impact is substantial. The interesting effect is that the Kuo scheme is basically pushed towards the behavior of the BM scheme. Therefore, this result hints that the relationship between convective precipitation and relative humidity (or the moisture field more in general) is a dominant factor in determining the properties of the tropical mean state and of the tropical dynamics.

7.2 Future perspectives

Starting from this work, several future lines of research can be proposed. First of all, different quantities could be used in order to determine the transition rates. Previous works have introduced CIN, CAPE and measures of the dryness of the atmosphere (Majda and Khouider (2002), Khouider et al. (2003), Khouider et al. (2010), Frenkel et al. (2012)). The length of the list of the possibilities is almost arbitrary, but the experiments should be designed in an extremely careful way, since as we have seen in the last chapter a convective parameterization realizes its own

dependency of precipitation on a certain field, and conflicts between this natural relationship and the one induced via the definition of the transition rates of the stochastic model could create inconsistencies which could make questionable the effective usefulness of this method.

Another possibility could be to introduce multiple convective regimes, and the correspondent life cycle of convective events. This will be needed in particular in applications to more realistic convective parameterizations. Operational convective parameterizations 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, on the line 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 said, clustering of convective events are indeed observed in studies of tropical dynamics. The nature of the interactions between clouds is nevertheless still unclear. 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. We also point out that adopting a mean-field description of the system implies that critical processes at a grid box scale (if they exist at all) cannot be represented. Moreover, the introduction of interactions will lead to a non linear mean-field transition matrix, which could make problematic the coupling strategy we have adopted in this paper. It is anyway encouraging the fact that most of the processes by which the presence of a cloud influences the properties of other clouds are not direct cloud-cloud interactions, but rather cloud-environment-cloud interactions: the presence of a cloud modifies the environment in a certain region, thus influencing the probability of having new clouds in that region. This is, in a sense, the spirit of the mean field approximation, which seems then not only the simplest way to take into account the effects interactions between members of a cloud population, but also to some extent a physically motivated one.

Bibliography

Adams, D. K. and N. O. Renno, 2003: Remarks on Quasi-Equilibrium Theory. *J. Atmos. Sci.*, **60**, 178–181.

Arakawa, A., 2004: The cumulus parameterization problem: Past, present and future. *J. Climate*, **17 (13)**, 2493–2525.

Arakawa, A. and W. H. Schubert, 1974: Interaction of a cumulus cloud ensemble with the large-scale environment, part I. *Journal of the Atmospheric Sciences*, **31**, 674–701.

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

Berner, J., G. Shutts, and T. Palmer, 2005: Parameterising the multiscale structure of organised convection using a cellular automaton. *ECMWF Workshop on Representation of Sub-grid Processes Using Stochastic-dynamic Models, Reading, UK*.

Betts, A. K. and M. J. Miller, 1986: A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX, ATEX and arctic air-mass data sets. *Quart. J. R. Met. Soc.*, **112**, 693–709.

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

Bowler, N. E., A. Arribas, K. R. Mylne, K. B. Robertson, and S. E. Beare, 2008: The MOGREPS short-range ensemble prediction system. *Q. J. R. Meteorol. Soc.*, **134**, 703–722.

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

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

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

Cohen, B. G. and G. C. Craig, 2006: Fluctuations in an Equilibrium Convective Ensemble. Part II: Numerical Experiments. *Journal of the Atmospheric Sciences*, **63**, 2005–2015.

Coles, S., J. Heffernan, and J. Tawn, 1999: Dependence measures for extreme value analyses. *Extremes*, **2 (4)**, 339–365.

Craig, G. C. and B. G. Cohen, 2006: Fluctuations in an Equilibrium Convective Ensemble. Part I: Theoretical Formulation. *Journal of the Atmospheric Sciences*, **63**, 1996–2004.

Dahms, E., H. Borth, F. Lunkeit, and K. Fraedrich, 2011: ITCZ Splitting and the Influence of Large-Scale Eddy Fields on the Tropical Mean State. *Journal of the Meteorological Society of Japan*, **89 (5)**, 399–411.

Eliasen, E., B. Machenhauer, and E. Rasmussen, 1970: On a numerical method for integration of the hydrodynamical equations with a spectral representation of the horizontal fields. *Report No. 2, Institute for Theoretical Meteorology, Copenhagen University*.

Faranda, D., V. Lucarini, G. Turchetti, and S. Vaienti, 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. *Journal of the Atmospheric Sciences*, **30**, 408–413.

Fraedrich, K., 2012: A suite of user-friendly global climate models: Hysteresis experiments. *Eur. Phys. J. Plus*, **127 (53)**.

Fraedrich, K., H. Jansen, U. Luksch, and F. Lunkeit, 2005: The Planet Simulator: Towards a user friendly model. *Meteorologische Zeitschrift*, **14**, 299–304.

Fraedrich, K., E. Kirk, and F. Lunkeit, 1998: PUMA: Portable University Model of the Atmosphere. *Technical Report 16, Deutsches Klimarechenzentrum*.

Frank, W. M., 1983: The cumulus parameterization problem. *Monthly Weather Review*, **111**, 1859–1871.

Frenkel, Y., A. J. Majda, and B. Khouider, 2012: Using the stochastic multicloud model to improve tropical convective parameterization: A paradigm example. *Journal of the Atmospheric Sciences*, **69**, 1080–1105.

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

Ghil, M., et al., 2011: Extreme events: dynamics, statistics and prediction. *Nonlin. Processes Geophys*, **18**, 1–56.

Gnedenko, B. V., 1943: Sur la distribution limite du terme maximum d'une série aléatorie. *The Annals of Mathematics*, **44 (3)**, 423–453.

Gregory, D., 1997: The mass flux approach to the parameterization of deep convection. *The Physics and Parameterization of Moist Atmospheric Convection, edited by: Smith, R. K., Kluwer Academic Publishers, Netherlands*, 297–319.

Gregory, D. and P. R. Rowntree, 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, R. A., 2004: Mesoscale convective systems. *Rev. Geophys.*, **42**, RG4003.

Kain, J. S., 2004: The Kain-Fritsch convective parameterization: an update. *J. Appl. Meteorol.*, **43**, 170–181.

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

Khouider, B., J. Biello, and A. J. Majda, 2010: A stochastic multicloud model for tropical convection. *Commun. Math. Sci.*, **8 (1)**, 187–216.

Khouider, B., A. J. Majda, and M. A. Katsoulakis, 2003: Coarse-grained stochastic models for tropical convection and climate. *PNAS*, **100 (11)**, 941–946.

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

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

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

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

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

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

Lin, J. W. and J. D. Neelin, 2003: Toward stochastic moist convective parameterization in general circulation models. *Geophys. Res. Lett.*, **30 (4)**, 1162.

Louis, J., 1979: A parametric model of vertical eddy fluxes in the atmosphere. *Bound. Layer Meteor.*, **17**, 187–202.

Louis, J., M. Tiedtke, and J. F. Geleyn, 1981: A short history of the operational PBL-parameterization at ECMWF. *Proceedings, ECMWF workshop on planetary boundary layer parameterization, Reading, 25-27.*

Majda, A. J. and B. Khouider, 2002: Stochastic and mesoscopic models for tropical convection. *PNAS*, **99 (3)**, 1123–1128.

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

Mapes, B., 1997: Equilibrium vs. activation controls on large-scale variations of tropical deep convection. *The Physics and Parameterization of Moist Atmospheric Convection, edited by: Smith, R. K., Kluwer Academic Publishers, Netherlands*, 321–358.

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

Mapes, B. E., 1993: Gregarious tropical convection. *J. Atmos. Sci.*, **50**, 2026–2037.

McKane, A. J. and T. J. Newman, 2004: Stochastic models in population biology and their deterministic analogs. *Phys. Rev. E*, **70 (041 902)**.

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

Neale, R. and B. Hoskins, 2001a: A standard test for AGCMs including their physical parametrizations. I: The proposal. *Atmospheric Science Letters*, **1**, 101–107.

Neelin, J. D., O. Peters, and K. Hales, 2009: The transition to strong convection. *J. Atmos. Sci.*, **66 (8)**, 2367–2384.

Neelin, J. D., O. Peters, J. W.-B. Lin, K. Hales, and C. E. Holloway, 2008: Rethinking convective quasi-equilibrium: observational constraints for stochastic convective schemes in climate models. *Phil. Trans. R. Soc. A*, **366**, 2581–2604.

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

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

Peters, O., A. Deluca, A. Corral, J. D. Neelin, and C. E. Holloway, 2010: Universality of rain event size distributions. *J. Stat. Mech.*, P11 030.

Peters, O., C. Hertlein, and K. Christensen, 2002: A complexity view of rainfall. *Physical review letters*, **88 (1)**, 18 701.

Peters, O. and J. D. Neelin, 2006: Critical phenomena in atmospheric precipitation. *Nature Physics*, **2 (6)**, 393– 396.

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

Peters, O., J. D. Neelin, and S. W. Nesbitt, 2009: Mesoscale convective systems and critical clusters. *J. Atmos. Sci.*, **66 (9)**, 2913–2924.

Plant, R. and G. C. Craig, 2008: A stochastic parameterization for deep convection based on equilibrium statistics. *Journal of the Atmospheric Sciences*, **65**, 87–105.

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

Plant, R. S., 2012: A new modelling framework for statistical cumulus dynamics. *Philosophical Transactions of the Royal Society A*, **370 (1962)**, 1041–1060.

Ragone, F., K. Fraedrich, H. Borth, and F. Lunkeit, 2013: Coupling a minimal stochastic lattice-gas model of a cloud system to an agcm. *Quart. J. R. Meteorol. Soc.*, submitted.

Randall, D. A. and D. M. Pan, 1993: Implementation of the Arakawa-Schubert cumulus parameterization with a prognostic closure. *The Representation of Cumulus Convection in Numerical Models, Meteorol. Monogr., vol. 46, edited by K. A. Emanuel and D. J. Raymond, Am. Meteorol. Soc., Boston*, 137–144–385.

Randall, D. A., D. M. Pan, P. Ding, and D. G. Gripe, 1997: Quasi-equilibrium. *The Physics and Parameterization of Moist Atmospheric Convection, edited by: Smith, R. K., Kluwer Academic Publishers, Netherlands*, 359–385.

Randall, D. A., et al., 2007: Climate models and their evaluation. *Climate Change 2007: The Physical Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., Cambridge University Press, Cambridge, UK and New York, NY, USA.*

Redelsperger, J., F. Guichard, and S. Mondon, 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, Vol. 18. Springer.

Roeckner, E., et al., 1992: Simulation of the present day climate with the ECHAM model: Impact of model physics and resolution. *Report 93, Max-Planck-Institut fur Meteorologie.*

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

Seiffert, R., R. Blender, and K. Fraedrich, 2006: Subscale Forcing in a Global Atmospheric Circulation Model and Stochastic Parameterisation. *Q. J. R. Meteorol. Soc.*, **132**, 1627–1643.

Slingo, A. and J. M. Slingo, 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.

Stechmann, S. N. and J. D. Neelin, 2011: A stochastic model for the transition to strong convection. *J. Atmos. Sci.*

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

Stephens, G. L., S. Ackerman, and E. A. Smith, 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. *Atmospheric Research*, **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. and M. J. de Oliveira, 2009: Role of noise in population dynamics cycles. *Physical Review E*, **79**, 061 128.

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

van Kampen, N. G., 2007: *Stochastic processes in physics and chemistry*. Elsevier, 3rd edn.

Wheeler, M. and G. N. Kiladis, 1999: Convectively Coupled Equatorial Waves: Analysis of Clouds and Temperature in the Wavenumber–Frequency Domain. *J. Atmospheric Sci.*, **56**, 374–399.

Yanai, M., S. Esbensen, and J. Chu, 1973: Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. *J. Atmospheric Sci.*, **30**, 611–627.

Yano, J. I., 1999: Scale-separation and quasi-equilibrium principles in Arakawa and Schubert's cumulus parameterization. *J. Atmos. Sci.*, **56**, 3821–3823.

Yano, J.-I., K. Fraedrich, and R. Blender, 2001: Tropical Convective Variability as 1/ f Noise. *Journal of Climate*, **14**, 3608–3616.

Yano, J.-I., W. W. Grabowski, G. L. Roff, and B. E. Mapes, 2000: Asymptotic approaches to convective quasi–equilibrium. *Quart. J. Roy. Meteor. Soc.*, **126**, 1861–1887.

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

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

Acknowledgments

I am grateful to my supervisor Klaus Fraedrich for his help and suggestions during these four years of PhD. I also would like to thank my co-supervisor Hartmut Borth for the huge amount of time he has spent with me during the first two years, and my panel chair Hartmut Grassl for his help and advices in some critical moments of the PhD.

I would like to thank all the people who have worked or are working in the SICSS for the support that the school has given me during my time as a SICSS member. I am thankful to all the members of the Theoretical Meteorology Group for their help and for being always open to my questions and perplexities.

Many people have been important in order to keep me sane of mind through these sometimes rather difficult years. The first person I have to thank in this sense is Salvatore (I said you would have been the first!), whose patience and wisdom has been a reference point in the most stressful periods of the past two years. I want also to thank all the other people who have made my time in Grindelberg 5 pleasant and fun in different parts of this journey: Robert, Sebastian, Davide, Markus, Shabeh, Jeroen, Tamas, Nicolas, Teresa, Dan, Huang, Xiuhua, Valerio, Giovanna and my current officemates Jonas, Anja and Maida.

My life in Hamburg would have not been possible if I didn't have also very good friends outside the University, in particular Eleftheria, Suvarchal, Olga, Thomas, Nasia, Marianna, Filippo, and my little Hamburg family, Francesco, Fabio and Lorenza.

Despite the distance, there's a lot of people in Italy that has been important during my time in Germany. In particular Pollo, who I guess is the only one who really understands, Eugenio (I'm coming to see the little Carlo as soon as possible!), my soulmate Lape, il Cast, with whom I have shared more than with anybody else, and of course, as always in the past fifteen years, Matilde.

Finally, I would like to thank my parents and my family for their love, support and infinite patience.

This thesis, but most importantly these four years of learning and fighting are dedicated to Mattia Tondelli, who was the only person who suggested me not to quit with the PhD back in summer 2011, during what would have been our last conversation. I've done one of the two things you told me that night; I'm not sure about the other one. Two years have passed already since that day, but you keep visiting me, in dreams and tears, and you are always welcome.

Erklärung/Declaration

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutz habe.

I hereby declare, on oath, that I have written the present dissertation by my own and have not used other than the acknowledged resources and aids.

Hamburg, den Unterschrift