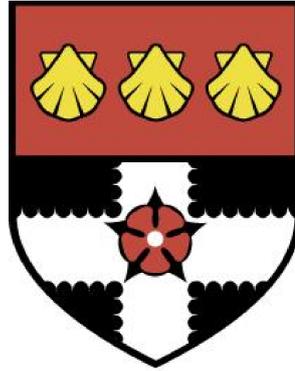


THE UNIVERSITY OF READING

Department of Meteorology



**Using Stochastic Parameterisations to
Study the Sensitivity of the Atmosphere
to its High-Frequency Variability**

MICHAEL WHITALL

A thesis submitted for the degree of Doctor of Philosophy

29 / March / 2012

DECLARATION

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Michael Whittall

ABSTRACT

It has been suggested that the atmosphere's climatic behaviour exhibits sensitivity to its high-frequency variability, associated with sub-grid processes parameterised in global atmospheric simulations. The parameterisations are a major source of model error, and uncertainties in their formulation must be sampled to give representative spreads in ensemble forecasts. Some component of variability / uncertainty in resolved-scale effects of parameterised processes derives from sub-grid details of the atmosphere's state, which are undefined in models. This has motivated various stochastic parameterisation methods aiming to sample sub-grid uncertainties in ensemble forecasts and give more realistic high-frequency variability in models.

This thesis explores the atmosphere's sensitivity to high-frequency variability, by comparing different stochastic methods from the literature in ensemble Single-Column Model simulations of convection over the Tropical West Pacific. A comparison of a smaller sample of methods is then made in an aqua-planet simulation, to investigate dynamical sensitivities.

The Single-Column Model's ensemble variability is dominated by the triggering behaviour of the convection scheme, giving different profiles of variability depending on details of its formulation. All the stochastic schemes investigated scaled-up this internal convective variability, giving an increased range of states, which caused increases in mean layer-cloud condensate and lower-troposphere moisture fluxes. The ensemble-mean response was also strongly influenced by changes in convective activity, which were highly dependent on the atmospheric state, modelling framework, and type of stochastic scheme used. Physical mechanisms have been identified for some of these changes, but others maybe model artefact. The aqua-planet's large-scale circulation was found to be insensitive to simple stochastic schemes. But the cloud responses gave significant surface radiative forcings, so the coupled atmosphere-ocean system would be likely to experience noise-induced drift. Investigation of a stochastic scheme designed to sample sub-grid variations in convection suggested sub-grid uncertainties dominate parameterisation uncertainties at grid-scales of 50km or less.

ACKNOWLEDGEMENTS

This work was funded by the Natural Environment Research Council, with additional CASE award funding from the Met Office (UK Meteorological Office). Simulations were performed on the HPCX and HECTOR supercomputing facilities, using part of the NCAS allocation on those platforms. Many thanks to Ricky Wong and Sarah Beare at the Met Office for providing the Single-Column Model and aqua-planet simulations used in this study. Additional thanks to Sarah and Alberto Arribas (also at the Met Office) for providing the code for the Random Parameters stochastic parameterisation scheme and bug-fixes for the models. Thanks to Colorado State University for making the TOGA-COARE IFA mean observed data and estimated dynamical forcings freely available online. And thanks of course to the University of Reading and its Meteorology department for taking me on as a research student, providing all the facilities I've needed to carry out this work, and being very patient with my inability to meet deadlines!

Many thanks to the department's IT support team for their tireless work keeping the system running, and for the many times they've resuscitated the machine I use to store and analyse my data whenever my codes crashed it! Endless gratitude also to the department's Met Office Unified Model support staff and other UM veterans who have helped with my simulations; Lois Steenman-Clark and Jeff Cole especially.

Many thanks to those on my monitoring committees, for all their advice and their patience with my late submission of reports; to Julia Slingo for sharing her world-renowned insights in tropical meteorology, to Maarten Ambaum for sound advice and heart-warming encouragement, and to Dan Kirshbaum for stepping in to offer helpful positivity and valuable insights on convection. Thanks also to Dan and to Michael Herzog at Cambridge University for examining my thesis following its initial submission in 2010, and for granting me a deferred resubmission despite the very incomplete and patchy nature of that work.

Many, many thanks to my supervisor Robert Plant, for so many years of vital technical expertise, practical advice, undying enthusiasm and encouragement for my work, and unceasingly relaxed and friendly attitude. His patience and positive, helpful outlook in the face of my disorganisation have been inspiring! He also provided the model code for the Plant & Craig (2008)

stochastic convection scheme and the Kain (2004) convection scheme. Thanks also to my CASE supervisors at the Met Office, Neill Bowler and Alberto Arribas; they have both offered many a helpful piece of advice or fascinating discussion, repeatedly travelled to Reading from Exeter to meet me and exchange results, and I'll fondly remember all the nice lunches together at the Met Office canteen.

Thanks to all in mesoscale group for the useful feedback and delicious termly dinner outings. Thanks to everyone who shared office 2U06 with me; the innumerable UNIX and matlab questions answered, conversations both meteorological and non-meteorological and moments of collective appreciation for clouds out the window will live with me forever! As will the Christmas pantomimes, conkers tournaments, BBQ and barn-dances, Met's Got Talent shows, pub trips, fantastically random conversation threads on Met Social mailing list, and enough things that made me laugh that I'd have to write another entire thesis to recount them! Reading University Meteorology Department is easily the most friendly and sociable academic institution I've experienced. Special thanks to the secretaries and also to all the hard-working, often unseen people who maintain the department; they vacuum the floors, clean the toilets, fix the heating and water systems and do all-manner of chores to keep the place really pleasant to work in.

Many thanks to my parents and in-laws for all the moral support, encouragement, much-needed holidays and for not letting me give up on this project. I feel a lot happier having finished it, whatever the outcome turns out to be! Endless gratitude to my amazing wife Helen for continuously supporting me all this time, for her patience with all my late home-comings, the yummy home-cooked hot meals brought to the department along with life-saving supplies brought in for those overnight last-ditch attempts at writing-up on-time, all the hilarious ESSC / Met Department email exchanges, oh and for providing the photos of sub-grid weather phenomena in chapter 1! Eternal thanks also to Zion & Aaron, Chris "hat" and everyone who's heard of "The Brooms of Destruction", Chris W, all in SPEAK Network, Berkshire Greenpeace group and "home-group" for keeping me sane all these years, sticking with me despite me being a zoned-out burned-out wreck, and forever reminding me that there really is an amazing world of music, love and life to thrive in beyond the pages of this document! Special thanks also to everyone I've got to know at Reading University Chaplaincy Centre; nowhere else have I found such a friendly place and nurturing community I can nip out to for lunch breaks and cups of tea!

CONTENTS

1	INTRODUCTION - MOTIVATION AND THEORETICAL BASIS	2
1.1	OVERVIEW	2
1.2	CHAOS AND CLIMATE MODELLING	4
1.3	THE ROLE OF UNRESOLVED PROCESSES	8
1.4	THE ROLE OF HIGH-FREQUENCY / SMALL-SCALE VARIABILITY	12
1.5	STOCHASTIC PARAMETERISATIONS	18
1.5.1	MODEL UNCERTAINTY SCHEMES	18
1.5.1.1	PRAGMATIC BASIS	18
1.5.1.2	PERTURBED PHYSICS AND MULTI-MODEL APPROACHES	20
1.5.1.3	STOCHASTIC PHYSICS SCHEMES	22
1.5.2	PHYSICAL VARIABILITY SCHEMES	25
1.5.3	BACKSCATTER SCHEMES	29
1.6	EFFECTS OF STOCHASTIC FORCING ON DYNAMICAL SYSTEMS	32
1.7	DISCUSSION AND AIMS OF THIS THESIS	35
1.8	OVERVIEW OF METHOD	37
2	INVESTIGATING THE VARIABILITY OF PARAMETERISED PROCESSES IN A SINGLE-COLUMN MODEL	40

2.1	OVERVIEW AND CHOICE OF EXPERIMENT	40
2.2	EXPERIMENTAL SETUP	42
2.2.1	MODEL DESCRIPTION	43
2.2.2	MODEL VARIANTS	44
2.2.2.1	DEFAULT UM	44
2.2.2.2	RANDOM PARAMETERS	46
2.2.2.3	CONSTANT RANDOM PARAMETERS	49
2.2.2.4	MULTIPLICATIVE NOISE	50
2.2.2.5	KAIN-FRITSCH CONVECTION SCHEME	51
2.2.2.6	PLANT AND CRAIG STOCHASTIC CONVECTION SCHEME	52
2.2.2.7	DETERMINISTIC LIMIT OF THE PLANT AND CRAIG SCHEME	54
2.2.2.8	TIME-SMOOTHED CONVECTION	55
2.2.3	ENSEMBLE FRAMEWORK	56
2.2.4	ENSEMBLE STATISTICS	57
2.2.5	FORCING DATA	58
2.3	OVERVIEW OF METEOROLOGY FOR THE CASE	63
2.3.1	RAINFALL VARIABILITY	63
2.3.2	TEMPERATURE AND MOISTURE PROFILES	68
2.3.3	OUTGOING LONGWAVE RADIATION	70
2.3.4	SUMMARY	72
2.4	RESULTS	72

2.4.1	SENSITIVITY TO INITIAL CONDITION (IC) PERTURBATIONS	72
2.4.2	ENSEMBLE SPREAD AND VARIABILITY	79
2.4.3	SENSITIVITY TO STOCHASTIC PARAMETERISATION (SP) VERSUS INITIAL CONDITION (IC) PERTURBATION	81
2.5	CONCLUSIONS	84
3	COMPARISON OF STOCHASTIC PARAMETERISATION APPROACHES IN A SINGLE- COLUMN MODEL	90
3.1	INTRODUCTION	90
3.2	COMPARISON OF ENSEMBLE VARIABILITIES	92
3.3	COMPARISON OF PERTURBED ENSEMBLE SPREADS TO MODEL-UNCERTAINTY	101
3.4	COMPARISON OF ENSEMBLE MEAN-STATES	105
3.4.1	NOISE-INDUCED DRIFTS IN COLUMN WATER VAPOUR	105
3.4.2	NOISE-INDUCED CHANGES IN TEMPERATURE AND HUMIDITY PROFILES	108
3.4.3	NOISE-INDUCED CHANGES IN THE MOISTURE BUDGET	113
3.4.4	NOISE-INDUCED CHANGES IN CLOUD	122
3.5	CONCLUSIONS	127
3.5.1	ON THE NATURE OF PARAMETERISED PHYSICS VARIABILITY	127
3.5.2	ON THE RESPONSE OF PARAMETERISED PROCESSES TO CHANGES IN THEIR VARIABILITY	129
3.5.3	ON THE RELATIVE IMPORTANCE OF PHYSICAL AND MODELLING UN- CERTAINTIES	131

4	SENSITIVITY OF THE ATMOSPHERE TO HIGH-FREQUENCY VARIABILITY IN AN AQUA-PLANET FRAMEWORK	133
4.1	OVERVIEW	133
4.2	EXPERIMENTAL SETUP	135
4.2.1	MODEL DESCRIPTION	136
4.2.2	EXPERIMENTS	138
4.2.3	STATISTICAL METHODS	141
4.3	RESULTS	146
4.3.1	ANALYSIS OF THE CONTROL RUN	147
4.3.2	COMPARISON OF MOISTURE BUDGETS AND RAINFALL VARIABILITY	155
4.3.3	COMPARISON OF RADIATION, HEAT BUDGETS AND CLOUD	160
4.3.4	LARGE-SCALE CIRCULATION	170
4.4	DISCUSSION AND CONCLUSIONS	173
5	CONCLUSIONS	178
5.1	SUMMARY	178
5.2	IS THE ATMOSPHERE'S RESPONSE TO HIGH-FREQUENCY VARIABILITY HIGHLY DEPENDENT ON IT'S INTERNAL PROPERTIES?	180
5.3	HOW IMPORTANT ARE INHERENT SUB-GRID UNCERTAINTIES RELATIVE TO GENERIC PARAMETERISATION UNCERTAINTIES?	182
5.4	WHAT ARE THE KEY MECHANISMS INVOLVED IN THE GLOBAL ATMO- SPHERE'S SENSITIVITY?	184
5.5	FUTURE WORK	187

5.5.1	EXTENSION OF THE PRESENT METHODOLOGY	188
5.5.2	OTHER SUGGESTED EXPERIMENTS	191
5.5.3	SUGGESTED MODEL DEVELOPMENTS	192
APPENDICES		194
A	DETAIL OF THE PARAMETERISATION SCHEMES USED IN THE MET OFFICE UNIFIED MODEL	195
A.1	BOUNDARY LAYER	195
A.2	CONVECTION	196
A.3	LARGE-SCALE CLOUD	197
A.4	MICROPHYSICS	198
A.5	RADIATION	198
B	STATISTICAL METHODS USED TO TEST THE SIGNIFICANCE OF DIFFERENCES BETWEEN AQUA-PLANET SIMULATIONS	200
B.1	TEST FOR DIFFERENCES IN MEAN-STATES	200
B.2	TEST FOR DIFFERENCES IN VARIANCE	201
B.3	ESTIMATION OF EFFECTIVE SAMPLE SIZE	202
B.3.1	EFFECTIVE SAMPLE SIZE FOR THE SAMPLE VARIANCE	202
B.3.2	EFFECTIVE SAMPLE SIZE FOR THE SAMPLE MEAN	204
REFERENCES		206

CHAPTER 1: INTRODUCTION - MOTIVATION AND THEORETICAL BASIS

1.1 OVERVIEW

The ability of science to forecast the weather (an initial-value problem, or so-called “prediction of the first kind”) is limited by the chaotic exponential growth of errors in the initial state. Analogously, our ability to simulate the climate and predict its future nature under changing radiative forcings (a “prediction of the second kind”) is also limited by chaotically amplified manifestations of errors, this time in the representation of the atmosphere’s governing equations in models. This is described in section 1.2.

As described in section 1.3, the main source of these errors is thought to be in the parameterised representation of processes not resolvable in current discretised global model grids, and various studies have linked biases in climate models to deficiencies in parameterisations.

Section 1.4 discusses one common deficiency; the unrealistic (often under-estimated) variability of many parameterisations at the smaller scales / faster frequencies which are resolvable. This problem is expected from theoretical considerations; the deterministic bulk parameterisations employed in GCMs do not account for the uncertainties in the unresolved sub-grid variability of the atmosphere. Rather, they assume sub-grid fluctuations to be sufficiently small-scale to be in a statistical equilibrium, such that their contributions to the resolved-scale average to a well-defined form which is entirely predictable from the resolved state. Observations do not support this assumption in the real world. Also, some studies have linked the failure of models to correctly simulate some large-scale / slow modes of atmospheric variability to their inadequate high-frequency variability. This suggests that the correction of some atmospheric model

biases may require improvements to the realism of the variability of sub-grid parameterisations, which in turn may require the parameterisations to account for sub-grid uncertainties through non-deterministic formulation.

This paradigm has motivated the development of a range of stochastic parameterisations, described in section 1.5. Another, parallel motivation for such schemes has been in efforts to represent modelling uncertainties in ensemble prediction systems, which give over-confident forecasts when accounting for initial condition uncertainty alone. The deficient variability of parameterisations and the deficient spread of ensemble forecasts are thought to share some of the same theoretical under-pinnings, in the failure of deterministic parameterisations to represent sub-grid uncertainties. Modellers have attempted to tackle these problems using three broad approaches; generically sampling uncertainties in existing sub-grid parameterisations, developing new sub-grid parameterisations which are inherently stochastic in formulation, and stochastically re-introducing energy which is unphysically diffused from model fields.

Whilst methods have now been established to alter the high frequency variability in atmospheric GCMs in a range of different ways (often yielding improvements in forecast skill), the sensitivity of the atmosphere to the nature of its high-frequency variability is not well understood. The sensitivity of a much simpler dynamical system to stochastic forcing (illustrated in section 1.6) can be understood more trivially, and indicates several categories of system sensitivity to high-frequency variability which the atmosphere can also display.

This thesis addresses some of the unknowns regarding the sensitivity of the atmosphere to its own high frequency variability: Does the atmosphere respond generically to any form of “noise” at high frequencies, or to what extent does its response change depending on details of the noise or the atmospheric state? How important is sub-grid state uncertainty relative to overall parameterisation error? And by what mechanisms does high-frequency variability influence well-resolved climatic behaviour? These questions are discussed further in section 1.7. Shedding some light on these issues should help to guide the choice of stochastic parameterisations for GCMs, and indicate how sophisticated such schemes need to be. It would also be of general interest in understanding scale interactions in the atmosphere.

A general experimental methodology to address these questions is outlined in section 1.8.

Sensitivity of the atmosphere to high-frequency variability will be investigated by comparing several different stochastic parameterisations from the literature in a common host model, along with some alternative deterministic modifications which also affect the high-frequency variability. Complimentary experiments are performed in a Single-Column Model (SCM) and a 3D Aqua-Planet framework, to disentangle aspects of the atmospheric response due to resolved dynamical processes and sub-grid / non-dynamical processes.

In chapter 2, the stochastic and deterministic model configurations to be studied are described, the SCM framework is described and validated, and its high-frequency variability is investigated. In chapter 3 the SCM is then used to compare all the model configurations, and key sensitivities of the parameterised sub-grid and non-dynamical processes to the differences in high frequency variability are investigated. Chapter 4 shows results in the 3D aqua-planet framework, with a horizontal resolution consistent with climate simulations, comparing two different stochastically perturbed configurations against a deterministic control simulation. Overall conclusions are drawn in chapter 5.

Some of the Single-Column Model work presented in chapters 2 and 3 was presented in the peer-reviewed literature in a paper co-written by myself and my supervisor Robert Plant (Ball & Plant 2008; Ball is this author's former surname pre-marriage). However, the work has been vastly expanded since then, with many additional results and conclusions from the Single-Column Model experiments presented in those chapters of the thesis.

1.2 CHAOS AND CLIMATE MODELLING

If a small perturbation is made to the state of the atmosphere at some specified initial time, its state at later times will diverge exponentially from what it would have been had the perturbation not occurred. Eventually the state initially perturbed only slightly will bare no recognisable similarity to the unperturbed version of events. This phenomenon, famously discovered by Edward Lorenz in 1961 and popularly known as the butterfly effect, is exhibited even in vastly simplified models of the atmosphere (as in Lorenz' first paper on the subject, Lorenz 1963). However, if the evolution of both the perturbed and unperturbed futures were somehow followed and accurately measured over sufficiently many years, one would expect that the mean state and the

statistics of the variability in each would be alike. In other words, we expect that the nature of the atmosphere's climate is determined purely by the basic parameters of the system (the Earth's radius, rotation-rate and the mass of the atmosphere), the external forcings at its boundaries (the properties of the Earth's surface and the distribution of solar and terrestrial radiation), and the governing equations. Perturbing the initial conditions simply selects a different realisation with the same statistical properties. On this basis we can make useful predictions of future climate under changing external forcings many tens of years ahead, despite the fact that the atmosphere's chaotic nature limits our ability to predict its state at a specific time beyond a few days ahead. One simply needs to create a model capable of correctly reproducing the climate as a function of the Earth's basic parameters and external forcings, and use it to investigate how the climate changes as the external forcings are varied according to some future scenario.

However, the construction of such a model has turned out to be a very non-trivial task. This can perhaps be illustrated by drawing a parallel to the popular "butterfly effect" allegory: Whilst a single stroke of a butterfly's wing will indeed alter the specific course of future weather, it will make no difference to the weather's general nature; the climate. However if one were somehow to systematically alter, say, the rules by which all the Earth's bees behave and interact, there could potentially be knock-on effects on the pollination rates of numerous plant species, and hence on vegetation cover and the wider ecology, and hence on transpiration and other factors that affect the climate. However, due to the level of complexity and non-linearity involved, it might be as hard to foresee the climatic impact of the change in the bees' behaviour as it is to predict the precise consequences of the butterfly's wing-beat. So, whilst weather forecasters recognise that a seemingly insignificant change in the initial state of the atmosphere can drastically alter the weather in a few days time, climate modellers also have to consider the possibility that a seemingly insignificant change in the rules governing the atmosphere could substantially alter its general behaviour. (see figure 1.1). Or in dynamical terms, just as the future trajectory of the atmosphere's state in phase space is extremely sensitive to small changes in the initial conditions, the shape of the attractor which the trajectory follows may be highly sensitive to small changes in the governing equations.

This is a frustrating problem because the representations of the atmosphere's governing equations used in models are not fixed to a precisely known set that represents the real world. Since

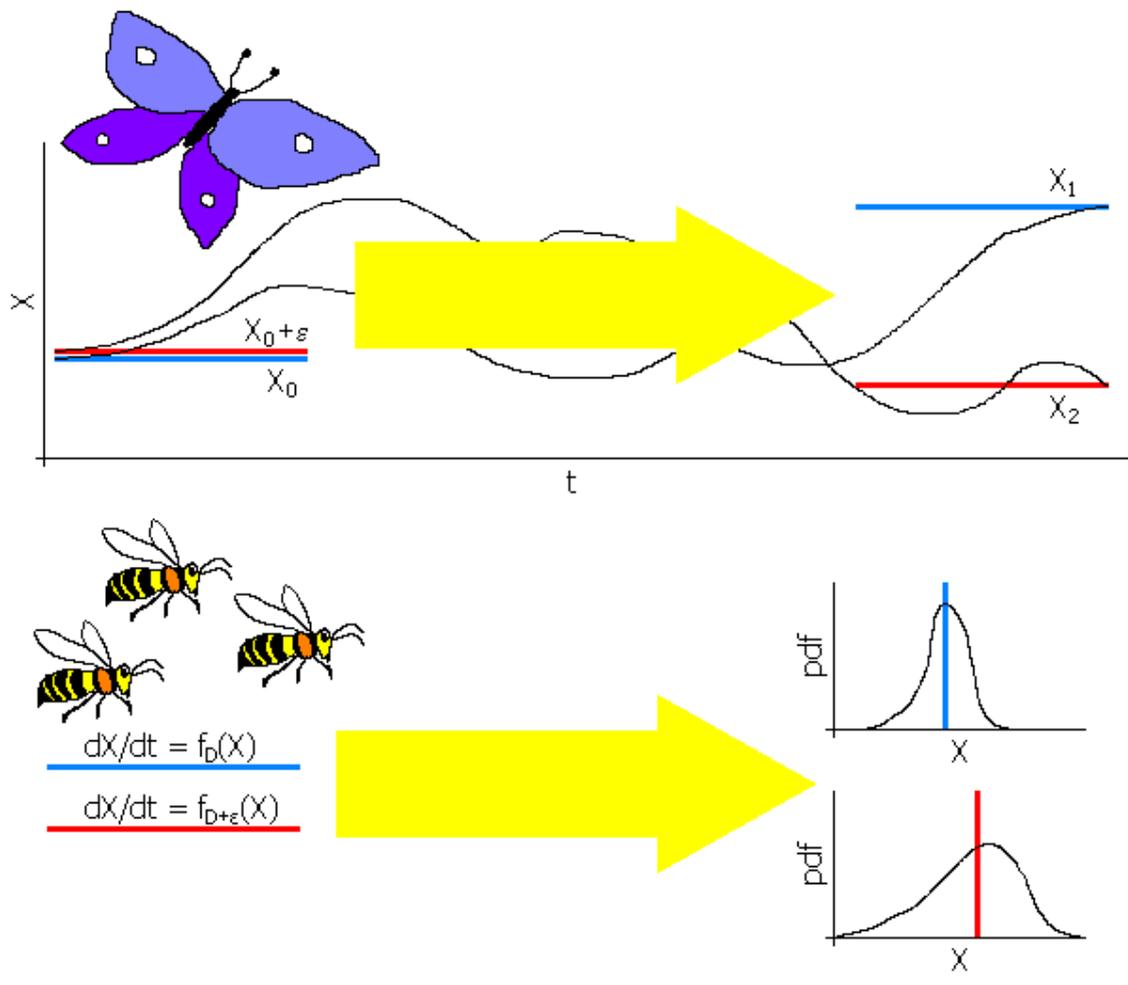


Figure 1.1: Schematics showing (top) the sensitivity of the future state of a system to its initial conditions, and (bottom) the sensitivity of a system's climate to its governing equations.

any simulation of the atmosphere must employ approximations in order to represent a continuous fluid with a finite set of variables in a computer program, all climate models contain some degree of error which could potentially manifest in an amplified way in the model's climatic behaviour. Even if a simulation gives a proper conservative treatment of mass, heat, momentum and moisture and can accurately predict the evolution of the weather over a few days, one cannot assume that it will accurately reproduce the climate. Indeed, if it is run over a longer time, even small inaccuracies in its representation of the governing equations may cause it to drift in to an unrealistic state and produce a climate which differs substantially from that in the real world. Even current state-of-the-art climate models do exhibit considerable biases in this way. For example, see the

IPCC fourth assessment report (AR4), Working Group I, chapter 8, on “Climate Models and their evaluation” (Randall *et al.* 2007).

Of course, deficiencies in the simulation of the atmosphere are by no means the only cause of biases in climate models. Inaccuracies in the modelling of the ocean, land-surface, cryosphere, biosphere, aerosols and chemistry all have roles to play, as does the way these components are coupled to each-other and uncertainties in external climate forcings. Non-linear coupled interactions between all these climate-system components can vastly amplify biases and raise the challenge of climate modelling to a whole new level of difficulty and complexity, but even if a state-of-the-art atmosphere-only model is run with all of the boundary conditions accurately specified, substantial climatic biases occur (this is discussed in the following section 1.3). It is these issues, which are internal to the atmosphere and relate to our understanding of it, with which this thesis is concerned.

In weather forecasts, it is now common practice to quantify the uncertainty caused by the growth of errors in the measurement and analysis of the initial state of the atmosphere by running an ensemble of forecasts in which each member has a different small perturbation added to its initial conditions, corresponding to the margin of error in the observed initial state. Since the availability of computational resources limits the number of ensemble members in these Ensemble Prediction Systems (EPS), it is important to know which possible initial perturbations the atmosphere’s future states will be most sensitive to, so that the divergent futures associated with these possibilities are not missed. Fortunately there are now well established analytical methods to determine the most sensitive modes of perturbation, such as Singular Vectors (see ch. 6 of Kalnay 2003) or Ensemble Transform Kalman Filters (Wang & Bishop 2003). Similarly, climate modellers want to know precisely which perturbations to a model’s formulation the model’s climate will be most sensitive to, so that efforts to improve model formulation can be well targeted. Unfortunately, to the author’s knowledge there is no analytical shortcut to calculate this information, so the only way forward for the climate modelling community has been to carry out a vast number of integrations with a plethora of different experimental model configurations. This has led to a wealth of literature identifying which atmospheric processes are poorly represented in models, and the implications of these deficiencies for simulations of the climate.

1.3 THE ROLE OF UNRESOLVED PROCESSES

As discussed in the previous section 1.2, the primary difference between the governing equations in atmospheric simulations and those in the real world is the finite resolution of the former. Of course the numerics of simulations are designed with the aim of quickly and smoothly converging towards the true solution of the model equations with increasing resolution. However, the residual inaccuracies will likely suffer large non-linear amplification, causing the simulated climate not to converge, if the resolution is too coarse to resolve some feature of atmospheric flow which plays a role in determining the climate. The atmosphere components of the current generation of climate models typically have a horizontal resolution of the order 100 km and a time-step of the order 30 minutes. This is not sufficient to resolve a number of key atmospheric features, including boundary layer turbulence, moist buoyancy-driven convection, the inhomogeneity of clouds and their interaction with radiation, cloud microphysical processes, a significant part of the gravity wave spectrum, and some forms of 2D turbulence in the large-scale flow (see figure 1.2). If these sub-grid processes were not represented at all in a model, obvious large inaccuracies would result. Fortunately, they can at least to some extent be predicted from larger-scale features of atmospheric flow, so modellers have had considerable success representing their effects on the resolved state of the atmosphere using parameterisations which are themselves functions of the resolved state (as represented by the state at individual model grid-points, or vertical columns of grid-points).

These parameterisation schemes typically combine well-defined physics with assumptions of dubious validity and empirical data in order to construct closed calculations of the “sub-grid tendencies” in atmospheric variables such as temperature, humidity and wind velocity. Since the above-grid scale component of the atmosphere’s dynamics that is resolved by a model is comparatively well-defined by its representation of the governing equations, the errors associated with a model’s formulation can largely be attributed to inaccuracies in the sub-grid tendencies, and hence to the assumptions and approximations made in the model’s sub-grid parameterisations. Parameterisations are therefore a major focus in efforts to improve the performance of climate simulations, and state-of-the-art models host an increasingly sophisticated and varied array of these schemes.

But even the most accurate parameterisation sets yet developed have substantial deficiencies,



(a) Heat and moisture mix from the surface in turbulent motions with scale a few m, visible in steam fog on Whiteknights Lake, Reading, October 2005.



(b) Air rises out of the boundary layer in a buoyant moist plume roughly 10 km across, near Cheltenham, August 2009.



(c) A layer of altocumulus displaying inhomogeneity on a range of scales, from a few km down to a few 10's of m, near Taunton, July 2010.



(d) Ice particles forming in altocumulus grow and fall from the clouds but sublimate before reaching the surface, over the Isles of Scilly, June 2008.



(e) Cloud forms on the crests of horizontally propagating gravity waves with wavelengths around 1 km and 100m over Reading, September 2009.



(f) A broad swirl pattern in stratocumulus visible from an aircraft indicates an eddy with a horizontal scale of several 10's of km, over the N Atlantic, August 2009.

Figure 1.2: Six examples of common features of atmospheric flow which are not resolved at a typical climate GCM resolution; (a) boundary layer turbulence, (b) moist convection, (c) cloud inhomogeneity, (d) cloud microphysical processes, (e) gravity waves, and (f) 2D turbulence in the large-scale flow. Photos courtesy of Helen Whitall.

which manifest as substantial biases in climate models. Randall *et al.* (2007) give an overview of biases that are common to many current climate models (those used in the IPCC's 4th assessment report). As they discuss, some of the most serious biases are associated with errors in the simulation of the tropical circulation, such as the large errors in annual mean rainfall patterns over all three tropical ocean basins.

Some of these biases are reduced or removed when Sea Surface Temperature (SST) patterns are constrained to near those observed. For example, Dai (2006) shows that the spurious double ITCZ pattern common to most coupled simulations does not occur in those models which employ surface flux-corrections to maintain realistic SSTs. Errors in the ocean model-component and its coupling to the atmosphere presumably play a role in these particular problems.

However, many other serious biases in tropical rainfall are purely atmospheric in origin. For example, Slingo *et al.* (2003) show that a climate model's dry-bias over the Maritime Continent associated with an under-estimation of the strength of the ascending branch of the Walker circulation in that region occurs even when SSTs are prescribed according to observations. They attribute the bias to the model's failure to represent interactions between sea-breeze circulations and deep moist convection. Also, Neale and Slingo (2003) show that these biases are not reduced even under a three-fold increase in a model's horizontal resolution. If the biases were caused by the coarse-grained representation of the pattern of land, sea and orography or some element of large-scale atmospheric dynamics, one would expect this increase in resolution to reduce them. But processes that are parameterised in climate models are still not resolved at the enhanced resolution. This supports the notion that such biases are caused by deficiencies in the parameterised representation of sub-grid processes.

Aside from biases in the pattern of mean rainfall, a number of important aspects of atmospheric variability in the tropical circulation are poorly simulated in most models. One of the key deficiencies is in the diurnal cycle of convective cloud and rainfall over both land and oceans. As shown in Yang and Slingo (2001), models tend to phase the diurnal cycle of deep convection several hours earlier than in the real world. Another common failing in the simulation of tropical rainfall variability is the tendency for models to over-estimate the frequency of rainfall in general, but to produce too much light rain whilst under-estimating the frequency of heavy rainfall events. A number of current state-of-the-art climate models are shown to have this problem in Dai

(2006). Both of these common model-errors are clearly linked directly to the parameterisations which generate convection, clouds and rain in the models.

There has also been a persistent failure by models to realistically simulate the Madden-Julian Oscillation (MJO). In models, this convectively coupled equatorial wave usually does not propagate coherently as it does in the real world, or it propagates at the wrong speed. Also, its amplitude is typically under-estimated so that it fails to stand out as a leading mode of tropical rainfall variability as observed. These and other errors in the simulation of the MJO are examined in IPCC AR4 climate models by Lin *et al.* (2006). They note that the ability of models to simulate the characteristics of the MJO has no systematic dependence on a model's horizontal resolution, but does exhibit dependence on the type of convection parameterisation used. Similarly, Rajendran *et al.* (2008) find that even increasing the horizontal resolution to a 20km grid-size does not improve the simulation of the MJO, and attribute its unrealistic weak amplitude and lack of coherent propagation in their simulations to "the basic deficiency of the parameterised convection". Inness *et al.* (2001) do find an improvement in the simulation of the MJO when a model's vertical resolution is increased. They attribute this to the resulting better representation of the stable layer around the melting level, which promotes the formation of cumulus congestus-type convection which terminates at this level and produces little rain, resulting in a moistening of the mid-troposphere during suppressed phases of the MJO. This observed mode of convection, which they suggest is important in reinforcing the MJO phase cycle, did not occur in the model with the lower vertical resolution. However, as they suggest in their conclusions, this merits a redesign of convection parameterisations so as to properly represent the congestus convection mode.

In summary, there are a number of key failings in climate models identified in the literature which are thought to be associated with erroneous or inadequate representation of unresolved processes in the atmosphere. It is perhaps unsurprising that the tropics are the region of the world where the circulation is most sensitive to errors in atmospheric parameterised processes; the tropical circulation is largely determined by the influence of the SST pattern (this is demonstrated in an aqua-planet framework by Neale & Hoskins 2001b), which is mediated through boundary layer mixing, the formation of boundary layer cloud, and deep convection, which are all parameterised in climate models. In order to improve the ability of models to reproduce the Earth's observed climate (and predict future climate), the key deficiencies in the sub-grid parameterisa-

tions for such processes must be addressed. This thesis is concerned with a particular deficiency that is common to the behaviour of many important sub-grid schemes used in climate models; the unrealistic nature of their variability near the scale of a model's grid-size and time-step length. This is discussed in the following section 1.4.

1.4 THE ROLE OF HIGH-FREQUENCY / SMALL-SCALE VARIABILITY

Despite the increasing success of parameterisations in producing realistic average sub-grid tendencies and climate mean-fields, many have been reported to produce deficient or unrealistic variability, especially on short timescales (that is, the shorter timescales resolvable by climate GCMs; of the order a few hours). In this regard, the main parameterised process discussed in the literature, and the most important one for generating high-frequency variability in the atmosphere (on the aforementioned timescales), is cumulus convection. As described in the previous section, convection parameterisations commonly fail to produce a realistic diurnal cycle of rainfall, and often yield unrealistic frequency distributions for the intensity of rainfall.

Shutts & Palmer (2007) found that the diabatic heating rates from a Cloud Resolving Model (CRM) simulation, if coarse-grained to a typical GCM resolution, still had much broader PDFs than the tendencies from a GCM convection parameterisation. Since the CRM explicitly resolves deep moist convection, it is assumed to give a reasonable representation of the variability of convective heating and rainfall, which the parameterisation severely underestimated in comparison.

Ricciardulli & Garcia (2000) suggest that momentum transported vertically by the smaller-scale gravity wave modes (which are excited by high-frequency variability in convective heating in the tropics) is important in generating observed low-frequency modes of variability in the middle-atmosphere, such as the Quasi-Biennial Oscillation (QBO).

It should be emphasised that the spectrum of convective heating in the troposphere is by no means the single primary factor influencing the QBO. Processes known to be of great importance include the excitation of a broad spectrum of wave modes (convective heating is one source, but there are others such as flow over orography), the vertical propagation of these waves, wave-

breaking in the stratosphere, and the stratosphere's response and internal dynamics. Studies such as Ricciardulli & Garcia (2000) point out that models must have sufficient vertical resolution in the stratosphere, with sufficiently little diffusion, in order to represent these processes. Employing a non-hydrostatic dynamical formulation is also vital for simulating the resolved-scale wave-modes, and good parameterisations for sub-grid gravity waves are important.

However, Ricciardulli & Garcia (2000) do suggest that the deficient variability of convective parameterisations, and the resulting under-estimation of high-frequency gravity-wave activity, may explain why many GCMs poorly reproduce the QBO even when the stratosphere is well-resolved. They also hypothesise in their conclusions that convective variability may play an important role in exciting tropospheric modes such as the MJO through similar mechanisms.

Horinouchi *et al.* (2003) dispute the notion that most convective parameterisations underestimate high-frequency variability, showing that the range of models in their study are actually scattered around the best observational estimate of the frequency power spectrum of tropical precipitation. However, in their results it appears that those models which employ penetrative mass-flux convection schemes, and which produce the most realistic mean tropical precipitation patterns, are the ones that tend to under-estimate the convective variability. This view was supported by Amodei *et al.* (2001), who note in their conclusions that GCMs which use simple Moist Convective Adjustment schemes are able to generate a full spectrum of tropical wave activity close to that observed, whilst those which employ more sophisticated convective parameterisations tend not to.

Some models do indeed exhibit more realistic amounts of high-frequency variability associated with sub-grid processes, especially convection, but this is often unrealistic and can be generated by numerical artefact rather than successfully representing the high frequency variability of the relevant physical processes. A common example of this is the unrealistic discrete “on-off” behaviour of many mass-flux convection parameterisations. This is noted by Stiller (2009), who identified this spurious numerical noise behaviour as a major difficulty when developing a linearised version of the Met Office Unified Model's convection scheme for use in a Tangent Linear Model (TLM)¹. The mechanisms behind such behaviour are studied in detail in the Unified Model

¹A linearised version of a forecast model which is a vital component in 4D-Var data assimilation, and must behave smoothly to perform well

by Willett & Milton (2006) in their section 5.2 on the time-step by time-step behaviour of the convection scheme. Each time the parameterisation produces deep convection, it warms the column above the initiating layer through compensating subsidence, but cools the levels beneath it via downdrafts and evaporation of precipitation. This creates a stable convective inhibition layer at the source layer which must then be eroded by boundary layer mixing or shallow convection over the following time-step(s) before deep convection can trigger again.

Scinocca & McFarlane (2004) suggest that in models with penetrative mass-flux convection schemes, much of the variability in tropical rainfall and latent heating is actually driven by the large-scale precipitation scheme (the parameterisation that simulates rainfall associated with the resolved circulation, rather than unresolved features), whilst the convection parameterisation exhibits very deficient variability indeed. However, they show that the amount of high-frequency variability generated by the convection scheme can be greatly increased by altering the details of the scheme. In particular they find that if a prognostic closure for the convective mass-flux is used, the scheme generates realistic amounts of high-frequency variability. In the standard convection scheme in the model used in their study, the closure relates the convective mass-flux directly to Convective Available Potential Energy (CAPE), as in most so-called CAPE-closure schemes. However, they modified it by relating the rate of change of the mass-flux to CAPE instead, resulting in a large increase in the variability of the scheme, without worsening the modelled pattern of mean precipitation.

Approaches such as the Scinocca & McFarlane (2004) prognostic CAPE closure scheme essentially force the convection to respond to the generation of convective instability on longer time-scales (of order a few hours as opposed to one time-step), suppressing the near-instantaneous removal of instability and time-step to time-step “on-off” noise that occurs in many CAPE-closure mass-flux convection schemes. This should allow any instability that arises to persist over longer timescales before it is removed, which should lead to a stronger, more coherent atmospheric response to the convective instability. It is therefore perhaps unsurprising that such an approach amplifies the high-frequency variability overall, despite suppressing the unphysical variability at the very fastest timescales. However, the physical justification for forcing parameterised convection to respond on longer timescales is unclear, given that timescales of less than an hour are consistent with typical cumulus plume life-times in observations and Cloud-Resolving Models.

One could interpret such efforts as an ad hoc means of tuning an existing numerical source of variability (the on-off noise of the convection parameterisation) to behave more like the unrepresented physical sources. Also, some models may be more prone to numerical instability associated with “grid-point storms” if convective instability is allowed to build up over longer timescales.

Variability in convection and tropical rainfall are not the only issues concerned with high-frequency variability. Many models display unrealistically low mesoscale eddy activity in the mid-latitude storm-tracks, due to effects of the grid-scale truncation on the energy spectrum, the numerical diffusion associated with the dynamical scheme, or artificial diffusion which is added to aid numerical stability. This is evident in the unrealistically rapid drop-off in the energy-spectra of model winds at small scales (e.g. Shutts 2005), relative to observed spectra which actually become less steep at small scales (Gage & Nastrom 1986). This deficiency is seen even at synoptic scales much larger than the grid-resolution and worsens as the grid-scale is approached. It has been suggested that this could play an important role in the common failure of GCMs to predict the frequency of mid-latitude atmospheric blocking events (a phenomenon characterised by the occurrence of a strong, persistent anticyclone in the mid-latitude storm track and the interruption or reversal of the prevailing Westerly winds).

D’Andrea *et al.* (1998) found a systematic tendency for atmospheric models to underestimate the occurrence of blocking in a study of 15 different GCMs as part of the Atmospheric Model Intercomparison Project (AMIP). Interestingly, their study did not suggest any tendency for the models with higher resolutions to give better simulation of blocking frequency. Although model skill in forecasting blocking has in fact improved over subsequent model upgrades, under-prediction of blocking has remained a key deficiency in medium-range forecasts despite increases in model resolution. For example, Pelly & Hoskins (2003) found that the ECMWF ensemble prediction system gives a modest systematic under-prediction of Northern hemisphere blocking frequency, which worsens with increasing forecast lead-time, suggesting that the ensemble drifts towards a climatological blocking frequency significantly lower than that observed.

Meanwhile, a number of studies have shown that momentum transports by synoptic and smaller scale eddies in the mid-latitudes play a key role in the formation and maintenance of atmospheric blocks (e.g. Luo *et al.* 2001), and that the amount of high-frequency variability in general could be crucial in determining the ability of the mid-latitude circulation to transition in

to a blocked state with the observed frequency (Molteni & Tibaldi 1990). It has therefore been suggested that up-scale backscatter of variability at unresolved scales on to resolved scales may be vital in reproducing the observed energy spectra for atmospheric flow, and in turn the frequency and nature of mid-latitude blocking events (e.g. Jung *et al.* 2005). This has motivated the development of schemes to reinject the kinetic energy dissipated by unphysical diffusion in models back in to the resolved flow; these are discussed in section 1.5. One could think of the unphysical diffusion present in a GCM as a crude parameterisation for mixing by the unresolved eddies. And just as with the convective parameterisation, there is then a strong argument for improving it in order to better represent the high-frequency, small-scale variability associated with the sub-grid features.

There are sound theoretical reasons to expect that parameterisations will underestimate the variability associated with unresolved processes. Lin & Neelin (2002) state in their introduction that “In the atmosphere it is reasonable to hypothesise that for a given large-scale temperature and moisture field, there is a contribution to the variability of convection that arises inherently from small-scale motions, but which are not well represented by large ensemble means.” The “ensemble” referred to here is the ensemble of sub-grid features (e.g. convective plumes) that would be present in the real atmosphere within an area equivalent to a model grid-box. Parameterisations that represent the effect of sub-grid features on the resolved state purely in terms of the resolved variables must consider the features they represent in an ensemble mean sense, since they know nothing about the sub-grid state or the associated variability in the ensemble of features. As discussed by Williams (2005), this is a valid treatment of the unresolved features provided that the grid-box area would contain a sufficiently large ensemble of them to yield a stable ensemble-mean statistic. Such a case, in which a sample is large enough that its statistics are stable and reliably match those of a hypothetical infinite background population, is known as a statistical equilibrium.

However observations indicate that, whatever model resolution is chosen, the real atmosphere does not exhibit any clear “scale-break” between atmospheric features which are resolvable and those that are small enough to parameterise according to a statistical equilibrium assumption. For example, Gage & Nastrom (1986) presented wavelength power-spectra for winds in the upper troposphere which show a continuous spectrum of atmospheric variability, with spectral power

decreasing according to a split power-law from the planetary scale down to a few kilometres. Therefore as Williams (2005) points out, whatever resolution a model employs, the most energetic unresolved features will likely be those with spatial scales just below the smallest resolvable scale (in the case of climate GCMs, the most energetic unresolved features are organised convective systems). Thus only a small number of these features can be contained in a grid-box area, and large deviations from statistical equilibrium will be inevitable. This has the consequence that the effect of unresolved processes on the resolved state cannot be expressed purely as a function of the resolved state, but will inevitably have some variability dependent on fluctuations in the state of the ensemble of sub-grid features. The sub-grid state is of course undefined in models, so the component of grid-scale variability associated with it is not represented by the parameterisations and is absent in atmospheric simulations.

In summary, a range of modelling and observational studies suggest that atmospheric GCMs have difficulty in reproducing the variability observed in the real atmosphere at the smaller space- and time- scales they can resolve. This model deficiency has commonly been found to be highly sensitive to the details of the sub-grid parameterisations employed, suggesting it is indeed associated with common failings of sub-grid parameterisations. Theoretical considerations regarding the formulation of sub-grid parameterisations, which contain no information about the uncertainty of the sub-grid state in each grid-box, could explain the ubiquity of this problem. Some studies have also suggested that the deficiency of high-frequency variability can cause larger, slower modes of atmospheric variability, such as the QBO, the MJO and mid-latitude blocking, to be poorly represented in models as well. Thus this issue not only relates to the realism of the statistics of rainfall and other atmospheric variables on short time-scales, but has implications for the fidelity of climate simulations overall. This has motivated numerous efforts to improve the variability of the parameterisations used in GCMs on small space- and time- scales in recent years, by introducing random elements to them to account for the inherent uncertainty in the sub-grid processes they represent. These so-called stochastic parameterisations are discussed in the following section 1.5.

1.5 STOCHASTIC PARAMETERISATIONS

The stochastic parameterisations so-far developed broadly tend to fall into three categories, in terms of their basis and motivation: those designed to account for modelling uncertainties in Ensemble Prediction Systems, those which aim to represent physical sources of variability associated with likely variations in the unknown sub-grid state, and those which attempt to counter the unphysical effects of the grid-scale truncation in models through "backscatter" of unphysically diffused kinetic energy back on to the resolved scale. Each of these is described below, with examples from the literature given. As we will see, there turns out to be a certain amount of overlap between these three motivations. For example, sub-grid *variability* is itself a source of forecast *uncertainty*, so stochastic schemes designed to address the former are routinely implemented in ensemble weather forecasts to account for the latter.

1.5.1 MODEL UNCERTAINTY SCHEMES

1.5.1.1 PRAGMATIC BASIS

As noted in section 1.2, in ensemble prediction systems it is desirable to sample the range of possible future realisations as fully as possible, with a spread of ensemble members consistent with the true uncertainty of the forecast. If this is achieved, the spread of ensemble members at some lead-time (as measured by their standard deviation from the ensemble mean) should be equal to the root-mean-square error of the ensemble mean forecast (relative to the actual observed state of the atmosphere at that future time). Traditionally, the uncertainty in the forecast was quantified entirely by the margin of error in the analysis state used to initialise the forecast, and the growth of this error with time.

However, in practice even the most rigorous treatment of initial condition uncertainty (accounting for observational error, the sparsity or lack of representivity of observations, and error in the data assimilation system used to generate the analysis) tends to yield ensemble predictions in which the error of the ensemble mean is on average somewhat larger than the ensemble spread. This means the future observed state falls outside the range predicted by the ensemble more often than it statistically should if the ensemble were an unbiased sample of possible future realisations.

This has two possible interpretations; either some genuine source of uncertainty has not been accounted for in the spread of ensemble members, or the ensemble members are systematically biased away from the atmosphere's true evolution due to systematic biases in the model. In fact, both are almost certainly the case. As discussed earlier, there is an unaccounted-for source of uncertainty in that the resolved-scale effects of unresolved processes at-least partly depend on the unknown sub-grid state. And of course atmospheric GCMs do indeed exhibit systematic biases.

As also aforementioned, these biases can largely be associated with deficiencies in sub-grid parameterisations. Arguably, any sub-grid parameterisation must make assumptions not based on physical laws in order to obtain closure, and it is unclear what the most appropriate choice of assumptions for a given parameterised process should be. A pragmatic approach is to treat the model error associated with sub-grid parameterisations as an additional source of forecast uncertainty; that associated with the uncertain formulation of the optimal parameterisation set possible, given the limits of current scientific understanding and computational resources. If a single, fixed, deterministic set of parameterisations is chosen for all members of an ensemble forecast, this "parameterisation uncertainty" has not been sampled adequately, and one would expect the ensemble to have deficient spread.

These considerations have led to the development of stochastic schemes aiming to sample the structural uncertainty of model parameterisations within ensemble forecasts. They are designed simply to pragmatically account for this source of uncertainty, analogously to the initial condition perturbations used to account for initial condition uncertainty. As such there is no direct requirement for them to behave like known physical sources of variability, provided that they don't induce unfavourable model biases. Indeed, it is important that such a scheme samples the plausible range of tendencies that could be produced by sub-grid processes at each time-step as fully as possible. Thus fairly broad, generic methods are often considered appropriate for this type of stochastic parameterisation.

The un-accounted-for forecast uncertainty associated with the unknown sub-grid state doesn't immediately appear to be represented by such a scheme. However, this uncertainty is likely to manifest in the deficient variability of sub-grid parameterisations, which, as discussed in the previous section, is thought to contribute to other systematic model biases. And many stochastic methods aimed at addressing parameterisation uncertainty do so by boosting the variability of the

parameterisations. So in practice, the theoretically-based concept of sub-grid state uncertainty and the more pragmatically-based concept of parameterisation uncertainty have a high degree of overlap. As such, many stochastic schemes designed to represent specific physical sources of variability, or backscatter unphysically diffused energy on to the resolved flow (described in later subsections) are often applied in ensemble forecasting contexts to account for associated uncertainties. Here however, we discuss methods aimed specifically at sampling model uncertainty.

1.5.1.2 PERTURBED PHYSICS AND MULTI-MODEL APPROACHES

Before discussing stochastic methods, it should be mentioned that many studies have aimed to represent modelling uncertainty in ensemble forecasts without introducing any stochastic element. Instead, they sample a range of different deterministic model configurations within an ensemble. Such systems are often referred to as “perturbed physics” ensembles. Some examples from the literature are discussed below.

Houtekamer *et al.* (1996) created an ensemble forecast system in which different versions of a model’s convection, radiation, and gravity-wave drag schemes were applied in each ensemble member. Whilst this “multi-parameterisation” approach yielded some increase in ensemble spread, they found the ensemble was still under-spread relative to observations, and concluded there were further sources of uncertainty not captured by it.

Mylne *et al.* (2001) showed that an ensemble combining forecasts from the Met Office Unified Model and the ECMWF Integrated Forecast System out-performed both of the two centre’s individual forecast systems, giving weight to the idea of representing model uncertainty through the use of a multi-model ensemble.

Perhaps the most-used perturbed physics ensemble method is to sample the range of plausible values for free parameters within a given parameterisation set. The assumptions made to obtain closure in sub-grid parameterisations usually involve the introduction of parameters, which control the rate or relative importance of some aspects of the sub-grid process being modelled. For example, convection parameterisations often contain a prescribed timescale over-which convective instability is removed, whilst cloud schemes often contain a parameter controlling the degree of sub-saturation at which cloud may begin to form due to sub-grid inhomogeneity in the

moisture-field. The values of such parameters cannot be defined exactly from theory and must be estimated, based on observations, high-resolution modelling studies, or by tuning them so as to optimise model performance. Such considerations may place some constraints on parameter values, but in most cases a broad range of values can still be considered appropriate. By assigning different parameter values sampled from within this range to each member of an ensemble, one can account for the associated uncertainties in a forecast.

Of course, this paradigm has weaknesses. Many aspects of structural uncertainty in parameterisations cannot be accounted for through parameter uncertainty (for example, we would not expect that adjusting the parameters of a convection scheme closed on CAPE could encompass the behaviour of an alternative formulation closed on moisture convergence). Also, the appropriate range of values to sample for a given parameter may be no better defined than the value itself. However, modellers have had considerable success in accounting for at least a substantial component of model uncertainty by sampling parameter uncertainty.

Yang & Arritt (2002) tested an ensemble seasonal forecast of rainfall over North America, in which each ensemble member had a different fixed combination of values for the parameters in its convection scheme. They found that the mean of this “perturbed parameter” ensemble outperformed the forecast with default parameter values.

Murphy *et al.* (2004) also employed a perturbed parameter ensemble approach, to simulate the climate-response to a doubling of atmospheric CO₂. Their ensemble included perturbed parameters in the parameterisation schemes for all the sub-grid processes in the atmosphere climate model HadAM3, coupled to a mixed-layer ocean model. Using this approach, they demonstrated that the uncertainty in the climate response to a doubling of CO₂ at the regional scale was considerably larger than estimates based on scaling the response from an individual model configuration.

In Murphy *et al.* (2004), each ensemble member had only one of the model parameters perturbed from its default value; their distributions for climate sensitivity were based on the assumption that the effects of perturbations to different parameters would combine linearly and independently. Stainforth *et al.* (2005) and Piani *et al.* (2005) expanded the approach by running multi-thousand member perturbed physics ensembles, using the idle time on PCs belonging to thousands of volunteers around the world, in the climateprediction.net project. With the much

larger ensemble size, a more comprehensive sampling of the possible combinations of values for different model parameters was possible, and non-linear interactions between perturbations to different parameters could be accounted for.

Whilst ensembles which sample multiple models, configurations or parameter settings deterministically as described above are useful for elucidating the sensitivity of predictions to model uncertainty, they do not allow each ensemble member to explore the full range of model uncertainty being sampled. Rather, each member (and correspondingly each version of the initial state) accesses only one of the possible model configurations. In climate predictions, where the forecast has little sensitivity to the initial state, this is of little consequence. But in an ensemble weather forecast, it would be desirable to sample the full range of model uncertainty for each initial state, and vice versa. In principal this could be achieved by running a perturbed-physics ensemble of deterministic initial-condition ensembles, but the total number of ensemble members required to adequately sample both initial conditions and physics in this way may be impractically large.

Another potential problem often pointed out in the literature with the multi-model or perturbed physics approach is that each ensemble member in such a system has a different, unique attractor and different biases. This is beneficial if one is aiming to sample the model-uncertainties in the attractor (i.e. in climate predictions). But it yields an ensemble whose members are not equally skilful, due to their differing biases, making it difficult to obtain unbiased probabilistic forecast information.

1.5.1.3 STOCHASTIC PHYSICS SCHEMES

An alternative, natural extension of the perturbed physics approach is to sample model uncertainty in a random, time-varying manner within each member of an initial-condition ensemble. Thus the ensemble members are qualitatively alike and share the same attractor. This has led to the development of stochastic parameterisations in ensemble forecasts, as described earlier. In such schemes, model uncertainty is no longer sampled in a time-mean sense, but its effects on the time-wise evolution of weather systems are better sampled. As such, we expect a “stochastic physics” ensemble to have a less-spread climatology than a perturbed physics ensemble, but diverge towards its climatological spread more rapidly. In a weather forecasting context this is

beneficial. It should avoid the tendency for some perturbed physics ensemble members to drift towards divergent, unrealistic climatologies (this had to be accounted for in perturbed physics climate studies such as Stainforth *et al.* 2005), and it should give larger, more representative ensemble spreads at weather-forecast lead-times.

Such a method was pioneered by Buizza *et al.* (1999), who describe a “Multiplicative Noise” stochastic physics scheme which was implemented operationally in the ECMWF’s medium-range ensemble forecasting system. In their scheme, the net effect of all the model’s parameterisations is considered to carry uncertainty proportional to its magnitude. To account for this, the net tendencies (in temperature, moisture and winds) produced by all parameterised processes are multiplicatively scaled by a random factor which varies in time, space, and between ensemble members.

Sub-grid processes, such as deep convection, often exhibit coherent organisation on scales greater than the grid-size / time-step, and Buizza *et al.* (1999) felt that the perturbation scalings applied in their stochastic scheme should reflect this. They therefore applied autocorrelation by holding the perturbation scalings constant over “tiles” of multiple neighbouring grid-points / consecutive time-steps. The scheme then has three tuneable parameters; the amplitude of the perturbation scaling’s departures from unity, its autocorrelation timescale, and its autocorrelation spatial scale. They chose values for these so-as to optimise the probabilistic skill of the ensemble. They found the stochastic scheme increased the ensemble spread at all lead-times between 0 and 10 days, especially in the tropics, and gave significant improvements in the skill of the ensemble forecast.

A similar stochastic multiplicative perturbation method is now also used in the ensemble forecast system run by the Meteorological Service of Canada, as described by Charron *et al.* 2010. But unlike the Buizza *et al.* (1999) scheme, they perturb the parameterised tendencies using a random pattern generated using spherical harmonics. Using this scheme, combined with a stochastic backscatter scheme, the ensemble produced much more representative spreads than its predecessor (which represented model uncertainty using a multi-parameterisation approach), such that a previous artificial inflation of the initial condition perturbations is no longer needed. The ECMWF ensemble has also recently been updated to generate perturbations using spherical harmonics (rather than the “tiling” method described in Buizza *et al.* 1999) and employ a

stochastic backscatter scheme alongside the Multiplicative Noise scheme (Palmer *et al.* 2009).

Another simple stochastic physics method for sampling model uncertainty in an ensemble forecast is the time-varying form of the perturbed parameter ensemble approach described earlier. Arribas (2004) describes such a “Random Parameters” stochastic scheme, which has been implemented operationally in the Met Office Global and Regional Ensemble Prediction System (MOGREPS, a short-range ensemble forecast system, described by Mylne *et al.* 2005). MOGREPS, and the stochastic schemes it uses, have also been described more recently in the peer-reviewed literature (Bowler *et al.* 2008,2009).

To sample parameterisation uncertainty, various model parameters are varied randomly in time and between ensemble members, with the values drawn from within plausible ranges specified according to advice from experts on each of the relevant parameterised processes. As in the “Multiplicative Noise” stochastic scheme of Buizza *et al.* (1999), it was felt that there should be some autocorrelation between perturbations from one time-step to the next. In this case, autocorrelation is specified by having the model parameters vary according to a first-order auto-regression model. Unlike the Buizza *et al.* (1999) scheme, there is no spatial component to the variability of the stochastic perturbations; complete spatial autocorrelation is forced by applying the same parameter values at all grid-points each time-step. As with the Buizza *et al.* (1999) scheme, the Random Parameters scheme was found to have a positive effect when introduced to an initial-condition ensemble, giving a significant increase in ensemble spread.

The Multiplicative Noise and Random Parameters schemes described above consider model uncertainty in a generic manner, applying stochastic perturbations consistently across all parameterised processes. Some other methods have treated model uncertainty in different parameterised processes using different methods, often with a focus on the convection parameterisation.

For example, Bright & Mullen (2002) - added a stochastic term to the vertical velocity input to a convection scheme’s trigger function, to account for uncertainty in where and when deep convection should occur. They also stochastically perturbed the critical Richardson number in the boundary layer scheme, to account for uncertainty in transitions between different turbulence regimes. These stochastic schemes were applied in ensemble forecasts of the North American SW monsoon, and found to give some increase in the spread and probabilistic skill of the ensemble.

Similarly, Li *et al.* (2008) demonstrated a regional ensemble prediction system, perturbing the convection scheme as in Bright & Mullen (2002), and also the threshold relative humidity for cloud-formation due to sub-grid moisture inhomogeneity. They perturbed these using a stochastic forcing field based on spherical harmonics, so that autocorrelations in space and time could be specified in a qualitatively realistic way. When testing the impact of various upgrades to the ensemble system, they found that the stochastic physics schemes produced the greatest improvements in ensemble skill for dynamical variables.

Teixeira & Reynolds (2008) tested a stochastic scheme which uses multiplicative perturbations similar to those in the Buizza *et al.* (1999) scheme, but only applies them to the tendencies produced by a model's convection parameterisation. They found that the scheme initially produced the highest levels of ensemble spread in the tropics, but that the perturbations in the tropics spread to the mid-latitudes and migrated up-scale as the forecast progressed. Again, when the scheme was introduced to an initial-condition ensemble, it was found to increase the spread of the ensemble beneficially, reducing the frequency with which subsequent observations lie outside the range predicted by the ensemble forecast.

1.5.2 PHYSICAL VARIABILITY SCHEMES

As discussed in section 1.4, deficient variability in parameterised processes can cause models to give inaccurate statistics for the variability of observables such as rainfall, and may contribute to systematic model biases. And without any information about the sub-grid state in each grid-box and its likely variations, one would expect some physical sources of variability to be absent from deterministic parameterisations.

Although the sub-grid state is by definition unknown to the modeller, constraints on the statistics of its variability can often be provided by theory, observations and high-resolution modelling studies. New sub-grid parameterisations can then be designed which treat the sub-grid state statistically and account for its likely variability in their closure assumptions. In doing so they aim to introduce stochastic variability on the resolved scale which is physically consistent with an un-biased random sampling of the plausible sub-grid states. This approach should produce realistic parameterised variability on the resolved scale, provided the statistics of the sub-grid state's

variability are accurately specified. Those statistics clearly include the amplitude of the sub-grid variability, but also other factors such as its dependence on the resolved-scale and on the sub-grid state at previous times and neighbouring locations.

Although such schemes are sometimes complex in nature, there has been increasing interest in developing them to address deficient variability and associated model biases in climate simulations. Several such schemes have already been implemented in ensemble prediction systems, so as to accurately account for sub-grid uncertainty in the spread of ensemble members, by selecting a different random sample of sub-grid states in each ensemble member.

Lorenz (1975) foresaw, from a theoretical point of view, a need for the development of stochastic parameterisations to represent variability at unresolved scales, in order to accurately simulate the climate. More recently, Palmer (2001) again showed theoretically that neglecting variability at unresolved scales can lead to large errors in the climatology at well-resolved scales in models. As was confirmed by various studies mentioned in section 1.4, he suggested that some biases exhibited by GCMs could be caused by this neglect. He proposed that addressing this problem required the development of non-local dynamically-based stochastic schemes to represent the effects of sub-grid variability on the resolved scales.

As explained in section 1.4, the most energetic scales of unresolved variability (and which have the greatest impact on resolved flow) will be those just below the smallest resolvable scale. And for weather and climate GCMs, these are dominated by mesoscale variability in convection. As such, most stochastic schemes aimed at representing specific sources of unresolved variability have focused on this.

Among the earlier attempts was the simple stochastic scheme of Lin & Neelin (2000), who added a stochastic term to the CAPE input to the convection parameterisation in an intermediate-complexity tropical atmosphere simulation. This input influenced both the triggering and magnitude of convective events in the model, and the stochastic term followed a first-order auto-regression model. This stochastic scheme is qualitatively similar to some of those developed to represent model uncertainty in ensemble forecasts (described in the previous subsection), but was implemented explicitly to “examine the impacts of convective variance arising intrinsically at the unresolved scales”, rather than to boost ensemble spread to levels more representative of fore-

cast error. They found the scheme had a systematic effect on the model, and could dramatically increase the amplitude of slow modes of variability in tropical rainfall and dynamical fields. However, this response was highly sensitive to the autocorrelation timescale of the stochastic term.

The UK Met Office developed a stochastic scheme specifically to represent the effects of organised convective systems on the resolved-scale vorticity field, (Gray & Shutts 2002). This was implemented in the Met Office Global and Regional Ensemble Prediction System (MOGREPS, later described in the peer-reviewed literature by Bowler *et al.* 2008). The Stochastic Convective Vorticity (SCV) scheme adds a vorticity dipole, with an upper-level anticyclone associated with the anvil outflow and a mid-level cyclone, to the model winds during convective events. The magnitude of the added dipole is related to the CAPE and the precipitation output by the convection parameterisation, and scaled by a random number. It was found to have a neutral effect on the model climatology, but was able to perturb the position and intensity of tropical cyclones and affect the Southern-hemisphere storm-track via sub-tropical convective disturbances that feed in to it.

Plant & Craig (2008) designed a stochastic convection scheme which accounts for the fact that, within a finite grid-size, the ensemble of possible convective plumes maybe poorly sampled. The true sample of plumes present maybe small and thus likely to deviate far from statistical equilibrium and vary considerably (even under constant large-scale forcings). The result is a source of variability which cannot be captured by grid-scale variables alone and so must be represented stochastically.

Most global weather and climate models in use today employ penetrative mass-flux convection schemes which parameterise the ensemble of possible convective plumes within each grid-cell using a “bulk plume” calculation with a single entrainment rate (or two parallel schemes with distinct entrainment rates for “shallow” and “deep” convection). It is generally accepted that the rate of entrainment of air from a convective plume’s environment (per unit mass-flux) is crucially dependent on the size of the plume; small cumulus clouds have much higher entrainment rates than broad cumulonimbus updrafts. In reality, a whole spectrum of plume-sizes may occur, and the overall updraft mass-flux in a given area will be spread over a corresponding spectrum of entrainment rates. If that area is sufficiently large for the ensemble of convective plumes to be considered to be in a statistical equilibrium, the spectrum of entrainment rates on which the mass-

flux occurs should follow some equilibrium distribution. But if the area considered is smaller, then at any given time the finite number of plumes present will represent only a limited sampling of that distribution, and the total mass-flux and entrainment rate spectrum will vary depending on that sampling.

So, the conventional “bulk plume” parameterisation approach doesn’t account for the likely spectrum of entrainment rates, or the likely fluctuations in that spectrum or in the total mass-flux. To overcome these shortcomings, the Plant & Craig (2008) stochastic convection scheme explicitly simulates individual convective elements, modelling within each grid-cell an ensemble of plumes of varying sizes (and corresponding entrainment rates) which interact only via their effects on the resolved-scale variables, as proposed by Arakawa & Schubert (1974). Craig & Cohen (2006) used theoretical considerations and cloud-resolving model data (Cohen & Craig 2006) to find the expected form of the equilibrium distribution of convective plume sizes, and the Plant & Craig (2008) scheme generates a realistic ensemble of plumes by sampling this distribution. The distribution is normalised by the area of the grid-cell, such that more plumes are generated for larger grid-sizes, and the sampling of the distribution becomes more complete with decreasing model resolution. So for very coarse grid-sizes, the scheme tends towards a statistical equilibrium, giving a deterministic spectral convection scheme as in Arakawa & Schubert (1974). But at smaller grid-sizes there may only be a few plumes present in each grid-cell, and large fluctuations in convective tendencies may occur depending on the chance sampling at a given time.

The Plant & Craig (2008) scheme’s response to instability in the environment profile is parameterised by scaling the sampled distribution linearly with Convective Available Potential Energy. Thus the scheme is closed on CAPE as in most mass-flux convection schemes. One problem with the conventional point-wise CAPE-based closures employed by such schemes is that in reality the scales influencing convective plumes may extend beyond the arbitrary bounds of a model grid-cell / time-step. For example, real clouds can affect each-other’s development over distance via gravity-wave propagation, and may persist for longer than a model’s time-step. There is no physical justification for pinning the scale within-which convective plumes can be thought of as having a shared environment to that of the model’s grid-size and time-step. The Plant & Craig (2008) scheme allows this scale to be specified in the CAPE closure, by calculating the CAPE using an environment profile which is averaged over nearby grid-points and recent time-steps

within a specified horizontal distance and time-lag. Thus the scheme can simulate fluctuations of the convection about an equilibrium which exists over a wider scale than the model's resolution.

Plant & Craig (2008) tested the stochastic convection scheme in a Single-Column Model with constant large-scale forcings, comparing its equilibrium behaviour to that of a companion Cloud Resolving Model simulation. The scheme produced stochastic fluctuations in total convective mass-flux as designed, and over time yielded probability distributions of mass-flux consistent with those obtained by spatially-averaging the Cloud Resolving Model simulations.

Keane & Plant (2012) implemented the scheme in a 3-dimensional modelling framework, using a bicyclic 512 km square domain, a range of horizontal resolutions between 16 and 51 km, and constant large-scale forcings as in the SCM in Plant & Craig (2008). In the 3D framework, they found that the scheme only produced stable, realistic distributions of convective mass-flux when the CAPE used in the closure calculation was averaged over a sufficient area and time-period to cover (approximately) a statistical equilibrium. If not enough averaging was employed, the closure calculation was affected directly by the point-wise fluctuations produced by the convection scheme, rather than being based on a wider equilibrium state as the scheme assumes. As a result, the convective mass-flux statistics would be affected by numerical "on-off" triggering behaviour. They found that the closure-averaging needed to cover at least around 33 individual plumes to yield stable convective statistics; the averaging length and time scales required to achieve this vary depending on the degree of convective instability, but amounted to multiple grid-lengths / time-steps in their simulations. If this condition was met, they found that the scheme produced realistic variability and mean-states at a range of different heights and model-resolutions, and gave much more realistic rainfall statistics than conventional deterministic bulk-plume convection schemes.

1.5.3 BACKSCATTER SCHEMES

In fluid-dynamics generally, there is a tendency for kinetic energy to "cascade" to progressively smaller scales, as motions on any given scale are dissipated by the growth of eddies on a smaller scale. Kinetic energy in a turbulent fluid will eventually cascade down to the very small scale at which molecular viscosity dominates, and be dissipated by diffusion. In models, kinetic energy

must be removed much sooner in its down-scale cascade, at the scale of the grid-resolution. In fact, models exhibit numerical diffusion, or have diffusion artificially added to suppress numerical non-linear instability associated with the build-up of energy at the grid-scale truncation. This acts to remove kinetic energy at scales somewhat larger than the grid-size.

This would be of no consequence to the ability of a model to simulate the well-resolved scales at which no such diffusion applies, provided kinetic energy always and only cascades *down*-scale. In practice, although the down-scale cascade is usually dominant, there is some up-scale cascade as well. Or, as put by Frederiksen & Davis (1997), “the larger scales are also randomly forced by their nonlinear interaction with the subgridscale eddies.” So some of the kinetic energy on scales near or below a model’s grid-size would in reality “backscatter” upscale, creating or amplifying motions at well-resolved scales. But in a model this energy is dissipated before the backscatter can take place. As a result, atmosphere models generally under-estimate the amount of kinetic energy in a broad part of their resolvable spectrum, to a worsening degree as the grid-scale is approached (eg Shutts 2005). The implications of this problem for the fidelity of atmospheric GCMs were discussed in section 1.4.

Considerable efforts have therefore been made to design schemes to estimate how much energy has been unphysically diffused from model winds at poorly-resolved scales and re-inject it, so that backscatter of energy from these scales on to the well-resolved scales may occur as in reality. The pattern of winds re-injected in this way should itself be dependent on backscatter from the sub-grid scale. Since the form of this is unknown in the model, backscatter schemes stochastically generate a plausible pattern of winds to re-inject.

One could perhaps think of such schemes as physically-motivated stochastic turbulence closures. Like other stochastic schemes designed to physically represent sub-grid sources of variability, such as those described in subsection 1.5.2, they are being developed both to address systematic model biases and to represent associated uncertainties in Ensemble Prediction Systems.

Frederiksen & Davis (1997) first suggested the use of stochastic kinetic energy backscatter in General Circulation Models (it had already been developed in Large Eddy Simulations; Mason & Thomson 1992), and demonstrated it in an idealised 2D model of global atmospheric flow.

Shutts (2005) then developed a backscatter scheme for use in GCMs, which was tested in the ECMWF ensemble forecast system. In this scheme, energy backscatter is controlled by an estimate of the total energy dissipation, from numerical diffusion, convection, and mountain drag. This flow-dependent dissipation field provides the amplitude for a randomly-generated stream-function forcing field which is added to the model winds. The time-evolving pattern for this forcing is generated using a cellular automaton (that is, an algorithm in which the value at a given grid-cell is a simple function of the values at neighbouring grid-cells and at the previous time-step). The scheme was found to improve the realism of the energy spectrum of winds in the ECMWF model, and improve the probabilistic skill of the forecast by increasing ensemble spread.

Berner *et al.* (2008) then applied the Shutts (2005) Cellular Automaton Stochastic Backscatter (CASB) scheme in seasonal forecasts using the ECMWF model. They found that it significantly improved the skill of the forecast, not only by producing more representative ensemble spread, but by reducing model biases. It lessened the model's tendency to under-estimate the frequency of mid-latitude blocked flow regimes, and reduced some tropical rainfall biases.

Berner *et al.* (2009) improved on the Shutts (2005) scheme by generating the forcing pattern using a specified spectrum of spherical harmonics, time-evolved according to a first-order auto-regression model, making the spectrum of the forcing fully tuneable. Tuning can be done by comparison with coarse-grained output from high-resolution models. It was felt that the Cellular Automaton introduced ad-hoc dependence on the model resolution and was harder to tune to a desired forcing spectrum. As with the CASB scheme, the new scheme was found to improve the ensemble skill by giving more realistic rates of growth of ensemble spread and reducing biases. It was also found that replacing the flow-dependent calculation of energy dissipation with a constant mean dissipation-rate somewhat reduced the improvements in forecast skill produced by the scheme. This supports the theoretical arguments of Palmer (2001) that flow-dependent non-local stochastic schemes are required to properly represent the effects of sub-grid variability.

Meanwhile, Bowler *et al.* (2009) developed a similar Stochastic Kinetic Energy Backscatter scheme (SKEB) for the UK Met Office's MOGREPS ensemble forecasting system. The random stream-function forcing pattern was generated using a different method, using a 3D pattern generator, with autocorrelation scales which were tuned to match those in data from a Cloud-

Resolving Model. Initially, the unphysical energy dissipation was estimated offline and only from the diffusion and advection terms, and the constant mean dissipation pattern applied as a tunable parameter in the scheme. As with the Shutts (2005) CASB scheme, the SKEB scheme was found to bring the slope of the energy spectrum of model winds significantly closer to that in observations, and increase the ensemble spread.

Tennant *et al.* (2011) describe an improved version of SKEB which uses a flow-dependent dissipation rate similar to that in the Shutts (2005) scheme. Testing the scheme in the MOGREPS forecast system, they find further improvements in the growth-rate of ensemble spread and the forecast frequency of blocking. They also tested the scheme for a case where the ensemble forecast had suffered “jumpiness”, i.e. the forecast distribution of future weather jumped back-and-forth between different weather regimes with successive forecast cycles. The revised SKEB scheme was found to slightly reduce this tendency.

1.6 EFFECTS OF STOCHASTIC FORCING ON DYNAMICAL SYSTEMS

As described in section 1.5 a number of stochastic parameterisations have been designed and implemented in GCMs, motivated by subtly different theoretical or practical considerations which turn out to be somewhat synonymous in practice. Whilst such schemes have often been reported to affect the behaviour of a host model, often little attention has been paid to the underlying mechanisms involved.

From a theoretical perspective, a stochastic forcing may influence the behaviour of a dynamical system in several qualitatively different ways. This can be illustrated using a very simple dynamical system. Figure 1.3 shows four different 1-D potential well shapes, along with the probability densities for the location in these wells of a particle set to roll freely with some fixed total energy, and for a particle which has some stochastic variation in its total energy.

This system has 2 dimensions; the particle’s position and velocity, which experience an acceleration equal to the gradient of the potential field. It is not a chaotic system, but crucially it does exhibit an internal mode of variability with which a stochastic forcing may interact; the periodic

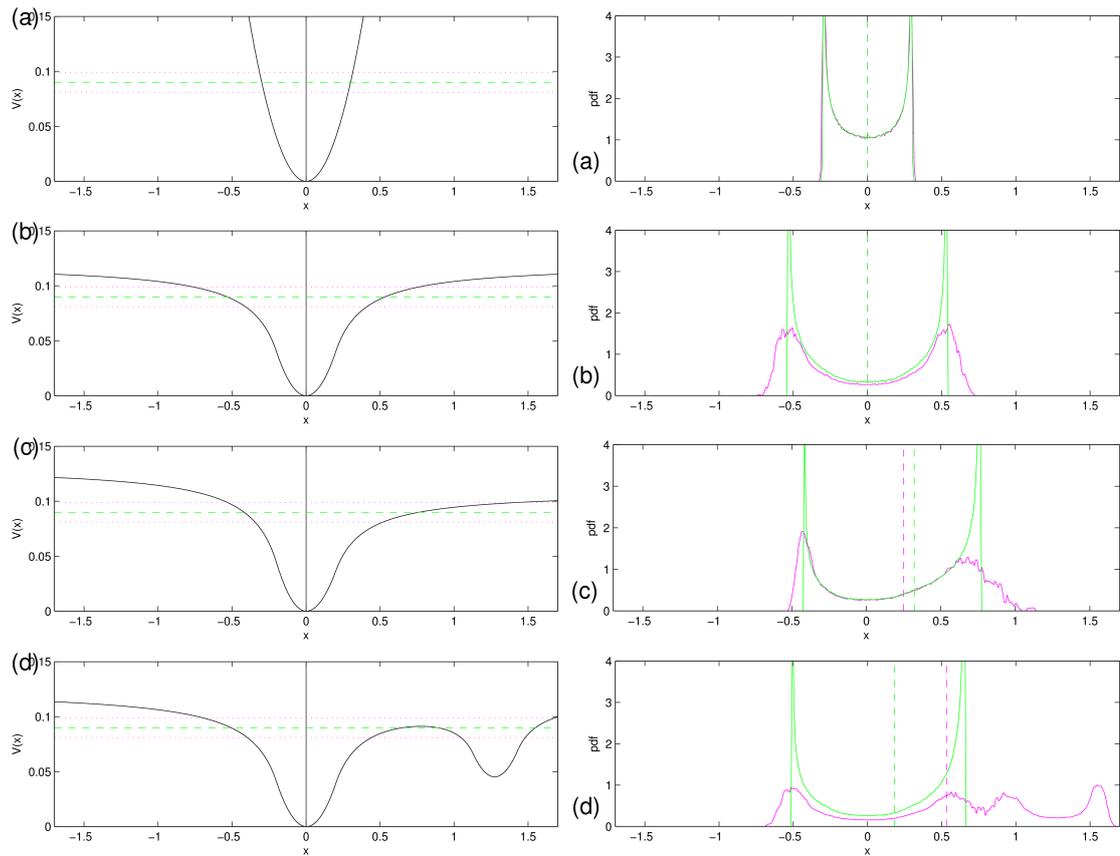


Figure 1.3: On the left; four potential well shapes; (a) Quadratic, (b) Gravitational, (c) Asymmetric gravitational, and (d) Double gravitational. On the right; the corresponding probability density functions for the position of a particle rolling freely in each well; (solid curves) the PDFs, and (dotted vertical lines) the means of the PDFs; for (green) the deterministic trajectory followed under the potential gradient force only, and (magenta) the trajectory when a stochastic forcing is also added. In the potential well plots on the left, the dotted green horizontal line shows the total energy for the deterministic trajectories, and the dotted magenta lines show the range of total energies explored by the stochastic trajectories.

oscillation of the particle in the well. The stochastic forcing here consists of perturbations to the particle's total energy, drawn randomly between limits. Autocorrelation is introduced simply by updating the random component once per period τ (where τ spans several time-steps, but a small fraction of the particle's period of oscillation) and having the stochastic energy term vary linearly between subsequent updates.

The four different potential well shapes represent four different forms for the system's dynamical feedbacks, and each represents a qualitatively different category of dynamical response

to stochastic forcing:

- *Type 0: linear response.* For the quadratic potential well (a), the gradient gives a linear restoring force, and the particle oscillates sinusoidally. The stochastic forcing has no qualitative effect on the system's internal mode or mean-state; the response is just a linear addition of the forcing which will only become significant if the amplitude of the forcing becomes comparable to the amplitude of the internal mode.
- *Type I: non-linear mode-amplification.* If the restoring force is non-linear (b), the system's internal mode may become more sensitive to the stochastic forcing, and its amplitude may be altered by a larger magnitude than that of the perturbations. The stochastic forcing makes new regions of phase-space accessible to the particle, increasing the range of extremes.
- *Type II: noise-induced drift.* If the restoring force is not only non-linear but asymmetric about its equilibrium state (c), the non-linear response of the system's internal mode to the stochastic forcing may also result in a change in the system's mean-state. Here, the stochastic forcing substantially reduces the residence time of the particle near the shallower extremity of its distribution, causing the mean-state to lie closer to equilibrium.
- *Type III: noise-induced transition.* Finally, if the system has multiple stable equilibria (d), the stochastic forcing may make it possible for it to shift to new internal modes or regimes altogether. Here, the stochastic forcing made it possible for the particle to overcome a potential barrier which was higher than its total energy in the deterministic case.

Even in this very simple system, the noise-induced response may be highly sensitive to details of the stochastic forcing. For example, figure 1.4 shows the noise-induced transition behaviour for the well with multiple equilibria (d), with three different update periods τ for the stochastic forcing. Interestingly, the longer the timescale of the stochastic forcing, the higher the frequency of transitions between the two regimes. This is because, in this case, a longer timescale increases the probability of the particle maintaining a positive excursion to its energy for long enough to cross over the potential barrier which separates the two regimes.

Whilst the atmosphere is clearly far more complex and has a vastly greater dimensionality than the simple dynamical system shown here, the same concepts can be extended. Replace the

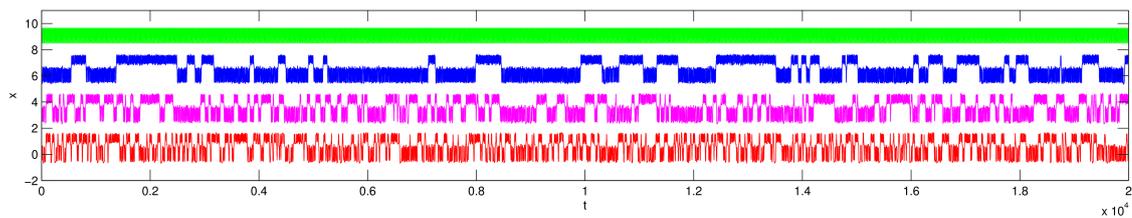


Figure 1.4: Long-term trajectories for the particle (position against time) for the double gravitational well, for (green) the deterministic system, and with the stochastic forcing with timescales of (blue) 0.5, (magenta) 1.0 and (red) 2.0.

particle position with the atmosphere's state in a multi-dimensional phase-space, and the well with the atmosphere's complex array of non-linear dynamical and physical feedbacks which act on that state, and adding a stochastic forcing to it should still yield one or more of the same basic kinds of response.

For example, the Random Parameters stochastic scheme tested by Bowler *et al.* (2008) was found to have no systematic effect on the model climate, suggesting that the atmosphere gave a relatively linear response (*type 0*) to stochastic parameter perturbations. The Lin & Neelin (2000) stochastic convection scheme gave a large increase in the slow modes of tropical variability, indicating a non-linear amplification (*type I*) of these internal modes in response to the stochastic forcing. In Berner *et al.* (2008), the Cellular Automaton Backscatter scheme was found to alter the mean rainfall pattern in the tropics, suggesting a noise-induced drift (*type II*) in the mean-state of the model. It was also found to alter the frequency of blocking in the mid-latitudes, indicating a noise-induced effect on the model's ability to transition between zonal and blocked flow regimes (*type III*). These changes in model behaviour were all found to make the models used more realistic, suggesting that the corresponding noise-induced effects play a role in determining the atmosphere's observed behaviour.

1.7 DISCUSSION AND AIMS OF THIS THESIS

Using theoretical considerations and experiments involving stochastic parameterisations, Various studies described in section 1.5 have found that the atmosphere's behaviour at the scales well-resolved by models is indeed sensitive to the variability at poorly-resolved and unresolved scales,

as hypothesised by some of the studies mentioned in section 1.4. They have also shown that this sensitivity covers more than just the amplitude of the small-scale variability; the atmosphere may behave differently depending on other properties, such as temporal and spatial autocorrelation of small-scale processes (Lin & Neelin 2000), and the flow-dependence of variability at small scales (Berner *et al.* 2009). However, the sensitivity of the atmosphere to such properties (and other related ones, such as the relationship between high-frequency fluctuations in different atmospheric variables, or how exactly the fluctuations are linked to small-scale processes) has not been investigated rigorously.

Rather, most studies have involved plugging one stochastic parameterisation (or a suite of stochastic parameterisations at the same time) in to a host model and reporting on the effects it has on the model, with a focus on potential improvements in forecast skill. Whilst some attempts have been made to tune stochastic parameterisations to match realistic high-frequency variability (e.g. by comparison with cloud-resolving model output), the design of structural configurations for the schemes are usually arrived at by a combination of guess-work (be it brilliantly intuitive!) and consideration for what will be most practical to code in to a given host model. As a result, many different stochastic methods have been developed, with a broad (and not entirely known) degree of overlap in terms of the variability and associated uncertainties they represent.

Indeed, modellers have had little *a priori* information about the detailed form of real high-frequency variability in the atmosphere, or even which properties of that variability a model atmosphere will be most sensitive to. Using Cloud-Resolving Models, substantial advances are being made in defining the properties of small-scale variability in the atmosphere (e.g. Shutts & Palmer 2007). But this information would be of greater use if the sensitivity of the resolved flows to such properties were better understood.

This study aims to shed light on some key un-answered questions concerning the sensitivity of the atmosphere to its high-frequency variability:

1. Does the atmosphere's response to high-frequency variability change drastically depending on the properties of the variability, the atmospheric state, or the particulars of the simulation? Or is there any robust, generic response common to any form of added noise? This question is highly relevant to the development of stochastic schemes. For example, if the

former is true it would justify developing stochastic schemes with a great deal of care and sophistication, but if the latter is the case, then very simple stochastic methods would suffice to represent at least a substantial component of high-frequency variability.

2. What is the relative importance of the inherent theoretical uncertainty associated with the unknown sub-grid state, compared to the pragmatically-based generic uncertainty associated with model error? This question is relevant to the representation of modelling uncertainties in ensemble prediction systems. If sub-grid state uncertainty is a major component of overall model uncertainty, then use of schemes which specifically represent physical sources of variability associated with sub-grid uncertainty (such as the stochastic convection scheme of Plant & Craig 2008, or the Stochastic Convective Vorticity scheme in MOGREPS, Bowler 2008) in ensemble forecasts is important.
3. Where high frequency variability in the atmosphere is found to influence its climatic behaviour, what are the key mechanisms involved? For example, is the atmosphere's response to high-frequency variability dominated by feedbacks from sub-grid and non-dynamical processes, or by a direct response from resolved-scale dynamics? And which types of system-response to stochastic forcing (out of those illustrated in section 1.6) tend to occur?

It is hoped that efforts to answer these questions will be of specific practical use to those in the weather and climate modelling community developing stochastic parameterisations for atmosphere GCMs, and also contribute to scientific understanding of multi-scale interactions in the real world's atmosphere. The experimental method followed is outlined in the following section 1.8.

1.8 OVERVIEW OF METHOD

To address the first question in the list of aims in the previous section 1.7, it will be useful to compare the response of atmospheric simulations to their high-frequency variability, for a variety of different model configurations, atmospheric conditions, and with varied properties of the high-frequency variability. It will then be informative to see what aspects of the response are similar across all the simulations, and which properties of the simulations lead to major differences.

Different atmospheric models contain a variety of different high-frequency variabilities (e.g. Horinouchi *et al.* 2003). However, they contain many other differences, so a difference between the resolved-scale behaviour of two models cannot easily be attributed to differences in their high-frequency variability responses. In order to investigate the sensitivity of these responses, it will be necessary to perform multiple experiments with the same model, with a variety of modifications designed specifically to alter the high-frequency variability. Fortunately, there is a wealth of literature on stochastic methods designed to do just that (see section 1.5). By comparing the behaviour of a common host model, using a variety of different stochastic parameterisations, the sensitivity to high-frequency variability can be explored.

Of course, the control simulation with no stochastic parameterisations is by no means devoid of high-frequency variability, so comparisons of stochastic runs to the control will not fully quantify the response. But the differences in the response when different stochastic parameterisations are compared can be explored as desired. These comparisons can be repeated for more than one model configuration and the responses can be compared for different atmospheric conditions within each simulation.

To investigate the importance of physically-based sub-grid uncertainty relative to generic modelling-uncertainties (the second question under the aims of the thesis), the comparison of different stochastic parameterisations described above will include some generic model-uncertainty schemes from the literature, along with the physically-based stochastic convection scheme of Plant & Craig (2008). It was felt that this represents what is likely to be a dominant source of sub-grid uncertainty, and is therefore a representative scheme to use. The stochastic methods so far developed to generically represent model-uncertainty (see section 1.5.1) are largely variations on two themes; stochastic multiplicative perturbations (e.g. Buizza *et al.* 1999), and time-varying stochastic perturbation of model parameters (e.g. Bowler *et al.* 2008). Therefore, these two schemes are also included in the comparisons in this study.

To investigate the mechanisms involved in the atmosphere's sensitivity to high-frequency variability (the third question), it would be useful to compare responses using one modelling framework which includes all the expected interactions within the atmosphere, and others which include only some of the processes interactively. It was felt that experiments in fully interactive atmospheric simulations run the risk of being difficult to interpret if the relative contributions from

resolved dynamical feedbacks and sub-grid / non-dynamical feedbacks are not known. Also, since this study is particularly interested in the sensitivity of the atmosphere, interactions with other components such as the ocean and land-surface will not be included.

Therefore, the methodology will be first to perform a comparison of different stochastic parameterisations in a Single-Column Model with prescribed dynamics, representing only the responses from parameterised processes. Results from this can then be compared to results from a comparison of stochastic parameterisations in a fully-interactive dynamical atmosphere simulation with prescribed surface properties. To keep the latter as simple, and interpretable, as possible, a highly idealised framework will be used in which orography, land-surface and seasonal variations are omitted. Such a framework, the “aqua-planet”, has been developed and has been used by the modelling community for some time. It simulates the atmosphere over a hypothetical simplified Earth, whose surface is covered entirely by ocean with a prescribed, fixed sea surface temperature pattern, and whose axis of rotation is not tilted relative to the sun.

There is a very large range of possible stochastic methods and model configurations which could potentially be included in these comparisons, and it is not immediately obvious how to narrow down this choice so that a practical number of model integrations are performed. Fortunately, Single-Column Model (SCM) integrations are very computationally cheap to run, so a much broader range of configurations can be tested in the SCM than is possible in the 3D aqua-planet framework. It is hoped that a large number of SCM tests of stochastic parameterisation schemes can be used to narrow down which schemes / configurations would be most useful to test in a smaller set of aqua-planet experiments.

The SCM experiments, and the stochastic (and deterministic) parameterisations used, are described in detail in chapter 2. Preliminary results assessing the SCM’s high frequency variability and its sensitivity to model physics and stochastic perturbation are also presented in that chapter. In chapter 3, results of a more rigorous comparison of stochastic parameterisations in the SCM are presented. Chapter 4 gives details and results for preliminary aqua-planet experiments including stochastic model-uncertainty schemes, employing horizontal resolutions and parameterisation schemes consistent with climate simulations. Overall conclusions are drawn in chapter 5, along with suggested avenues for further research in this field.

CHAPTER 2:

INVESTIGATING THE VARIABILITY OF PARAMETERISED PROCESSES IN A SINGLE-COLUMN MODEL

2.1 OVERVIEW AND CHOICE OF EXPERIMENT

As described in chapter 1, the over-arching method of the work described in this thesis involves first investigating the atmosphere's sensitivity to high-frequency variability in isolation from the large-scale dynamics. This will inform the study of the interactive, dynamical response in an aqua-planet framework, described in chapter 4. Various stochastic parameterisations are employed to explore sensitivities to the aspects of high-frequency variability they represent.

Single-Column Model (SCM) tests, in which the large-scale dynamical effects are prescribed according to observational estimates, are a useful starting point for assessing the high-frequency variability and general behaviour of model parameterisations. They are advantageous in that they allow parameterised processes and their interactions to be studied separately from large-scale dynamical responses, and are computationally very cheap to run. They have long been a standard tool in the development, testing and intercomparison of physical parameterisations for weather and climate GCMs (for example, Lord 1982, Grell *et al* 1991, Randall *et al.* 1996, Ghan *et al.* 1999, Xie *et al.* 2002). However, stochastic parameterisations have very seldom been tested in this way. Indeed, many stochastic parameterisations cannot be meaningfully applied in an SCM, because they act on the horizontal momentum equations which are not simulated (e.g. Frederiksen & Davies 1997), or they also have an explicitly non-local formulation (e.g. Shutts 2005). There are nonetheless many stochastic methods aimed at addressing the local thermodynamic variability of parameterised processes, for which the SCM framework seems a well-suited test-bed.

Model intercomparisons using the SCM framework have formed a major part of the Global Energy and Water cycle EXperiment (GEWEX) Cloud System Study (GCSS), which aims to support the development of improved physically based parameterisations for cloud processes. This collaboration has included a working group on Precipitating Convective Cloud Systems (PCCS, also known as Working Group 4) for many years, and an overview of their intercomparison methodology is given in Moncrieff *et al.* (1997). Work under the GCSS PCCS, and in this field generally, has focused a lot on the behaviour of deep moist convection over the tropical oceans. As discussed in chapter 1, the variability of this key climate process is often poorly simulated on a range of time-scales, largely due to deficiencies in the convection parameterisations (which, among other things, GCSS aims to improve). Extensive observational studies of this variability were made during the Tropical Ocean Global Atmosphere - Coupled Ocean-Atmosphere Response Experiment (TOGA-COARE) observation campaign in the tropical West Pacific, from November 1992 to February 1993. This is described by Webster & Lukas (1992).

One feature of rainfall variability over the tropical oceans which has attracted particular attention is the observed tendency for distinct transitions between periods in which convection is suppressed, giving dry conditions for many days, and periods in which it is active, giving widespread heavy rain. This behaviour is a key feature in many poorly simulated modes of tropical variability, such as the MJO, and is studied using TOGA-COARE data by Petch *et al.* (2007). They compare simulations using a Cloud Resolving Model (CRM), a global Numerical Weather Prediction (NWP) model, and an SCM. Their simulations cover 3 sub-periods each of the order 10 days selected from within the 4-month TOGA-COARE observation period, each containing a marked transition from suppressed to active convection. The modelling framework for these experiments, including the observation-derived prescribed dynamical forcings, have been designated by the GCSS as PCCS case 5. This case simulates an interesting and fundamental aspect of atmospheric variability associated with sub-grid processes, includes a representative range of the atmosphere's behaviour over the tropical oceans, and has a well-documented and readily available framework for running SCM experiments. It has therefore been selected for the present study as an ideal test-case in which to compare stochastic parameterisations and investigate the variability associated with parameterised processes.

In this chapter, the behaviour of an SCM simulation very similar to one of those used by

Petch *et al.* (2007) is assessed, with particular attention given to the high-frequency variability of the parameterised convection. A number of modified versions of the model are described, each incorporating a different stochastic parameterisation approach from the literature, and several alternative deterministic formulations are implemented for comparison. The variability in some of these model versions is briefly investigated. This is to assess the usefulness of the SCM for comparing stochastic parameterisations, and how best to compare them in order to gain insights about the underlying physics driving the variability. A more rigorous comparison of the stochastic and deterministic SCM variants, investigating differences in their variabilities and mean-states, is made in the next chapter.

Detail of the modelling framework and model variants used is given in section 2.2. Section 2.3 gives an overview of the meteorology of the case simulated by these experiments, investigating the key physical processes involved and the SCM's ability to capture them. Preliminary results examining the SCM's variability and sensitivity to initial conditions and stochastic perturbations are presented in section 2.4, and a summary and concluding discussion are given in section 2.5.

2.2 EXPERIMENTAL SETUP

The model runs simulate a column of the atmosphere in the tropical West Pacific warm pool region, at $2^{\circ}S$, $156^{\circ}E$, over the time period 9th - 28th January 1993 (this is the consecutive periods B and C of GCSS PCCS case 5; see Petch *et al.* 2007). A description of the model used is given in subsection 2.2.1 below. A number of alternative configurations of the SCM have been used, each with differing convection parameterisations or stochastic elements introduced. These are summarised in table 2.1 and described in subsection 2.2.2, with detail of the different parameterisations used. In order to quantify the variability of the SCM, investigate sensitivity to initial conditions and gain statistically robust results, ensembles of SCM runs have been performed for all the model variants. Subsection 2.2.3 describes the ensemble framework, with consideration of ensemble statistics given in subsection 2.2.4. To run an SCM simulation of a real case-study as in these experiments, the effects of the resolved-scale dynamics (which are not simulated) must be prescribed according to observation-derived estimates. The large-scale forcing data-set used in the present study is the same as that used in the SCM runs of Petch *et al.* (2007), and is described

in subsection 2.2.5.

2.2.1 MODEL DESCRIPTION

Experiments have been carried out using the UK Met Office Unified Model (UM, Cullen 1993) in its Single-Column configuration. The default model configuration used is very similar to that used in the SCM experiments of Petch *et al.* (2007), including the same set of parameterisations (essentially those used in the atmosphere component of the Hadley Centre Global Environment Model (HadGEM1) and described by Martin *et al.* 2006). The same vertical resolution is used, with 38 model levels. Further detail of the parameterisations is given in subsection 2.2.2 under the description of the default UM configuration.

There are however a few differences between the SCM described here and that used by Petch *et al.* (2007). Firstly, a time-step of 30 minutes is used for all model routines except the radiation scheme, rather than the 20 minute time-step with 10 minute sub-step for convection used in Petch *et al.* (2007). The longer time-step used in this study makes it more consistent with a climate simulation, whereas Petch *et al.* (2007) were comparing the SCM to an NWP simulation. The radiation scheme uses the same time-step of 3 hours.

Also, a minor improvement was made to the SCM for the present study. Since the model contains no dynamical adjustment component, the height and pressure of each model level were held constant in the original SCM such that they were consistent with the initial temperature profile. However, as the temperature profile is then allowed to evolve according to the model physics routines, the model heights, pressures and temperatures would become thermodynamically inconsistent. A scheme was therefore added to modify the heights of the model levels each time-step so as to maintain hydrostatic balance, given fixed pressure at each level including the surface. Tests confirm that this modification had a noticeable but very small impact on the simulation (not shown).

2.2.2 MODEL VARIANTS

Table 2.1 gives an overview of the different SCM variant runs performed in this study. The motivation and context for the stochastic parameterisations included were described in the literature review on stochastic parameterisations in chapter 1; details of their implementation in the UM SCM are given in the following subsections. Two generic model uncertainty schemes from the literature (Random Parameters and Multiplicative Noise) are compared, along with a scheme designed to physically represent variability associated with sub-grid uncertainty in convection (the Plant & Craig 2008 stochastic convection scheme). A deterministic perturbed parameter ensemble (based on the Random Parameter scheme) is also tested for comparison. Structural uncertainties in convection parameterisations are explored by comparing three different deterministic convection schemes in the SCM. The effects of smoothing the convective variability are investigated using a simple time-smoothing scheme designed for this study. All of the SCM configurations used are fully described below.

2.2.2.1 DEFAULT UM

The default Unified Model SCM has the prognostic variables temperature (T), zonal, meridional and vertical wind components (u, v and w), specific humidity (q), cloud liquid water content (qcl) and ice water content (qcf). Fluxes of heat and moisture in and out of the model due to radiative transfer, surface fluxes and precipitation are all calculated explicitly by parameterisations and are not constrained to observations. An overview of the parameterisation set is given by Martin *et al.* 2006, and includes:

- A boundary-layer turbulent mixing and surface flux scheme, which includes a Richardson number based local mixing component and a non-local buoyancy-driven component.
- A penetrative mass-flux convection scheme, based on an entraining-detraining bulk plume model (Gregory & Rowntree 1990), closed on dilute CAPE for deep convection, and on surface buoyancy flux for shallow convection.
- A large-scale cloud scheme, in which cloud begins to form at grid-mean relative humidities above a threshold somewhat below 100%, due to sub-grid inhomogeneity.

Table 2.1: Summary of the model variants used.

Name	Abbv.	Description	Type
Default UM	DefUM	UM SCM as in Petch <i>et al.</i> (2007), physics as in Martin <i>et al.</i> (2006)	Deterministic
Random Parameters (Varying)	RP	As DefUM but with randomly varying model parameters, following Arribas (2004)	Stochastic
Random Parameters (constant)	RPconst	As RP, but with the randomly selected parameter values held constant	Deterministic
Multiplicative Noise	MN	As DefUM but with randomly scaled physics tendencies (Buizza <i>et al.</i> 1999).	Stochastic
Kain-Fritsch	KF	As DefUM but with Kain & Fritsch (1990) convection scheme instead of the default	Deterministic
Plant & Craig (stochastic)	PC	As DefUM but with Plant & Craig (2008) stochastic convection scheme (uses KF plume)	Stochastic
Plant & Craig (deterministic)	PCdet	As PC but set to simulate deterministic limit of well-sampled cumulus ensemble	Deterministic
Time-Smoothed Convection	TSC	As DefUM but with time-smoothing applied to the convective tendencies.	Deterministic

- A cloud microphysics scheme, which parameterises the fall-speed for ice particles and conversions of water between ice, cloud-liquid-water, and rain species.
- A fully interactive 2-stream radiative transfer scheme with several spectral bands in the Shortwave and Longwave, accounting for the radiative effects of various gases, water-vapour and cloud.

All of these parameterisation schemes are described, with appropriate references, in appendix A.

2.2.2.2 RANDOM PARAMETERS

Here one of the stochastic schemes employed in the Met Office Global and Regional Ensemble Prediction System (MOGREPS, described by Bowler *et al.* 2008) is implemented in the SCM. This is the Random Parameters (RP) scheme described therein and in section 2.2 of Arribas (2004). It is designed to represent parameterisation uncertainty in an ensemble forecast by randomly sampling various free parameters in the model, in a time-varying manner within each ensemble member. The motivation and context for the scheme were described in detail in the literature review on stochastic parameterisations in chapter 1.

The amplitude of the stochastic perturbations introduced by the scheme is set by the bounds of the allowed range of values for each time-varying model parameter. These bounds must be chosen somewhat subjectively and with consideration for the numerical stability of the parameterisations in which they reside, but have been set after consulting experts on each of the relevant parameterised processes (Arribas 2004). Some of the parameter bounds used in the RP scheme in this study differ from those given in Arribas (2004), as the values used operationally in MOGREPS were revised since and it was felt that the most up-to-date values should be used here. Details of the model parameters perturbed by the scheme are given below, and refer to elements of the UM parameterisation schemes mentioned in subsection 2.2.2.1. See appendix A for descriptions of each parameterisation detailing the elements controlled by these parameters. The parameter bounds chosen are given in table 2.2.

- *Neutral mixing length parameter, par_mezcla* : This parameter is a coefficient in the formula to estimate the neutral mixing length scale, which is in-turn a coefficient in the calculation of vertical turbulent fluxes in the boundary layer scheme.
- *Stability function parameter, G_0* : A coefficient in the empirical stability functions used to relate boundary layer turbulent fluxes to the Richardson number.
- *Entrainment rate coefficient, $entcoef$* : A single parameter which scales the rate of entrainment of air from the environment profile in to convective plumes as they ascend. This applies to the deep, shallow and mid-level convection schemes.
- *CAPE timescale, τ_{CAPE}* : Sets the timescale over which convection dissipates Convective

Available Potential Energy. It is used to scale the convective mass-flux and therefore scales the overall convective response for a given amount of instability. Only applies to the deep and mid-level convection schemes.

- *Critical Relative Humidity, RH_{crit}* : The threshold Relative Humidity for cloud formation; sets the width of the distribution of saturation used in the large-scale cloud scheme.
- *Cloud-to-rain conversion threshold over land, CW_{land}* : The threshold cloud liquid water content above which auto-conversion to rain begins according to the “warm rain” parameterisation in the microphysics scheme. This value only applies over land grid-points.
- *Cloud-to-rain conversion threshold over sea, CW_{sea}* : As CW_{land} , but a lower value which applies over sea grid-points, to account for the fact that rain forms more readily over the sea where aerosol concentrations are lower.
- *Ice fall-speed coefficient, CI* : Coefficient in the formula relating ice particle fall-speed to ice particle size.

The RP scheme of Arribas (2004) also perturbs two parameters used in the Gravity Wave Drag scheme, but this is not applicable in the SCM framework. Also, since the case simulated in this study is over open sea, the CW_{land} parameter has no effect.

Vertical correlation of the parameter values is ensured by using the same set of random parameters at each model level at a given time-step (except for RH_{crit} , which varies with height by default, but the same stochastic term is added to it at each model level). Temporal correlation of the random parameters is set using a first-order auto-regression model:

$$P_{n+1} = \mu + r(P_n - \mu) + k\varepsilon \quad (2.1)$$

where P_n is the value of a parameter P at one time-step, P_{n+1} is its value at the next time-step, μ is the mean value of P , r is the auto-correlation of P , and $k\varepsilon$ is a stochastic shock term (see section 3.2 of Mylne et al 2005). The value of ε is randomly drawn from a flat distribution between -1 and 1 , with k setting the maximum amplitude of the shock term. In the configuration used

Table 2.2: RP scheme parameter values.

Parameter name	Parameterisation scheme which uses it	Standard value (in default UM)	Bounds in RP scheme (min/max)	Units
<i>par_mezcla</i>	Boundary layer	0.15	0.05 / 0.5	
G_0	Boundary layer	10.0	5.0 / 20.0	
<i>entcoef</i>	Convection	3.0	2.0 / 4.0	
τ_{CAPE}	Convection	1800	1800 / 3600	seconds
RH_{crit}	Large-scale cloud	80.0	77.5 / 82.5	%
CW_{land}	Microphysics	8.0E-4	1.0E-4 / 1.0E-3	kg kg ⁻¹
CW_{sea}	Microphysics	2.0E-4	5.0E-5 / 5.0E-4	kg kg ⁻¹
CI	Microphysics	25.2	17.0 / 33.0	

in MOGREPS, and implemented here, the temporal correlation is forced to act on longer time-scales by simply applying the auto-regression formula (2.1) to recalculate the parameter values once every 3 hours, rather than every time-step. Each parameter value is constrained to within its allowed bounds P_{min} , P_{max} by applying a simple check to restore it to P_{min} (P_{max}) if it goes below P_{min} (above P_{max}). The maximum amplitude of the shock term is set to $k = 1/3(P_{max} - P_{min})$. Strong autocorrelation is applied, with $r = 0.95$. Figure 2.1 shows an example time-series of a model parameter generated by the RP scheme, and figure 2.2 shows the resulting probability density for the same parameter. The scheme produces a relatively flat distribution across most of the parameter's range, but the thresholding applied gives the scheme a tendency to draw the max and min values particularly often. Each parameter sits on its allowed bound values about 20% of the time, preferring the boundary closer to its standard value.

Correlations between different parameters are forced by using a single value of ϵ for all perturbed parameters at a given time. To remove any bias towards the standard parameter values near the start of the SCM runs, the RP scheme is called to recalculate the parameter values 30 times before the first time-step.

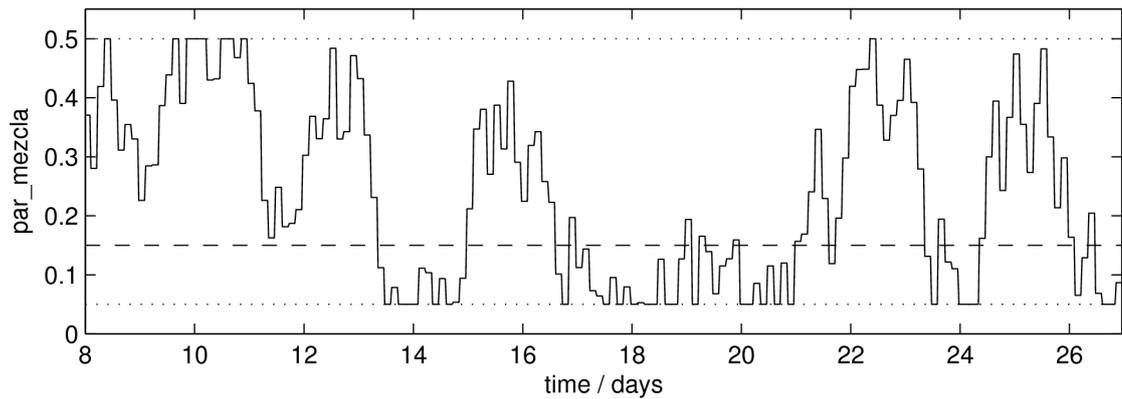


Figure 2.1: Time-series of the model parameter `par_mezcla` generated by the RP scheme in the SCM (solid), standard value of `par_mezcla` (dashed) and RP scheme max/min values (dotted).

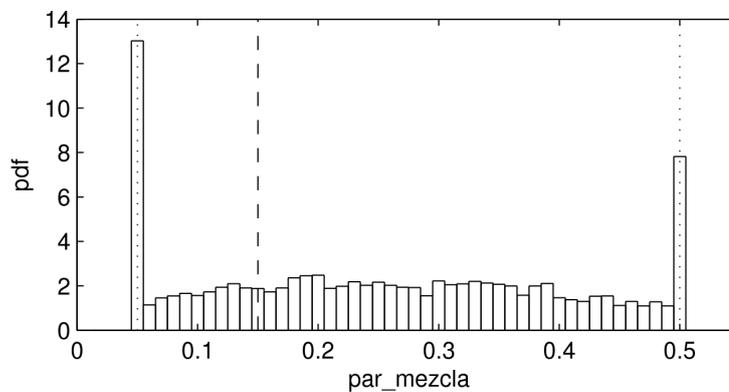


Figure 2.2: Probability Density Function for the model parameter `par_mezcla` generated by the RP scheme. Dashed and dotted lines are as in figure 2.1. The PDF was generated from an ensemble of 40 independently perturbed 19-day runs.

2.2.2.3 CONSTANT RANDOM PARAMETERS

Another approach to sampling model parameterisation uncertainty is to run an ensemble of model runs, each with a different but constant-in-time set of parameter values. This approach was discussed in chapter 1, along with examples from the literature. Here, such a “perturbed parameter” ensemble of SCM runs has been generated by running the RP scheme described above to determine initial parameter values, which are then held constant for the whole run. Different random values are drawn independently for each ensemble member. As in the time-varying RP runs, the RP scheme is repeated 30 times before the first time-step, to remove any autocorrelation with the initial standard parameter values. The ensemble therefore samples the same distribution for

the parameters as the time-varying RP scheme described above, but removes the stochastic component, making an interesting comparison. Note that this approach does not explore the full possible parameter space, as it is constrained by the correlations between parameters forced by the RP scheme.

2.2.2.4 MULTIPLICATIVE NOISE

A stochastic multiplicative perturbation scheme following Buizza *et al.* (1999) has been implemented in the UM SCM. An updated version of this scheme is used operationally in the ECMWF Integrated Forecasting System ensemble. The Multiplicative Noise (hereafter MN) scheme simply scales the model's total parameterised tendencies by a factor which varies randomly but symmetrically about unity. It is designed to account for the uncertainty associated with the parameterised processes in an ensemble prediction system (see the literature review on stochastic parameterisations in chapter 1 for further discussion of its theoretical basis and use in forecasting). In this implementation, stochastic perturbations to temperature T , specific humidity q , and winds u and v are added at the end of each time-step, following the formulae

$$\Delta T_{stoch} = \beta \epsilon (\Delta T_{BL} + \Delta T_{conv} + \Delta T_{LScloud} + \Delta T_{micro} + \Delta T_{rad}) \quad (2.2)$$

$$\Delta q_{stoch} = \beta \epsilon (\Delta q_{BL} + \Delta q_{conv} + \Delta q_{LScloud} + \Delta q_{micro}) \quad (2.3)$$

$$\Delta u_{stoch} = \beta \epsilon (\Delta u_{BL} + \Delta u_{conv}) \quad (2.4)$$

$$\Delta v_{stoch} = \beta \epsilon (\Delta v_{BL} + \Delta v_{conv}) \quad (2.5)$$

The subscript *stoch* denotes the stochastic perturbations, whilst other subscripts denote the unperturbed increments from each of the default UM parameterisations listed in subsection 2.2.2.1; *BL* for the boundary layer scheme, *conv* for the convection scheme, *LScloud* for the large-scale

cloud scheme, *micro* for the microphysics scheme and *rad* for the radiation scheme. ϵ is a random number drawn from a flat distribution between -1 and 1, and β is a parameter which controls the amplitude of the stochastic perturbations. There is also a check to restore q to zero if the stochastic perturbation implies a negative value.

The random number ϵ is not redrawn for the perturbations to different model-levels or variables, forcing the stochastic perturbations to replicate the correlation structure of the original parameterised tendencies. Temporal correlation is forced by holding ϵ constant over multiple time-steps, the correlation timescale τ_C of the scheme being set by the number of time-steps. The scheme of Buizza *et al.* (1999) also forces spatial correlation by holding ϵ constant over multiple neighbouring grid-points to span a correlation length-scale L_C , but clearly this is not applicable to the SCM framework. They found that the scheme yielded the greatest improvement in the performance of an Ensemble Prediction System with $\beta = 0.5$, $\tau_C = 6$ hours, and $L_C = 10^\circ$. The same values for β and τ_C have been used here.

2.2.2.5 KAIN-FRITSCH CONVECTION SCHEME

In this model variant the Default UM's convection scheme has been entirely replaced by a different commonly used deterministic penetrative mass-flux convection scheme; that of Kain & Fritsch (1990, hereafter KF), with subsequent modifications. The version implemented here is the one described by Kain (2004). This convection scheme is in many ways similar to the UM's standard convection scheme (described in appendix A), but there are nonetheless some structural differences:

- The KF scheme's convective trigger function is crucially dependent on the grid-point vertical velocity w , whilst the Default UM convection scheme only uses w to determine whether to call the deep or shallow version of the scheme.
- The KF and default UM convective plume models have somewhat different formulations for the exchange of heat, moisture and momentum between the plume and its environment.
- The KF scheme explicitly models the plume vertical velocity, terminating the plume when this falls to zero, whilst the Default UM scheme simply terminates the plume when the

ascent parcel ceases to be buoyant. Therefore, the KF scheme may produce convective overshoots, whilst the default UM scheme does not.

- The KF scheme diagnoses whether to apply a CAPE-based closure (for deep convection) or a closure based on boundary layer turbulent kinetic energy (for shallow convection) depending on the depth of the convection, *after* the plume ascent has been performed. The default UM scheme applies similar closures, but diagnoses which one to use *a priori*. Note that, whilst the Kain (2004) scheme parameterises the timescale for dissipation of CAPE according to a cloud lifetime which is a function of the grid-size and the horizontal wind-speed, in this implementation the timescale has been set to a constant value similar to that in the default UM scheme. This is because the grid-size is not defined in the SCM.

The KF convection scheme was designed for use in mesoscale NWP models rather than climate GCMs, so it may seem an unusual choice to compare with other schemes in this study. Its plume model design is more consistent with a single convective cloud (as is plausible in a mesoscale model grid-box), rather than a bulk plume representing an ensemble of many clouds (as is more likely in a climate GCM grid-box, and as assumed in the default UM convection scheme). However, this property made the KF plume model an ideal choice for use in the Plant & Craig (2008) stochastic convection scheme, which simulates multiple individual plumes simultaneously. This scheme has also been used in this study and is described in subsection 2.2.2.6 below. The deterministic KF scheme has therefore been implemented here for comparison with that stochastic scheme. Also, the UM convection scheme and the KF scheme are both commonly in use in NWP models, so a comparison of the two is instructive regarding structural uncertainty in convection parameterisations. Such a comparison is made in the UM in a forecasting context by Done (2002).

2.2.2.6 PLANT AND CRAIG STOCHASTIC CONVECTION SCHEME

Here the stochastic convection scheme of Plant & Craig (2008, hereafter PC) has been implemented in the SCM. This scheme is designed to account for departures from statistical equilibrium in the ensemble of convective plumes within each model grid cell, given the likely small size of this sub-grid ensemble of clouds. The context and motivation for this scheme was described in

the literature review on stochastic parameterisations in chapter 1.

At statistical equilibrium convective cloud sizes are assumed to follow the theoretical distribution of Craig & Cohen (2006). A population of clouds in each grid box is generated by randomly sampling this distribution. Clouds with different sizes behave differently because those with larger radii should have a lower rate of entrainment of air from their environment per unit cloud mass-flux than those with smaller radii.

Each cloud in the population is independently simulated using an adaptation of the KF plume model. In this model, the mass of air from the environment δm_e made available for entrainment in to the plume as it rises over a pressure interval δp is parameterised as $\delta m_e = m_0(-0.03\delta p/r)$, where m_0 and r are the mass-flux and radius respectively of the plume at cloud-base. In the KF scheme, m_0 is obtained from the grid-scale closure calculation and r is set to a fixed value. However, in the PC scheme m_0 is drawn from the equilibrium distribution (2.6), r is calculated assuming that m_0 is proportional to r^2 , and entrainment rates are calculated accordingly for each cloud. Following Craig & Cohen (2006), the probability distribution function (PDF) for an individual cloud's cloud-base mass flux m_0 is:

$$p(m_0)dm_0 = \frac{1}{\langle m_0 \rangle} \exp\left(\frac{-m_0}{\langle m_0 \rangle}\right) dm_0 \quad (2.6)$$

where $\langle m_0 \rangle$ is the equilibrium ensemble-mean cloud-base mass-flux per cloud and is assumed to have a constant value of $2.0E7 \text{ kgs}^{-1}$. In each grid-box, the PC scheme generates clouds randomly in time, and the likely magnitude of convective activity is set by scaling the overall probability per unit time of generating clouds. This is done by drawing clouds from the distribution 2.6 after normalisation by the equilibrium number of clouds within the grid-box, given by

$$\langle N \rangle = (\Delta x)^2 \frac{\langle \overline{M_0} \rangle}{\langle m_0 \rangle} \quad (2.7)$$

where $(\Delta x)^2$ is the grid-box area and $\langle \overline{M_0} \rangle$ is the equilibrium total cloud-base mass-flux per unit area, which is determined using a CAPE closure method. To obtain physically based convective variability, it is preferable for the closure quantity $\langle \overline{M_0} \rangle$ to vary smoothly according to

the resolved state of the atmosphere rather than responding to variability on the poorly-resolved scale of an individual grid-box / time-step. The CAPE closure calculation is therefore done using temperature and moisture profiles averaged over nearby grid-points and recent time-steps. The spatial averaging cannot be done in the SCM framework, but time-averaging is used to calculate the CAPE from which $\langle \overline{M_0} \rangle$ is derived. Without spatial averaging, a longer time-averaging period is needed, and a compromise had to be made between providing a smoothly varying closure quantity and capturing variations in the large-scale forcings. Initial tests suggested the scheme behaved most physically when averaging over 10 hours (20 time-steps), and this value has been used.

As is evident in equation 2.7, a larger (smaller) grid-size will give more (fewer) plumes on average, which will give a more complete (limited) sampling of the equilibrium distribution and hence smaller (larger) fluctuations in grid-box mean convection. To investigate this property, the scheme has been run with two different grid-sizes Δx of 50km and 100km. The SCM has no defined grid-size, but this can be set as a parameter within the PC scheme.

2.2.2.7 DETERMINISTIC LIMIT OF THE PLANT AND CRAIG SCHEME

In the limit of large grid-size Δx , the PC stochastic convection scheme will fully sample the equilibrium distribution of the ensemble of possible convective clouds, yielding deterministic behaviour. The scheme can be set to closely approximate this limit by operating as a spectral convective parameterisation. That is, the plume model is run for a large set of clouds which evenly samples the distribution of cloud-base mass-fluxes, and the tendencies produced by the clouds are weighted according to their occurrence probabilities and added up. This is similar to the approach originally proposed by Arakawa & Schubert (1974), which is used in other spectral convective parameterisations designed since (e.g. Hack *et al.* 1984). The PC scheme has been run in this deterministic mode so that the effect of the stochastic component of the scheme on the SCM can be assessed by comparison. Comparing this run to the UM implementation of the KF convection scheme described earlier will also be informative about the structural uncertainty associated with using a bulk versus spectral convection scheme.

2.2.2.8 TIME-SMOOTHED CONVECTION

It should be noted that deterministic parameterisations can exhibit considerable variability, even under constant forcings (see for example Plant & Craig 2008, figure 4, panel (d), in which the KF scheme varies considerably in a radiative-convective equilibrium SCM test). As shown by Scinocca & McFarlane (2004), it is possible to enhance this existing variability in GCM parameterisations without recourse to stochastic methods, by altering certain properties of the deterministic schemes. They show that forcing the convection parameterisation to respond on longer timescales, either by increasing the CAPE closure timescale or employing a prognostic closure, yields improvements in the realism of convective variability. Here a simple method to increase the timescale of convective variability has been implemented in the UM SCM, and will be referred to as the time-smoothed convection (TSC) scheme.

Each time-step, the tendencies from the Default UM convection scheme are not used to increment the model fields directly, but are added to separate convective heating and moistening fields ΔT_{tsc} and Δq_{tsc} which are then added to the model fields smoothly over future time-steps. This process follows the form

$$\frac{d\Delta T_{tsc}}{dt} = \left. \frac{dT}{dt} \right|_{conv} - \frac{\Delta T_{tsc}}{\tau_{tsc}}, \quad \left. \frac{dT}{dt} \right|_{convtsc} = \frac{\Delta T_{tsc}}{\tau_{tsc}} \quad (2.8)$$

$$\frac{d\Delta q_{tsc}}{dt} = \left. \frac{dq}{dt} \right|_{conv} - \frac{\Delta q_{tsc}}{\tau_{tsc}}, \quad \left. \frac{dq}{dt} \right|_{convtsc} = \frac{\Delta q_{tsc}}{\tau_{tsc}} \quad (2.9)$$

where the subscript *conv* denotes the tendencies from the Default UM convection scheme, and the subscript *convtsc* denotes the time-smoothed convection tendencies which are applied to the model. The amount of smoothing applied is controlled by the timescale τ_{tsc} , which has been set to 2 hours. This scheme has the advantage that, unlike altering the closure calculation, it should give the same mean convective response as the unsmoothed default convection scheme. So any changes seen when this scheme is employed should be the result of the change in variability, rather than an adjustment of the equilibrium state to maintain the convective heating rate under a different closure.

2.2.3 ENSEMBLE FRAMEWORK

Although SCMs have only one dimension and typically contain fewer degrees of freedom than 3D atmospheric models by several orders of magnitude, their solutions can exhibit very strong non-linearities. This is evident in Hack & Pedretti (2000), who demonstrated that an SCM version of the NCAR CCM3 model displayed high sensitivity to small initial condition perturbations. Ensembles of SCM runs each given a different random initial perturbation (of less than 1K in temperature) diverged to give ensemble spreads of several K at later times. They conclude that such large solution uncertainties render comparisons of single runs of different SCM formulations difficult to interpret; results drawn from such a comparison are of little relevance if they are not reliably reproduced under slight changes to the initial conditions. They suggest that to assess the effects of a change in model formulation, ensembles of runs with initial condition perturbations should be compared to check whether any statistically and practically significant changes in the model's behaviour occur. Such an approach is taken in the present study to compare different deterministic and stochastic parameterisations in the UM SCM.

The UM SCM's sensitivity to different initial condition perturbations will be investigated in section 2.4.1. After this, results are for ensembles containing a standard set of initial condition perturbations to temperature, drawn randomly from a uniform distribution between $\pm 0.5\text{K}$ on the lowest model level, and decreasing exponentially above with a decay height-scale of 1km. A corresponding perturbation is also applied to specific humidity so as to conserve relative humidity. Note that these perturbations are not aiming to quantitatively sample the uncertainty in the initial conditions used in the model, but are chosen so as to be large enough to force the ensemble to diverge into a set of effectively independent realisations during the run, but small enough not to give any of the ensemble members a lasting thermodynamic bias relative to the ensemble mean. Hume & Jakob (2005) used ensemble SCM simulations of the Tropical West Pacific to quantify parameterisation response to uncertainty in the initial conditions and large-scale dynamical forcings, and detect biases in the parameterisations relative to observations. These are not aims of the present study, but similar methods will be used to quantify the variability of various parameterisations and detect systematic differences between the different model variants.

When a stochastic parameterisation is employed, ensembles can be generated without initial

condition perturbations, by drawing the random numbers used in the stochastic component of the model independently for each ensemble member. This method is used to compare the parameterisation uncertainty to initial condition sensitivity in section 2.4.3. However, for completeness and fair comparison, all subsequent results for stochastic SCM variants include the initial condition (IC) perturbations described above as well as independently drawn stochastic parameterisation (SP) perturbations in each ensemble member.

2.2.4 ENSEMBLE STATISTICS

It will be useful to compare ensemble statistics such as the ensemble mean and ensemble spread (which will be quantified by the standard deviation of ensemble members) between different runs, to investigate any differences between them. In a finite ensemble both of these statistics will have sampling distributions of finite width, so a consideration of the uncertainties in these quantities is needed. The ensemble mean μ of some variable follows a sampling distribution which is approximately normal with a standard error of σ/\sqrt{N} , where σ is the standard deviation of the variable at that time and N is the number of ensemble members. It can be shown that for ensemble sizes N greater than about 25, the sampling distribution of σ is also approximately normal, with a standard error of $\sigma/\sqrt{2N}$. Most of the ensembles used consist of a control run with no initial condition perturbations, and 39 initially perturbed runs. In some cases a larger ensemble size of 100 was used to check the robustness of results. The standard errors and 95% confidence intervals for μ and σ for each of these ensemble sizes are given in table 2.3.

Hack & Pedretti (2000) used a much larger ensemble size of 500, which is preferable for robustly detecting more subtle effects. Given the low computational cost of SCM runs, this would

Table 2.3: Errors in ensemble statistics, expressed as percentage fractions of the ensemble spread σ .

Statistic	Ensemble size	Standard error	95% confidence interval
Ensemble mean μ	40	$\pm 16\%$	$\pm 32\%$
	100	$\pm 10\%$	$\pm 20\%$
Ensemble spread σ	40	$\pm 11\%$	$\pm 22\%$
	100	$\pm 7\%$	$\pm 14\%$

have been feasible here. However, smaller ensemble sizes are used in this study, in line with the typical size of operational ensemble systems. As will be seen in the results in section 2.4, the ensembles used are sufficient to detect any major differences which arise between different SCM configurations, with only minor, subtle changes being hard to robustly detect.

The confidence intervals for ensemble statistics can be narrowed by averaging over multiple model levels or time-steps, as is done for many of the results. However, the degree of robustness gained by averaging will depend critically upon the vertical and temporal autocorrelation of the ensembles used. In some of the results presented later, 95% confidence intervals have been estimated for vertical or temporal means of statistics by investigating the spread of the sampling distribution of the relevant averaged statistic when calculated from sub-samples of the ensemble. The method is described below.

For a statistic S_N calculated from an ensemble of size $N = n \times m$, a sample of m values of the statistic S_n can be calculated from m independent sub-samplings, each of size n . The values of S_n will have some scatter about S_N , from which the width of the sampling distribution of S_n can be calculated. The sampling distributions of the statistics in question are known to have widths proportional to $1/\sqrt{\nu}$ (for ν degrees of freedom), and the whole ensemble presumably has m times as many degrees of freedom as each sub-sample. So the standard error of the statistic for the whole ensemble σ_{S_N} can be estimated as $\sigma_{S_N} = \sigma_{S_n}/\sqrt{m}$, where σ_{S_n} is the standard deviation of the sub-sample values S_n . For statistics averaged over multiple time-steps or vertical levels, the Central Limit Theorem dictates that the sampling distributions should be approximately normal even if the point-wise values of the statistics are not. Therefore, normality has been assumed when estimating 95% confidence intervals from the standard errors σ_{S_N} . The confidence intervals shown in the results have been estimated using $m = 10$, $n = 4$. Sensitivity tests found they changed very little if different values of m and n were used.

2.2.5 FORCING DATA

A forcing data-set derived from TOGA-COARE observation campaign measurements is used to replace the dynamical component of the model in the SCM. The forcings consist of prescribed time-series for tendencies in temperature and specific humidity due to large-scale horizontal and

vertical advection (figure 2.3), and observed wind profiles (figure 2.4). The temperature and humidity tendencies are simply added to the model fields as increments each time-step, whilst the model winds are relaxed towards the observed values with a relaxation timescale of 1 hour. Large-scale advection of cloud condensate variables is assumed to be zero. The forcings were derived from observations over the TOGA-COARE Intensive Flux Array (IFA, see the map figure 14 of Webster & Lukas 1992). The method used to derive the large-scale dynamical tendencies from surface and radiosonde measurements at numerous sites across the IFA is described in Ciesielski *et al.* (2003), and involved estimating average heat and moisture budgets for the IFA. Whole IFA-means of the observations are used for the prescribed winds in the SCM, and also for the initial conditions. Sea Surface Temperatures (SSTs) are also prescribed according to an observed time-series from TOGA-COARE and are shown in figure 2.5. All of these data, and further information about their derivation, are freely available online from the TOGA-COARE SCM website (<http://kiwi.atmos.colostate.edu/scm/toga-coare.html>). The meteorology evident in the plots of the forcing data is described in section 2.3.

Emanuel & Zivkovic-Rothman (1999) found that corrections had to be made to the moist enthalpy budget of TOGA-COARE-derived forcing data to prevent errors in the data from causing large model drift in an SCM test of their convection scheme (their section 3). Wu & Moncrieff (2000) show that these forcing errors originate in the estimated large-scale advection tendencies (rather than the observed turbulent or radiative fluxes). Corrections to the moisture budget calculation to account for known biases in radiosonde humidity measurements have since been made by Ciesielski *et al.* (2003), but these have relatively little effect on the budget-derived rainfall during the period 9th-28th January 1993 simulated in the present study (see their figure 11). Petch *et al.* (2007) suggest that at times during TOGA-COARE there were insufficient observations to construct an accurate moisture budget, leading to large errors in both the budget-derived rainfall rate and the large-scale dynamical tendencies used to force SCM and CRM simulations. They conclude that many of the biases common to both SCM and CRM simulations of the case studied result from such errors in the large-scale forcings. This is in agreement with earlier model intercomparison studies using TOGA-COARE forcing data, such as Krueger & Lazarus (1999).

The errors in the forcing data discussed above have been frustrating to studies which attempt to validate CRMs or parameterisation sets in SCMs against observations. However, they do not

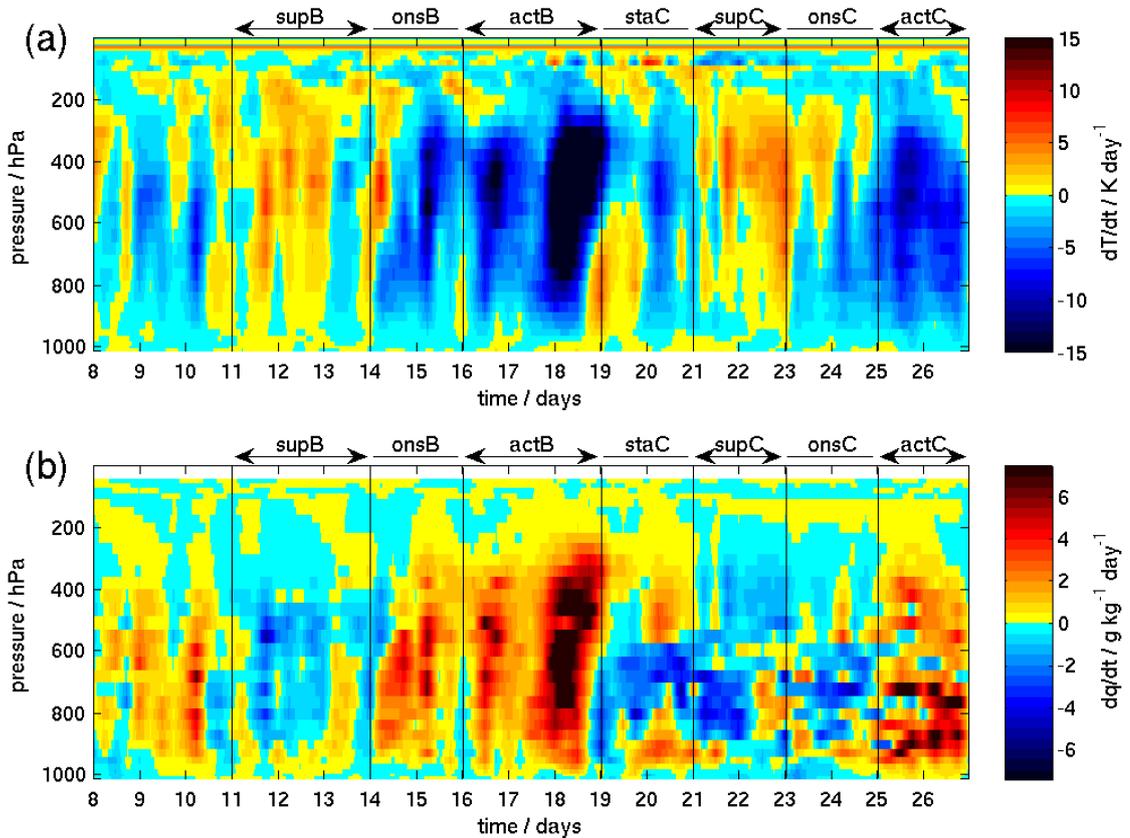


Figure 2.3: Time-pressure plots of large-scale dynamical tendencies in (a) temperature $dT/dt / \text{K day}^{-1}$ and (b) specific humidity $dq/dt / \text{g kg}^{-1} \text{day}^{-1}$ from the forcing dataset. The time axis is labelled in whole days (UTC) since the start of the month of January 1993, and vertical lines delineate the suppressed and active periods (described in section 2.3), as in later figures.

undermine the present study, which is more concerned with the variability of the physical processes represented by parameterisations, and in any differences which arise when the variability is artificially altered through the use of stochastic parameterisations. This study is not concerned with assessing whether such modifications to the model reduce biases relative to observations or otherwise improve model performance. Comparisons with TOGA-COARE observations are made in some of the following sections as a reality check and to aid in the understanding of the underlying physics.

Initial experiments with the default UM SCM and TOGA-COARE forcing data yielded a large drift in stratospheric temperature, with temperatures above 20km steadily declining by as much as 40K during the 19 day simulation (not shown). It is unclear whether this drift was present

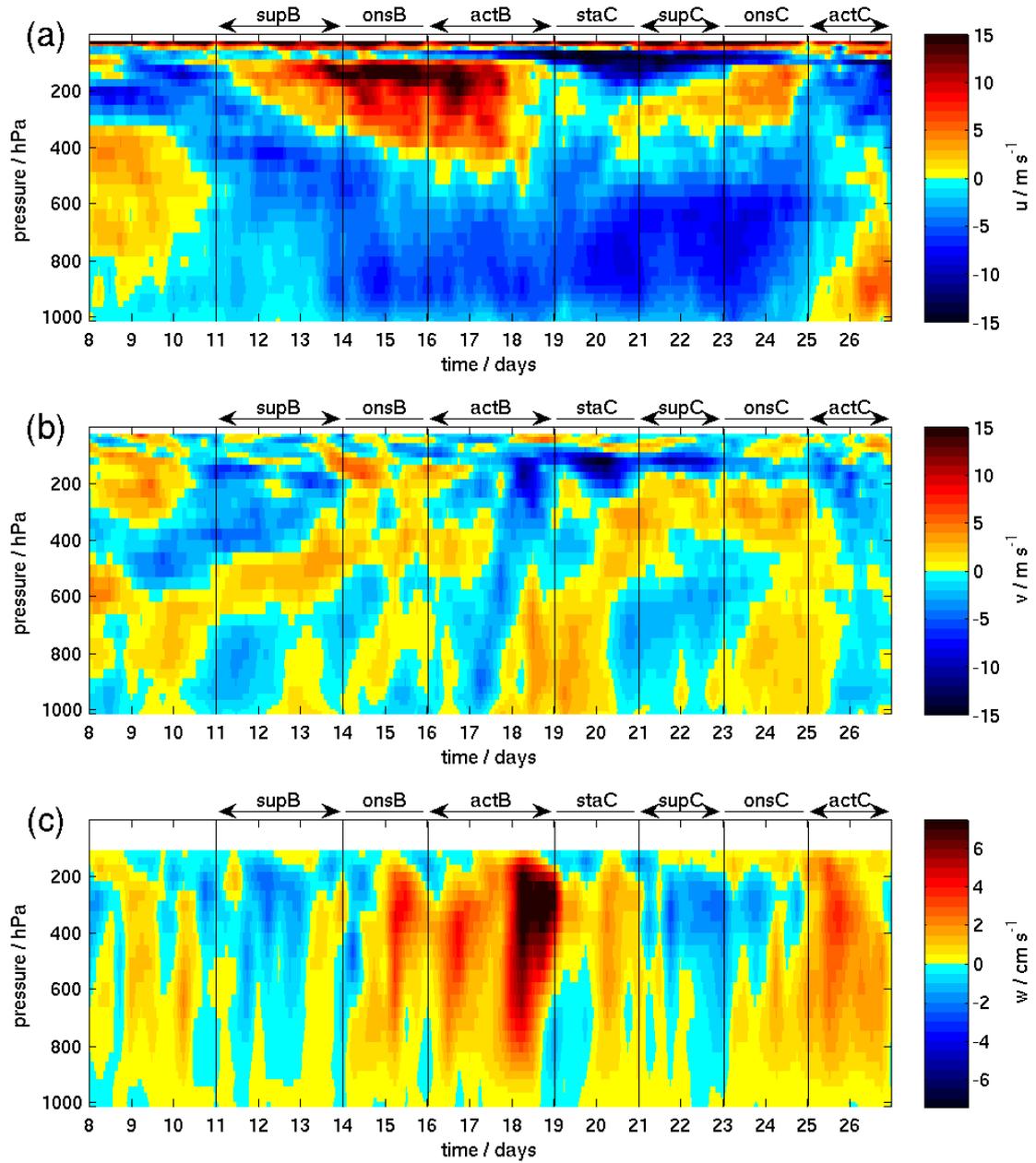


Figure 2.4: Time-pressure plots of observed winds; (a) zonal winds $u / \text{m s}^{-1}$, (b) meridional winds $v / \text{m s}^{-1}$ and (c) Vertical winds $w / \text{cm s}^{-1}$ from the forcing dataset.

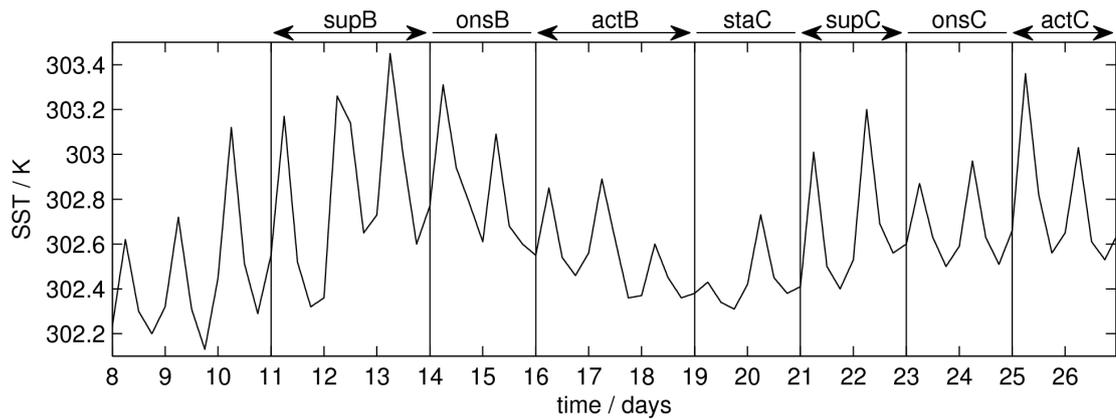


Figure 2.5: Time-series of Sea Surface Temperatures (SST) / K from the forcing dataset.

in other studies which use the same forcing data as they do not present results for stratospheric temperatures. Although the simulation appeared to behave realistically in the troposphere, the stratosphere drift caused unrealistic convection and cloud formation at around 20km. The drift was clearly a result of a radiative imbalance in the model, but the precise cause is unclear. It may be that the stratospheric water vapour is inaccurately specified in the initial conditions. But for the purpose of this study, the temperature drift was remedied by extending the large-scale forcing data-set into the stratosphere with constant positive values chosen to precisely offset the initial rate of erroneous stratospheric cooling.

Hack & Pedretti (2000) suggest that SCM frameworks should include a stabilisation component to prevent solutions from drifting into unphysical states due to a lack of large-scale dynamical feedbacks. A number of methods exist to parameterise the interaction of the thermodynamic profile with the large-scale dynamics in an SCM. The vertical advection terms can be calculated from the evolving thermodynamic profiles and the vertical velocity field (e.g. Hack & Pedretti 2000), the thermodynamic profiles can be relaxed towards observations (Hack & Pedretti 2000), the temperature profile can be held constant and vertical velocities calculated such that the vertical advection precisely balances the parameterised diabatic heating (Weak Temperature Gradient approximation, Sobel *et al.* 2007), or vertical advection maybe coupled to the diabatic heating via a gravity wave response model (Sardeshmukh 2004). However, as this part of the present study only aims to examine the non-dynamical component of the atmosphere's response to variability in parameterised processes, a dynamical stabilisation framework would seem an unnecessary

complication.

One of the SCM ensemble simulations of Hack & Pedretti (2000) (the ARM case) exhibited large deviations into unphysical states unless dynamical stabilisation was used. Although some drift in the temperature profile does occur in the experiments studied here (see section 2.3, figure 2.10), this does not seem to render the solutions unphysical, so no stabilisation framework is essential for a realistic simulation in this case. None has been employed here, and the large-scale dynamical tendencies are entirely prescribed.

2.3 OVERVIEW OF METEOROLOGY FOR THE CASE

To investigate the physical processes at work in the case simulated in this study, and check that the SCM is representing these processes, data from the Default UM SCM ensemble is presented alongside various observation-derived fields for the IFA in this section. The observed data, and information about their derivation, are available from the TOGA-COARE SCM website (<http://kiwi.atmos.colostate.edu/scm/toga-coare.html>). The observations are 6-hourly mean fields, so for fair comparison, 6-hourly means of the SCM data have been computed before calculating the ensemble ranges presented in this section. These will have less spread than the SCM data at a single time-step wherever variability on timescales shorter than 6 hours is important. The rainfall variability during the simulated period is investigated in subsection 2.3.1, and its relationship with the temperature and humidity profiles is explored in subsection 2.3.2. Variations in cloud are compared by studying the Outgoing Longwave Radiation in subsection 2.3.3, and a summary is given in subsection 2.3.4.

2.3.1 RAINFALL VARIABILITY

The case simulated by this experiment contains two suppressed phases, during which there is relatively little rainfall, each followed by an active phase with heavy convective precipitation. These are clearly visible in figure 2.6, which shows rainfall rate time-series from the SCM and from the observation-derived estimated moisture-budget of Ciesielski *et al.* (2003). The suppressed and active phases have been labelled here and on other figures for ease of comparison. These phases

are similar to those defined in Petch *et al.* (2007), except that they have been further delineated at times to distinguish the suppressed-to-active transition phases. Note that the time-axis in figure 2.6 and later figures is labelled with tick marks at 00h UTC each day, not local time. The time of local solar noon is 01:36 UTC, so the tick marks roughly correspond to the middle of the local day, rather than midnight.

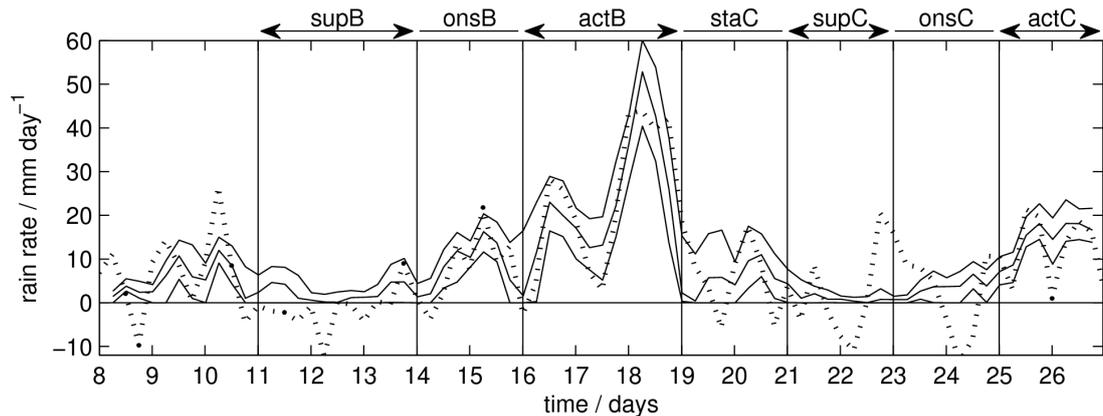


Figure 2.6: 6-hourly mean rainfall rates over the TOGA-COARE IFA; a budget-derived observational estimate (dotted), and the 5th, 50th and 95th percentiles from an ensemble of runs of the Default UM SCM used in this study (solid). The time axis is labelled in whole days (UTC) since the start of the month of January 1993. Vertical lines delineate the suppressed and active periods, as in other figures (see text).

During the period 12–14 January (supB), convection is largely suppressed. Moderate rainfall occurs on the 15th and 16th (onsB) during the onset of a vigorous active period from 17–19 (actB). This transition is accompanied by strengthening Easterly winds from the surface up to 500 hPa and Westerly winds aloft (see figure 2.4). After exceptionally heavy rains on the 19th, rains become moderate for the 20th and 21st (staC) as conditions stabilise, giving way to suppressed, dry conditions for the 22nd and 23rd (supC). Finally, light rains on the 24th and 25th (onsC) give way to moderately active conditions for the 26th and 27th (actC). At the same time the strong Easterlies give way to a Westerly wind burst near the surface. Moderate rains also occur during the first 3 days of the model runs, but this period will be treated as spin-up for the ensembles and so will not be studied in detail except to investigate initial condition sensitivity (section 2.4.1).

Negative values occasionally occur in the budget-derived rainfall rate shown in figure 2.6 because there are substantial errors in the estimated moisture budget used (these are discussed

in section 2.2.5). Since the forcings used to drive the SCM are derived from the same incomplete moisture budget, the model rain rates are likely also in error. So neither the modelled nor observation-derived rain-rates can be expected to be a precise representation of the true IFA mean, but it is reassuring that they are broadly in agreement over the timing and magnitude of rainfall events. An exception is the apparent heavy rain event at the end of supC in the budget-derived rainfall. However, negative values of comparable magnitude also occur before and after this event, suggesting the budget is not reliable during this period.

The vast majority of the rainfall in the SCM is produced by the convection scheme, with just a small fraction produced by the large-scale cloud / microphysics schemes during the active phases (not shown). To investigate the processes driving the rainfall variability shown in figure 2.6, time-series of other variables expected to affect convective rainfall are presented; the prescribed large-scale moisture convergence and the surface moisture fluxes are shown in figure 2.7, the rate of change of the total Column Water Vapour in figure 2.8 and the Convective Available Potential Energy (CAPE) and Convective Inhibition (CIN) in figure 2.9.

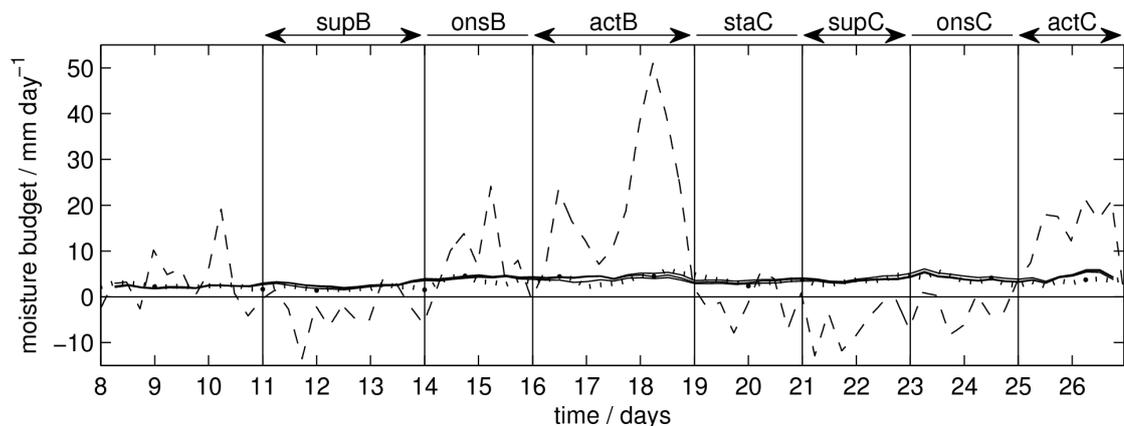


Figure 2.7: 6-hourly moisture budget; (dashed) IFA-average observed budget-derived large-scale moisture convergence (calculated as the column integral of the moisture forcing shown in figure 2.3b), (dotted) observed IFA-average surface moisture flux, and (solid) the 5th, 50th and 95th ensemble percentiles of surface moisture flux from the default UM SCM.

Comparing the rain rates in figure 2.6 to the large-scale moisture forcing in figure 2.7, it is clear that the rainfall closely follows the large-scale moisture convergence. However, the large-scale forcings in moisture and temperature (figure 2.3) are closely correlated (the primary term in both is the vertical advection), so this alone does not prove that the rainfall variability is being

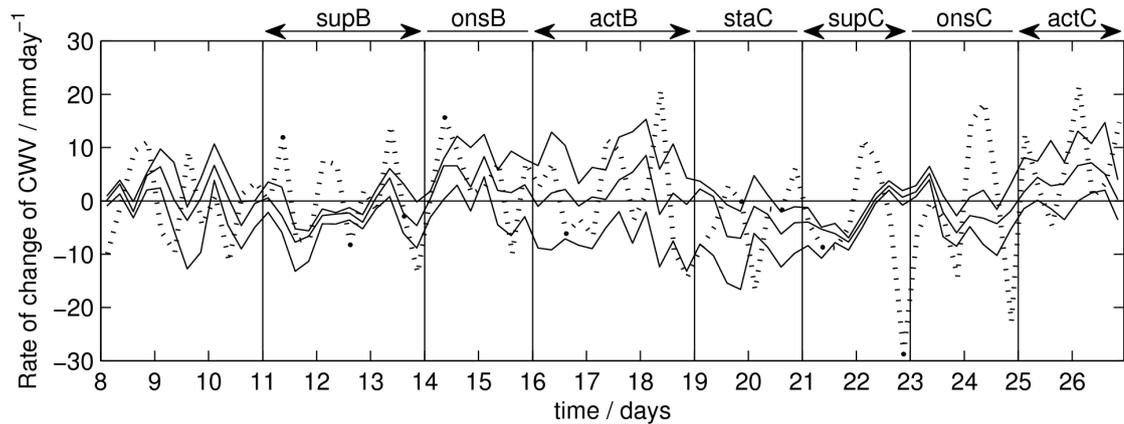


Figure 2.8: 6-hourly rate of change of the column-integrated water vapour; (dotted) observed IFA-average, and (solid) the 5th, 50th and 95th ensemble percentiles from the default UM SCM.

forced by moisture convergence, rather than variations in convective instability driven by the large-scale temperature forcing.

If the latter is the case, there should be a strong signal in the CAPE and/or CIN modulating the rainfall. Comparing the rain rates to figure 2.9a, there is no such signal in the CAPE; it is broadly similar in the observations and the SCM, and remains plentiful throughout the period, including the suppressed phases.

Considering CIN (figure 2.9b), in the observations there is no obvious influence of CIN on the rainfall variability either (e.g. during suppressed phases no significant rain occurs even when the CIN goes to near zero, and rainfall increases during actC despite the observed CIN being persistently high). There is some negative correlation between CIN and rainfall in the SCM (the highest correlation coefficient r relating 6-hourly CIN and rainfall was -0.36, found for the logarithms of the two variables, excluding zero values and the erroneous period onsC). However, a much stronger correlation was found between the 6-hourly large-scale moisture convergence and rainfall, with $r = 0.90$. Therefore, large-scale moisture convergence is presumably the primary factor driving the rainfall variability in figure 2.6, in both the observations and in the model.

The modelled CIN is generally lower than observed (except during onsC, when the CIN calculation in the model profiles is skewed by the appearance of a stable layer at 400 hPa with an unstable region above it, shown later in figure 2.10). The model and observed CIN both exhibit a

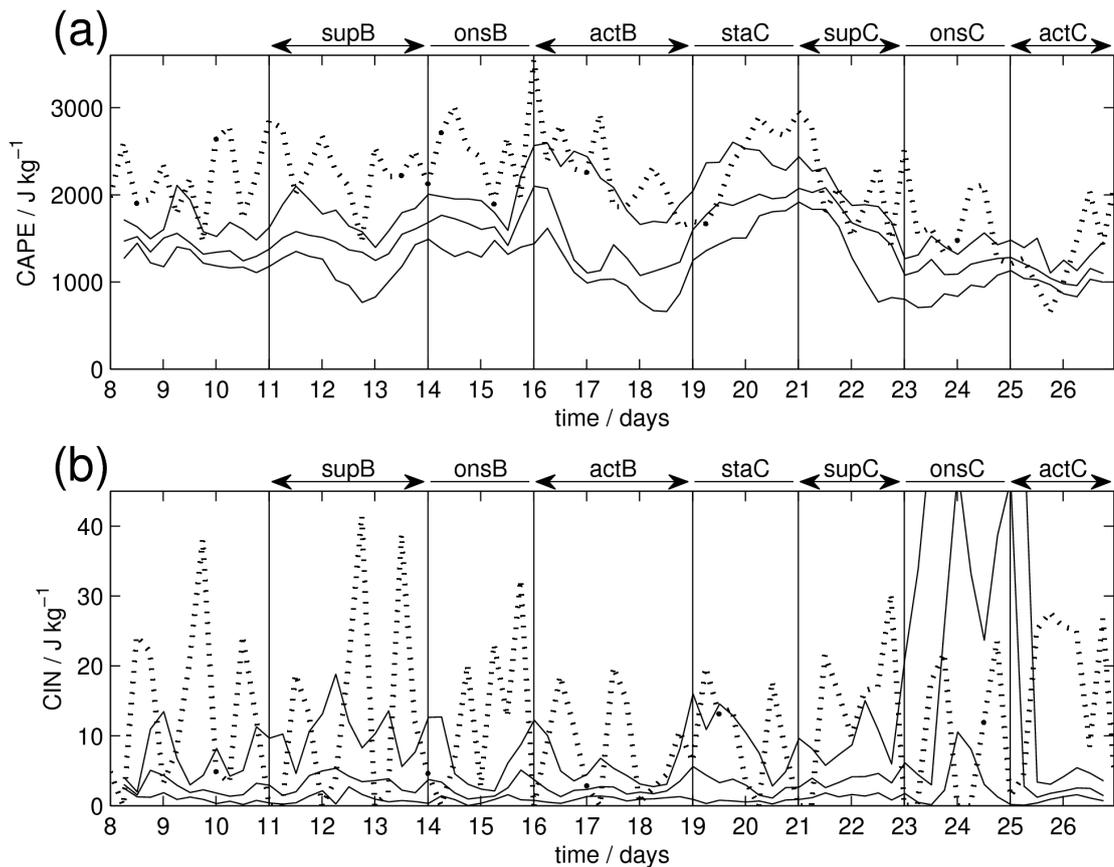


Figure 2.9: Time-series of (a) non-dilute Convective Available Potential Energy (CAPE), and (b) Convective Inhibition (CIN), calculated using a parcel ascent from the bottom level; (dotted) values computed from the 6-hourly IFA-mean soundings, and (solid) 5th, 50th and 95th ensemble percentiles from the DefUM ensemble. CAPE and CIN were computed from 6-hourly means of the modelled profiles before percentiles were calculated.

diurnal cycle, but they are not in phase; the observed CIN has maxima during the night or morning, whilst the SCM gives maxima during the day. This is presumably symptomatic of the failure of convection parameterisations to produce a realistic diurnal cycle (e.g. Yang & Slingo 2001).

Note in figure 2.7 that the surface moisture flux is relatively steady, with little ensemble variability in the SCM, and general agreement between the SCM and the observed IFA-mean. The time-variation in the moisture budget is dominated by a play-off between large-scale moisture convergence and rainfall, which are by far the largest terms but largely cancel each other during the active phases. During the suppressed phases, the drying (moisture divergence) forced by large-scale descent is partially off-set by fluxes of moisture from the surface. Figure 2.8 shows

that the residual of the rainfall, moisture convergence and surface flux still gives some significant variability in total Column Water Vapour (CWV). Some of the moisture introduced to the column during onsB and actB is not rained out but causes CWV to increase, whilst the column net loses water during staC, allowing rainfall to continue despite near-zero moisture convergence at that time.

2.3.2 TEMPERATURE AND MOISTURE PROFILES

To investigate the temperature and moisture variations during the case, figure 2.10 shows soundings averaged over each of the phases of convective activity. The observed temperature profile changed very little through the different phases of convective activity, in agreement with theories that suggest a Weak Temperature Gradient in the tropics (e.g. Charney 1963). The 6-hourly mean modelled temperature profile shows little ensemble spread, but has some systematic drift as the simulations progress, with a general cooling trend in the troposphere. Exceptions are the development of marked warm stable layers just above 200 hPa from onsB to staC, and at around 400 hPa during onsC.

As was noted in figure 2.8, there are considerable variations in the moisture profile in both the observations and the SCM. Whilst the dew-point in the boundary layer changes relatively little (constrained by contact with the sea-surface), the free troposphere shows a moistening trend from supB to actB, then a drying trend from staC to onsC, and a final re-moistening from onsC to actC. The SCM has a tendency to drift in to a somewhat drier state than the observations, with particularly dry episodes coinciding with the aforementioned stable layers during the onset phases onsB and onsC.

A number of studies have related the occurrence or not of tropical convection and precipitation to the humidity of the free troposphere. One such study is that of Holloway & Neelin (2009), who argue that the environment moisture profile through which a convective plume rises is a key control on its buoyancy via entrainment. The drier the environment air, the greater the evaporative cooling of the plume due to entrainment and the sooner it ceases to be saturated and buoyant. Convective rainfall then cannot initiate unless the environment is sufficiently moist to permit an entraining convective plume to remain buoyant over a sufficient depth for precipitation to form,

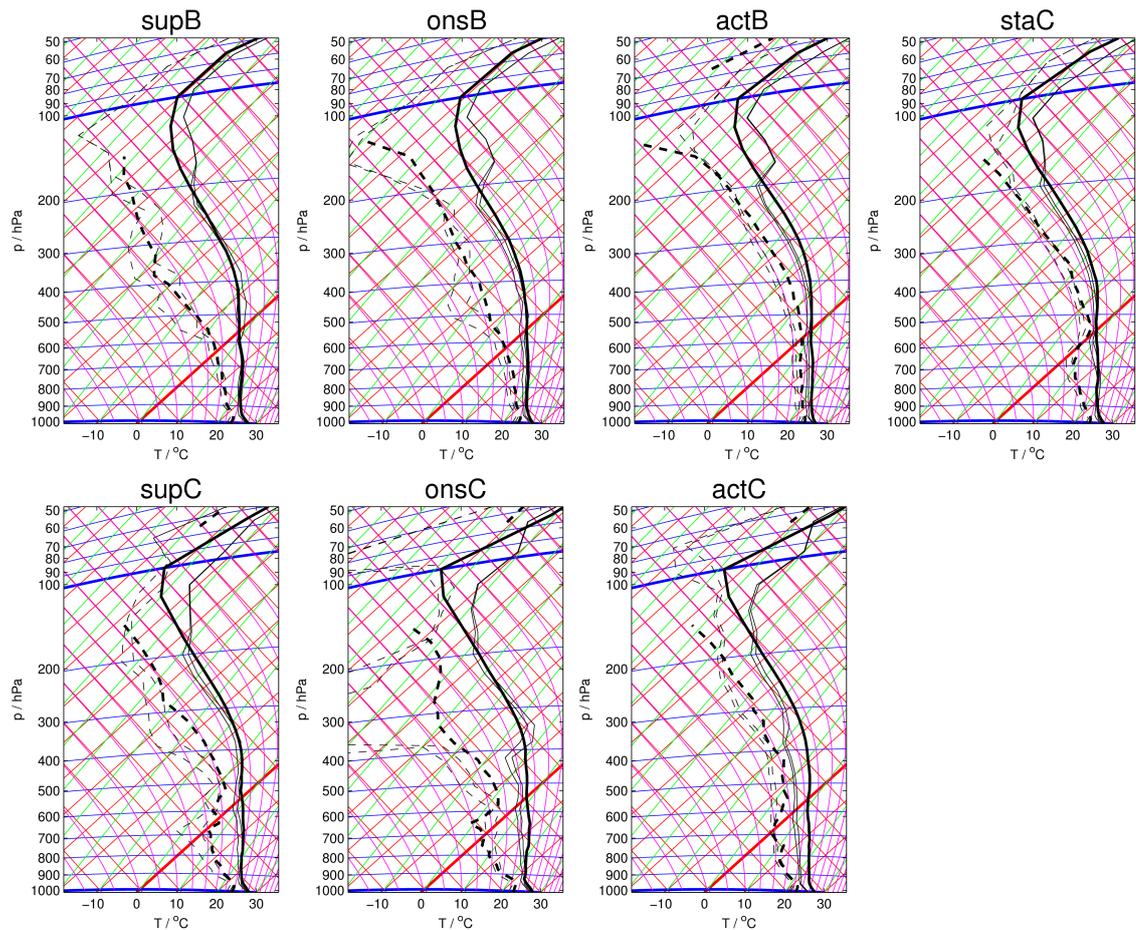


Figure 2.10: Tephigrams showing profiles of dry-bulb (solid) and dew point (dashed) temperatures averaged over each of the phases of convective activity shown in figure 2.6; (thick lines) average of soundings over the TOGA-COARE IFA, and (thin lines) 5th and 95th ensemble percentiles from the DefUM ensemble. Dew-points are not plotted where the observed specific humidity was rounded to zero in the available data. Model ensemble percentiles were calculated after time-averaging.

even if there is plentiful CAPE and low-level moisture convergence. This could explain the highly non-linear dependence of tropical rainfall on Column-integrated Water Vapour (CWV) observed by Bretherton *et al.* (2004), who suggested that rainfall only commonly occurs above some threshold value of CWV (scaled by a measure of free tropospheric temperature), beyond which it increases rapidly.

The modelled profiles (thin lines in figure 2.10) are indeed consistent with a relationship between free-tropospheric humidity and rainfall; the smaller the dew-point depression in the free troposphere in a given phase, the more rain occurs during that phase. This supports the notion that

the large-scale moisture convergence controls the rainfall variability in the SCM, via convective entrainment in the free troposphere as described above. Heavy convective rainfall only occurs when the large-scale forcings act to make the free troposphere sufficiently moist, and excess moisture convergence beyond this point is rained out. The SCM achieves this even though its convection scheme is closed on CAPE, not moisture convergence. This suggests the convection parameterisation successfully represents the environment humidity control on convective rainfall described above, via its representation of entrainment in the plume model.

The observed profiles broadly support rainfall dependence on dew-point depression as in the SCM, except towards the end of the period. The free troposphere remains very dry during onsC and actC and, unlike in the SCM, the profile does not cool so as to reduce the dew-point depression. It is unclear why substantial rains occurred during actC in reality (or equivalently, why the atmosphere equilibrated in a much drier state during actC than it did during other periods with similar forcings and rain-rates). The observed profile during actC is just as dry as during the suppressed phase supC, with slightly lower CAPE and greater CIN. Perhaps the observed moisture profiles during onsC and actC are in error due to the sparsity of observations at that time, as discussed by Petch *et al.* (2007). Or perhaps the IFA came under the influence of some organised convective system during actC.

Of course, in the real world there is a two-way interaction between the convective activity and the large-scale; large-scale ascent promotes cooling and moistening of the troposphere through vertical advection, which promotes deep convection, which produces diabatic heating which in turn promotes large-scale ascent. However, in these experiments there is no coupling of the large-scale forcings to the diabatic heating, so no such feedbacks occur. These feedbacks have been explored in SCM frameworks which parameterise the dynamical response to diabatic heating (e.g. Sobel *et al.* 2007).

2.3.3 OUTGOING LONGWAVE RADIATION

It would be preferable to verify that the model is producing cloud-fields which are reasonably realistic and representative. One diagnostic that is easily obtained is the IR brightness temperature T_b measured by the Japanese Geostationary Meteorological Satellite, which was observing the

tropical West Pacific at the time of TOGA-COARE. This should give an indication of cloudiness over the IFA, with lower brightness temperatures indicating the presence of high, cold clouds. To compare this with the SCM's Outgoing Longwave Radiation (OLR) diagnostic, it has been converted to an equivalent longwave flux F_{LW} using Stephan-Boltzmann's law $F_{LW} = \sigma T_b^4$, where $\sigma = 5.6704\text{E-}8 \text{ J s}^{-1} \text{ m}^{-2} \text{ K}^{-1}$ is the Stephan-Boltzmann constant. The time-series of F_{LW} is compared with the SCM ensemble range of OLR in figure 2.11.

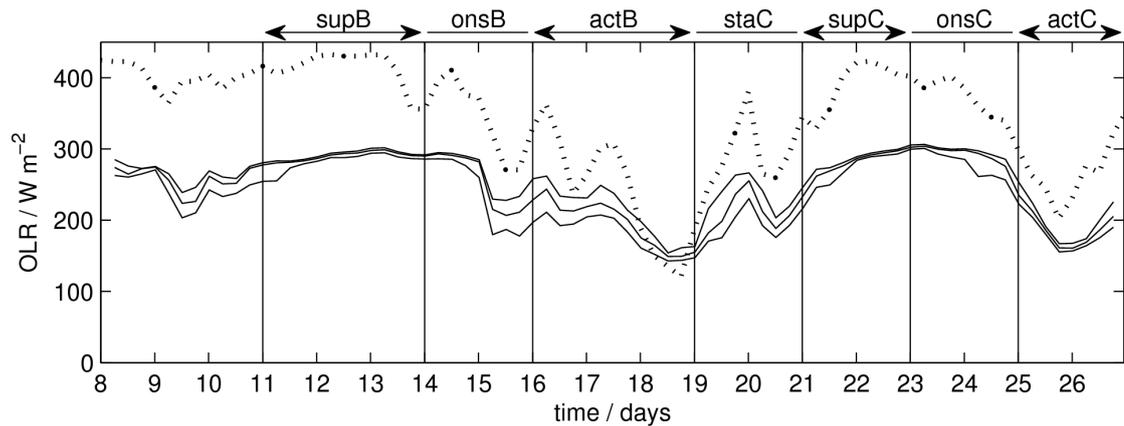


Figure 2.11: Time-series of 6-hourly mean Outgoing Longwave Radiation (OLR); (dotted) equivalent LW flux estimated from the IFA-average IR brightness temperature as observed by the Japanese Geostationary Meteorological Satellite (GMS), and (solid) 5th, 50th and 95th ensemble percentiles from the DefUM ensemble.

The equivalent LW flux gives unrealistically high values for OLR. The values nearly exceed the daily-mean downward solar radiation at top-of-atmosphere (429 W m^{-2}), implying a considerable radiative loss (accounting for albedo), whereas the Tropical West Pacific is in reality a major heat source in the climate system. It would be naive to assume a spectral brightness temperature measurement will be an accurate measure of the broadband OLR, so this discrepancy is not surprising. The key point to note is that the SCM OLR captures the variability in the observed brightness temperature well, suggesting that it is producing cold cloud in the right proportion at the right times. A minor exception is the observed slightly cloudy spell at the end of supB. Comparing with the rainfall time-series in figure 2.6, it seems the SCM does produce some rain at that time but there is no signature of any associated cold cloud in the model OLR.

2.3.4 SUMMARY

In summary, the default UM SCM is found to represent the observed phases of suppressed and active convection well. The diurnal cycle of convective inhibition is completely out of phase with the observations, but this turns out to be relatively unimportant, as the rainfall variability is primarily controlled by large-scale moisture convergence. This is via its forcing on the free-tropospheric humidity profile, which controls the occurrence of deep, rain-bearing convection via entrainment of environment air into convective plumes. The SCM appears to capture this well. The moisture budget is dominated by a play-off between moisture convergence and the convective rainfall response, with the rainfall accounting for most of the ensemble variability in the SCM's moisture budget. The surface moisture flux acts to replace the moisture removed from the column by the large-scale forcings during suppressed phases. The temperature profile in the SCM gradually drifts away from the fairly constant observed profile, with a modest cooling trend (and a slight drying). But it doesn't appear to diverge to any wildly unphysical state, suggesting that useful results can be obtained without using a dynamical stabilisation framework. The OLR in the SCM follows a very similar pattern of variations to the satellite-observed brightness temperature, suggesting that the timing and height of cloud in the SCM is broadly realistic.

2.4 RESULTS

In this section, preliminary results are presented investigating the variability of deterministic and stochastic parameterisations in the SCM. The variability of deterministic schemes is explored in subsection 2.4.1, by studying their sensitivity to initial condition perturbations. This is then related to the variations in convective activity in subsection 2.4.2, and compared to the SCM's sensitivity to stochastic physics perturbations in subsection 2.4.3.

2.4.1 SENSITIVITY TO INITIAL CONDITION (IC) PERTURBATIONS

To investigate the initial condition sensitivity of the SCM, ensembles have been performed with two different sets of IC perturbations. First, minimal perturbations were made, consisting of per-

turbations to temperature on the bottom model level only, drawn from a flat distribution between ± 0.25 K. Second, the standard IC perturbations described in section 2.2.3, which have a larger amplitude and vertical extent and act on specific humidity as well as temperature, were applied.

Time-series of temperature at 800 hPa for ensembles of Default UM SCM runs with the minimal and standard IC perturbations are shown in figure 2.12. In both cases the ensemble diverges over the first few days of the simulation, giving a temperature range of about 2K at many times. After a few days the spread of the ensemble saturates and stops increasing, but proceeds to vary considerably throughout the runs. The ensemble with the minimal IC perturbations takes slightly longer to saturate (about 4.5 days, compared to about 3.5 days for the standard perturbations), but thereafter the behaviour of the two appears qualitatively very similar. This suggests that once the ensembles saturate, their behaviour is no longer sensitive to the IC perturbations which initially caused them to diverge. Once this occurs, the spread of the ensembles is no longer a function of the initial conditions, so it must be a function only of the model and the large-scale forcings. The ensemble members vary rapidly within a wider envelope whose mean value and spread evolves with time as it is forced by the prescribed large-scale dynamics. Note that each member (looking at the control run for example) may be one of the warmest at one time and yet be one of the coolest a couple of days later, suggesting that the ensemble decorrelates with itself multiple times during the run.

In theory the control run (thick solid line), which has no IC perturbations, should be identical for both the ensembles shown in figure 2.12. However, despite using the same model and the same initial conditions, the control runs begin to diverge visibly from one-another after about 3.5 days. Other than the IC perturbations, the only difference between these two ensembles is that the standard IC perturbation ensemble was later rerun in order to obtain more diagnostics, after an upgrade had been made to the compiler on the computing platform used, which rendered results non bit-reproducible. So the tiny differences in numerical truncation error associated with the compiler upgrade were sufficient to force the realisations followed by the SCM control run to diverge. This is a sobering illustration of the capability of SCMs to exhibit very strong non-linear sensitivity to small perturbations, as discussed by Hack & Pedretti (2000), and reinforces the importance of using ensembles in this context. For example, just comparing the temperatures given by the two control runs, one could falsely conclude that the latter run is systematically

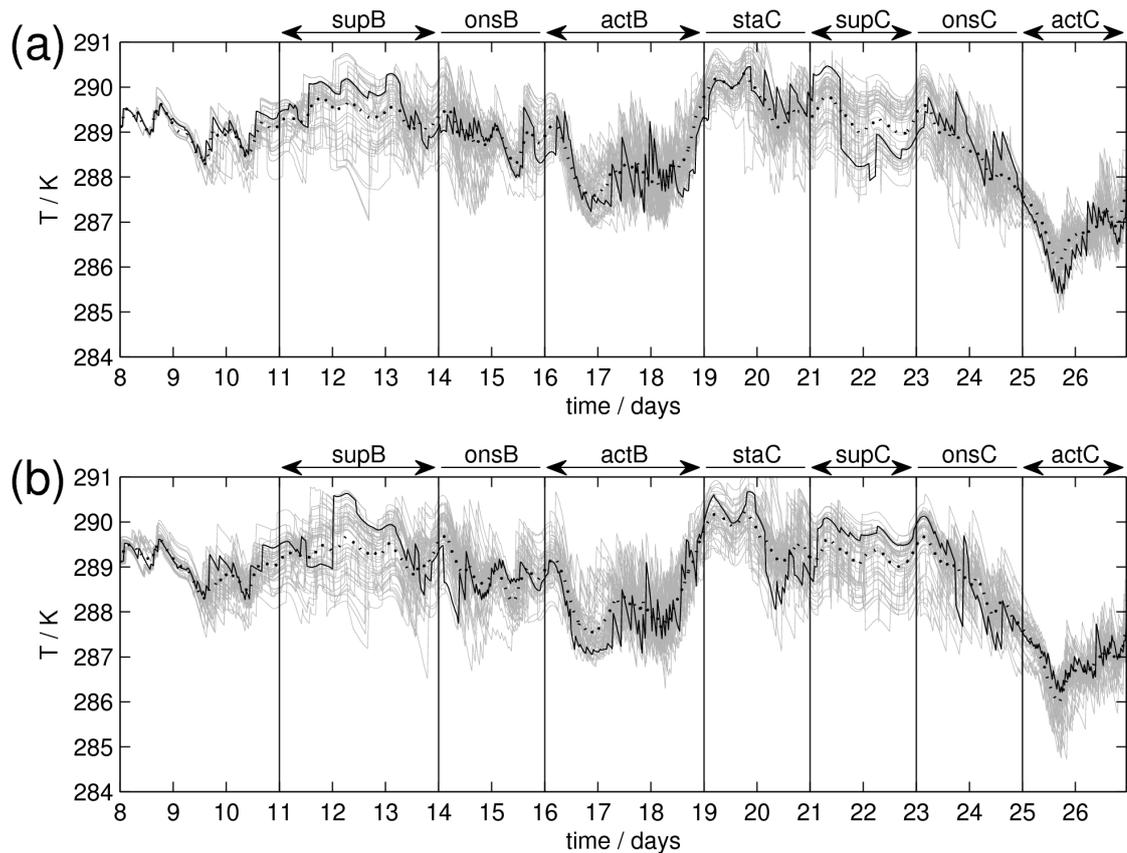


Figure 2.12: Time-series of temperature at 800 hPa from the Default UM SCM; (thick solid line) the control run, (thin solid lines) the 39 perturbed ensemble members, and (dotted) the ensemble mean, with (a) minimal, and (b) standard IC perturbations (described in text).

warmer and steadier during supC.

Figure 2.13 shows similar temperature time-series for ensembles of runs with the Kain-Fritsch convection scheme. Although these runs show a similar pattern of mean response to the large-scale forcings (i.e. cooling during the onset of active phases and warming during the stabilisation periods at the ends of active phases), the envelope of likely realisations is much narrower than in the Default UM ensembles for both sets of IC perturbations. For the minimal IC perturbation ensemble (figure 2.13a) the number of distinct trajectories on the plot appears much smaller than the size of the ensemble. This is clearly apparent in figure 2.14, which shows the two Kain-Fritsch ensemble time-series of temperature as in figure 2.13 but zoomed in on the 22nd and 23rd January. Even after 15 days of simulation, the 40 ensemble members remain tightly clustered, following only 6 distinct realisations. In this case the ensemble has failed to diverge, and statistics such

as the ensemble mean and spread will not representatively sample the full solution variability. However, when the standard IC perturbations are applied, a fully diverged set of realisations is evident (figure 2.14). But the spread is still considerably narrower than that seen in the Default UM ensembles.

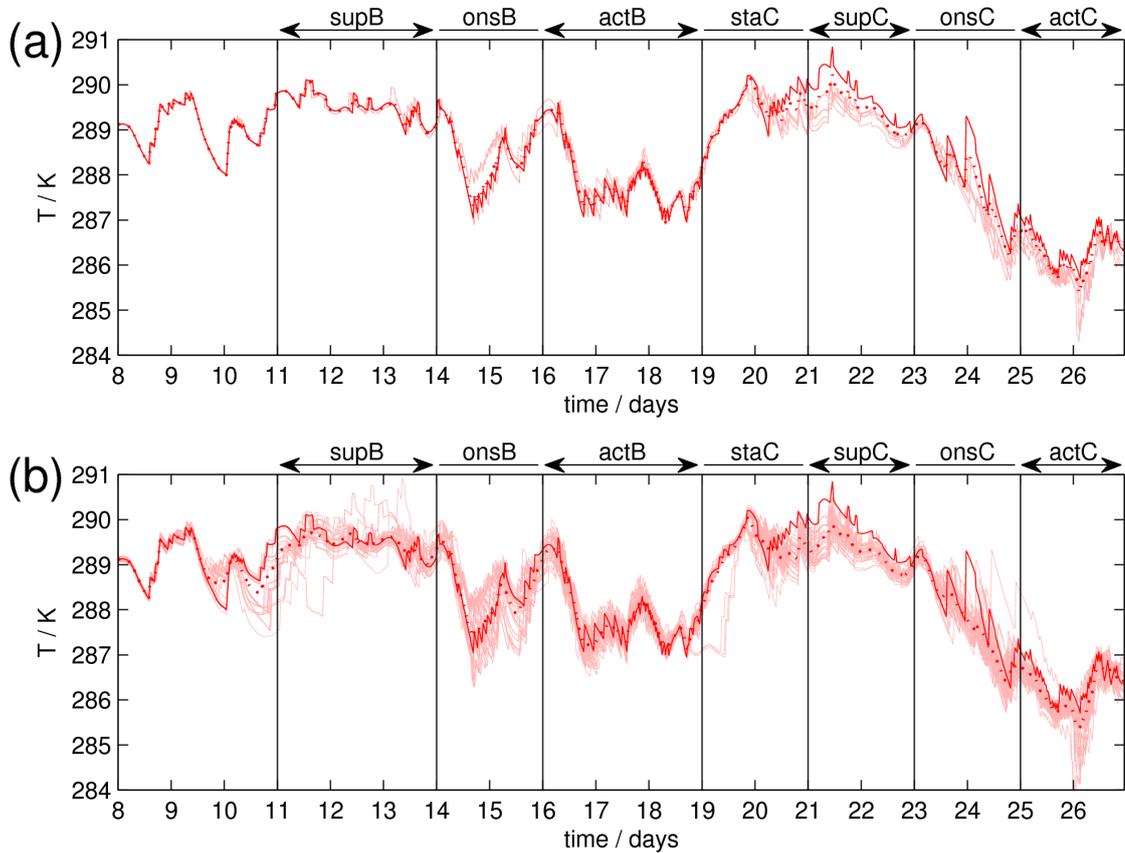


Figure 2.13: As figure 2.12, but for the Kain-Fritsch model variant.

Both the Default UM and Kain-Fritsch ensembles evidently exhibit discrete and sometimes large jumps in temperature, ranging in frequency from a couple of events per day during the suppressed phases to near-continuously noisy behaviour during active phases. The tendency of these jumps to happen more often during convectively active periods suggests that they are related to the convective parameterisation. Figure 2.15 shows the triggering behaviour of the convection scheme, for the Kain-Fritsch ensembles during the 22nd and 23rd January as in figure 2.14. Now comparing figure 2.15a with figure 2.14a, it is clear that the time-steps on which the convection scheme calls in many ensemble members correspond to those when the temperature jumps occur, for example just after $t = 21.4$ days and just before $t = 21.8$ days. Further, if one checks care-

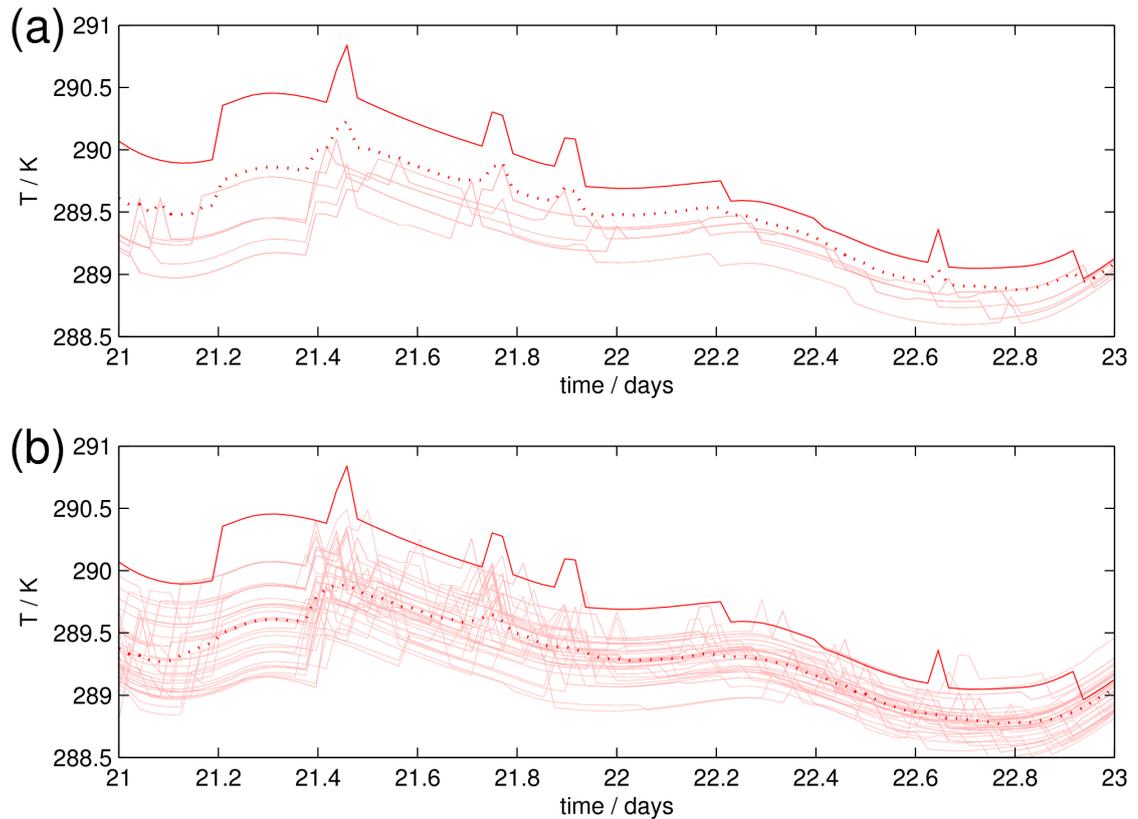


Figure 2.14: As figure 2.13, but zoomed in on the 22nd and 23rd January for clarity.

fully, the clustering behaviour of the temperature trajectories is precisely reflected in the timing of convective triggering; there are only 6 distinct sets of triggering time-steps in figure 2.15a, each corresponding to the jumps in one of the 6 trajectories in figure 2.14a. 16 of the ensemble members, including the control run, all trigger convection at the same times and follow the same trajectory, which fortuitously turns out to be an outlier when the ensemble better samples the likely realisations (when the standard IC perturbations are applied; figure 2.14b). When these larger IC perturbations are used, the clustering in convective triggering behaviour correspondingly breaks down, giving a much more random-looking scattering of triggering time-steps for different ensemble members in figure 2.14b. Some patterns remain, such as the tendency for convective triggering at $t = 21.4$ days; these are presumably driven by the large-scale forcings.

The discrete jumps associated with convective triggering appear to dominate the high-frequency variability in the SCM simulations studied here. Also, the divergence of ensemble members into distinct realisations often first occurs at such events. For example, the two control

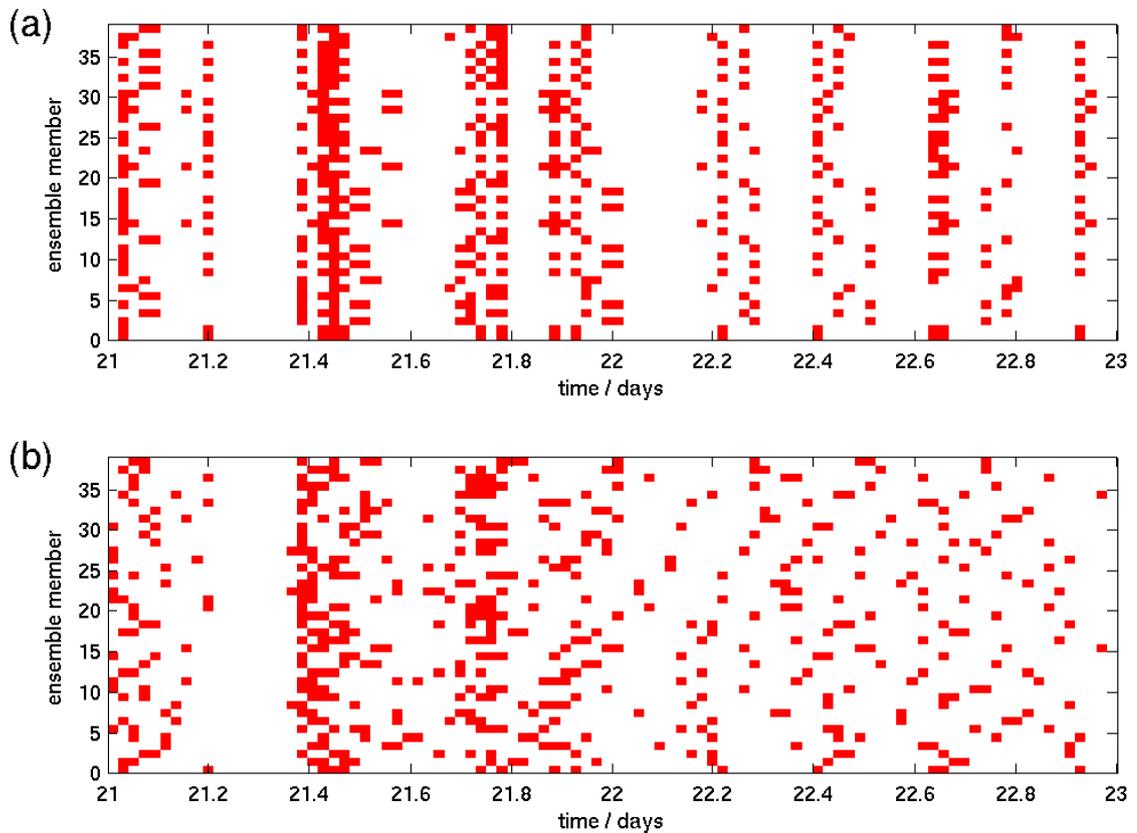


Figure 2.15: Time-series showing when the convection scheme is active in each ensemble member in the KF ensembles, for the same time period as in figure 2.14; (a) with minimal IC perturbations, and (b) standard IC perturbations. A box is coloured if there is any convective increment in temperature anywhere in the model column in the corresponding ensemble member and time-step.

runs in figure 2.12 first differ visibly when a small convective jump occurs at $t = 10$ days in (b) but not in (a). After this the timings and magnitudes of convective events differ increasingly between the two, and the large convective jump at $t = 11.5$ days finally forces them in to seemingly un-related trajectories.

Comparing figures 2.12 and 2.13, it is perhaps surprising that the SCM's spread response differs so much when a different convective parameterisation is employed. This is in contrast to the findings of Hume & Jakob (2005), who show in their section 4.1 that the two different SCMs used in their study produce very similar ensemble spread in response to initial condition uncertainty for a range of variables, including 850hPa temperature. Given the role the convection scheme plays in causing the SCM ensembles performed in the present study to diverge and vary,

and the large difference in spread between the DefUM and KF ensembles, it is clear that the ensemble spread is highly sensitive to the convective parameterisation. Presumably the SCMs used by Hume & Jakob (2005) had fortuitously similar convective behaviour.

The SCM runs of Hack & Pedretti (2000) were similar to those studied here in that the ensemble spread did not simply grow steadily and then saturate at some equilibrium value, but varied greatly in response to changes in the large-scale forcing and atmospheric state. At some times the ensemble of states diverges rapidly, whilst at others it reconverges. This behaviour suggests that after some time the ensemble spread may become more connected to the variability of the physics represented in the SCM than it is to the initial condition perturbations used to provide the initial range of states. The mechanisms responsible for generating variations in ensemble spread in the UM SCM will be explored further in the following subsection 2.4.2.

Hack & Pedretti (2000) also show that large bifurcations exist in their SCM solutions, with ensemble members clustering into multiple preferred modes at times, even after the ensembles have fully diverged. They point out that when such behaviour occurs, statistics such as the ensemble mean and standard deviation may not be representative or meaningful. They found that this was a particular problem in a case simulating very vigorous convection over continental land, in which the occurrence of multiple modes was associated with the development of unphysical model states in the boundary layer. However, their simulations of convection over the tropical oceans were more stable and not so prone to unphysical behaviour, and the same appears to be the case in the simulations studied here. The Kain-Fritsch ensemble does show a little tendency towards bifurcations (for example, during the 20th January in figure 2.13b), but there was no obviously unphysical mode associated with this, and the ensemble remains more evenly spread throughout most of the simulation. The occurrence or not of bifurcations appears to be a function of both the model (here they occur for the KF scheme but not the DefUM) and the large-scale forcings (in Hack & Pedretti 2000, they had a greater tendency to occur in the continental case than the oceanic one). But in this study, such behaviour is limited, and ensemble statistics are generally meaningful, as long as the ensembles fully diverge. The standard IC perturbations appear to be sufficient to achieve this.

2.4.2 ENSEMBLE SPREAD AND VARIABILITY

It was noted in section 2.4.1 that once an SCM ensemble saturates, its spread continues to vary, presumably as a function of the forcings and model physics rather than the initial conditions. This is explored further in this section.

Figure 2.16 shows a time-height plot of the ensemble spread in temperature in the default UM SCM ensemble. The spread appears to follow different characteristic regimes during suppressed and active phases, whilst behaving in a more unsteady manner during transition periods between the two. During the onset phases prior to each active phase, there is a peak in ensemble spread at mid-levels followed by a rapid decrease in spread.

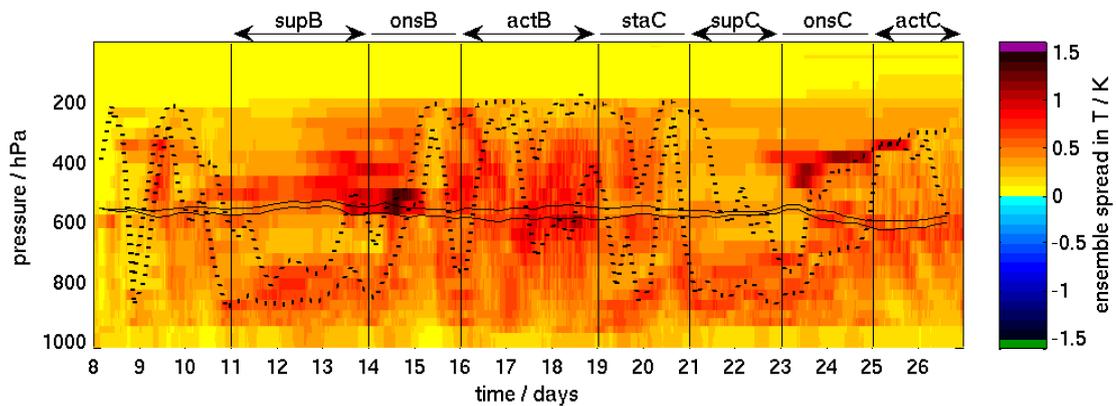


Figure 2.16: Ensemble spread (standard deviation) in temperature for the Default UM ensemble using the standard IC perturbations. Over-plotted are (dotted) the 25th and 75th percentiles of the convective cloud top pressure, and (solid) the 5th and 95th percentiles of the melting-level pressure. The cloud-top percentiles are determined at each time-step from the subset of ensemble members in which the convection scheme triggers, then smoothed for presentation here using 6-hourly averaging.

Most notably during the onset phases, the pattern of ensemble spread appears to be related to the convective cloud top height. For example, the peak in ensemble spread in the mid-troposphere on the 15th corresponds to where the ensemble contains considerable uncertainty regarding whether the convective cloud top penetrates the melting-level inversion. During the following day the ensemble members fall in to agreement that convection exceeds this height, and the ensemble spread drops throughout the troposphere. At many times during the run the maximum ensemble spread is at or near the melting level. This highlights the strong sensitivity

of deep convection there, and is consistent with the findings of Inness *et al.* (2001), who found that adequately resolving the melting-level inversion was important in the simulation of tropical rainfall variability.

Another interesting feature is the sloping layer of high ensemble spread around the 24th and 25th, which increases in height from around 450 to 350 hPa during onsC. This layer closely follows the 75th percentile of convective cloud top height, indicating an ascending lid on the convection (figure 2.10 in section 2.3 confirms a strong inversion was present in the model at this time / altitude). The ensemble spread is large here because the ensemble members produce a range of different heights for this lid, which has a sharp temperature gradient across it.

Aside from ensemble spread, the variability of the ensemble can also be quantified by the rate at which the ensemble decorrelates with itself. A decorrelation timescale τ can be calculated, following Dawdy & Matalas (1964), as:

$$\tau = \frac{1 + r_{\Delta t}}{1 - r_{\Delta t}} \Delta t \quad (2.10)$$

where $r_{\Delta t}$ is the ensemble autocorrelation at lag Δt . τ can be thought of as the time taken for the ensemble members to “forget” where they were relative to each-other. Figure 2.17 shows the ensemble decorrelation timescale for the default UM SCM. For comparison, figure 2.18 shows the temperature tendency due to convection in the SCM.

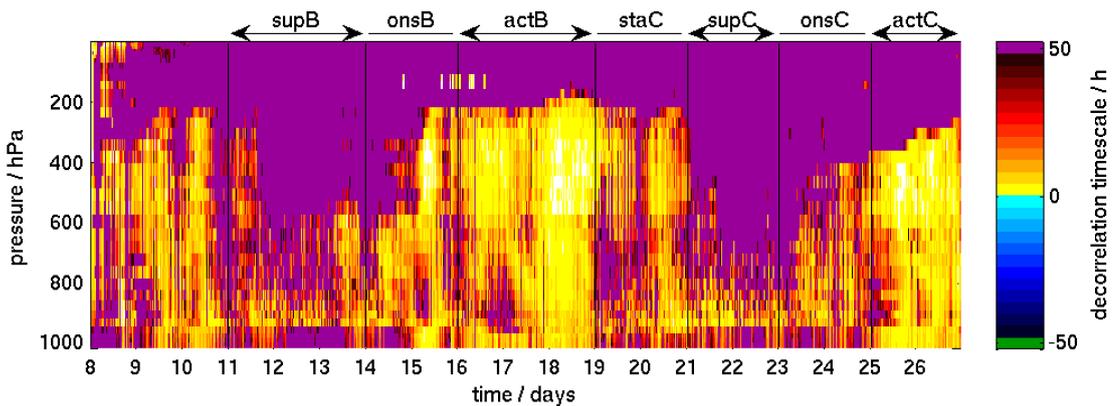


Figure 2.17: Ensemble decorrelation timescale for temperature for the Default UM ensemble. Derivation is described in the text.

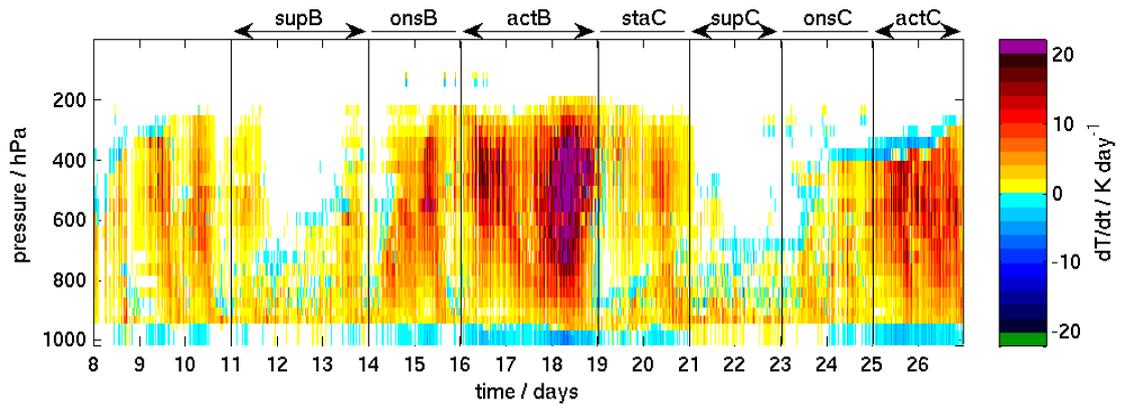


Figure 2.18: Ensemble-mean convective temperature tendency for the Default UM ensemble.

There is a striking correspondence between the convective temperature tendency and the ensemble decorrelation timescale τ . Where the convection is active, the ensemble decorrelates with itself in just a few hours; where convection does not occur, τ is greater than 2 days. This confirms the results from section 2.4.1, which suggested convection was the key process generating temperature variability in the ensemble. Similar results (not shown) apply for other fields such as specific humidity.

2.4.3 SENSITIVITY TO STOCHASTIC PARAMETERISATION (SP) VERSUS INITIAL CONDITION (IC) PERTURBATION

Here, the variability generated by a simple stochastic physics scheme implemented in the SCM (the Multiplicative Perturbation method of Buizza *et al.* 1999, described in section 2.2.2.4) is compared with the internal variability of the default parameterisation set (generated by initial condition perturbations as in the previous subsections).

Figure 2.19 shows time-series of the column-mean ensemble spread of temperature, for the Default UM ensemble with the minimal IC perturbations described in subsection 2.4.1 and the standard (larger and more extensive) IC perturbations described in subsection 2.2.3, and also for an ensemble which includes the multiplicative noise scheme, both with and without the standard IC perturbations added. The corresponding plots for relative humidity are shown in figure 2.20.

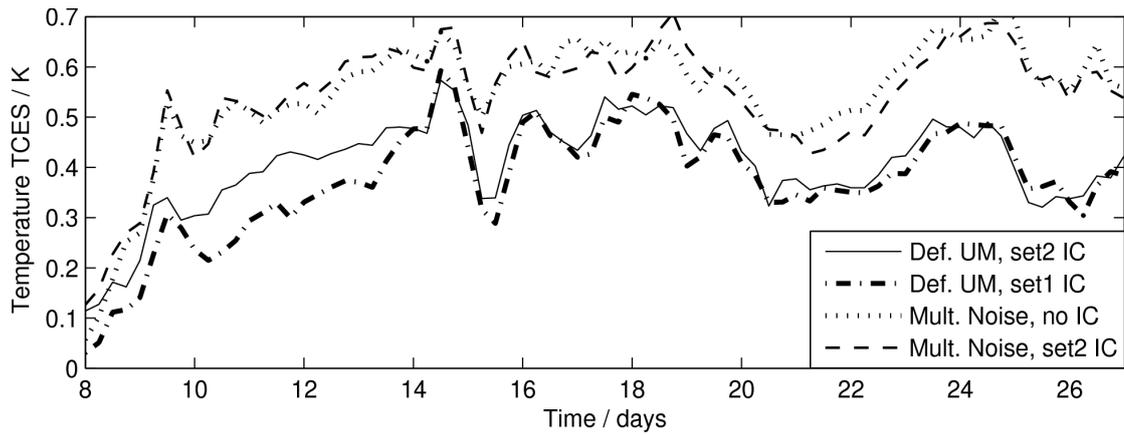


Figure 2.19: Column-mean ensemble spread of temperature in the default UM with minimal (dash-dotted) and standard (solid) IC perturbations, with multiplicative-noise perturbations (dotted), and with both standard IC perturbations and multiplicative noise (dashed).

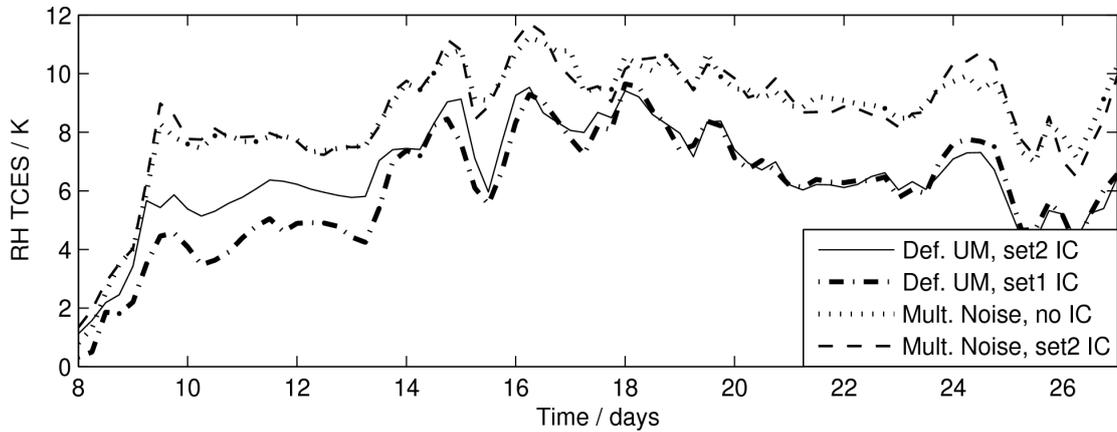


Figure 2.20: As Figure 2.19, but for relative humidity rather than temperature.

Looking first at the two Default UM ensembles, there is more spread over the first 6 days using the larger standard IC perturbations. However, the minimal and standard IC ensembles look very similar beyond 6 days, suggesting that the ensemble spread has saturated in both. This confirms that the saturated level of ensemble spread in temperature is independent of the size and nature of the IC perturbations, but rather provides a measure of the inherent variability of the SCM, as hypothesised earlier. For the Kain-Fritsch scheme, the larger IC perturbations produce larger ensemble spreads throughout (figure 2.13), but this is because the spread did not saturate when the minimal IC perturbations were used.

In a stochastic physics (SP) SCM ensemble, the stochastic method introduces some hitherto

neglected source of variability. One might anticipate that the physics perturbations would allow the SP ensemble to explore at least those realisations accessible to its deterministic analogue¹. If this is true, then IC perturbations should have little effect on ensemble spread when implemented in an SP ensemble.

It is clear from figures 2.19 and 2.20 that beyond the first 36 hours or so, the ensembles including multiplicative noise SP perturbations have spreads that are consistently larger than those occurring in the IC-only ensembles, typically by a factor of about a third. The inclusion of IC perturbations in addition to multiplicative noise slightly increases the spread during the first day, but has no significant effect thereafter. This is consistent with the idea that the IC perturbations allow one to sample different realisations, but do not affect the underlying distribution of likely realisations that emerges once the spread of the ensemble saturates. Similar conclusions apply for the other stochastic methods used (not shown).

Comparisons of the effects of IC and SP perturbations have been made before in the context of global GCM ensemble prediction systems. Buizza *et al.* (1999) found that IC-only ensembles produced consistently larger spread than SP-only ensembles, and that ensembles with IC and SP perturbations produced greater spread still². Teixeira and Reynolds (2007) found similar results over the tropics using a multiplicative noise scheme applied only to the moist convective tendencies (their figure 7a). Although these results differ from ours in placing far greater emphasis on IC perturbations, this is not surprising given the context. In particular, we focus on the saturated level of ensemble spread due to IC perturbations whereas in the cited studies, the runs do not reach saturation (see Teixeira and Reynolds (2007), figure 7a for example). Also those studies used much larger IC perturbations designed to sample IC uncertainty. Clearly the timescale for ensemble saturation is much longer in a 3D global simulation compared to the SCM, reflecting the slower, larger-scale dynamical modes which dominate the variability in the former.

Finally, before a rigorous comparison of the various deterministic and stochastic SCM configurations described in subsection 2.2.2 is made (in the next chapter), it would be desirable to

¹By which we mean the equivalent configuration with the stochastic component disabled, providing of course that such an equivalent is well-defined. For example, for a stochastic method in which model parameters are selected randomly, the deterministic analogue is simply a simulation with the default parameter set.

²This is shown for forecast days 3, 5 and 7 in their Table 1a.

check that the variability generated by all these schemes is reflected by changes in the ensemble spread, as it is for the multiplicative perturbation scheme shown in figures 2.19 and 2.20. 40-member ensembles, including the standard initial condition perturbations, were generated for all of the model variants. Figure 2.21 shows time-series of their column-mean ensemble spreads in temperature and specific humidity.

The various SCM configurations exhibit a broad range of ensemble spreads, which differ significantly from one-another. The spread time-series show large differences in the amplitude of variability in the different ensembles, and some differences in the time-variation of spread for the different schemes. The effects of the stochastic methods on the ensemble variability are clearly visible; both of the model uncertainty schemes (multiplicative noise and random parameters) are more spread than the default UM throughout, and the stochastic Plant & Craig scheme is significantly more spread than its deterministic counterpart. The differences between the schemes shown in figure 2.21 will be investigated in detail in the next chapter.

2.5 CONCLUSIONS

An SCM experiment simulating tropical convection, including transitions between suppressed and active phases, has been chosen for studying the high-frequency variability of parameterised sub-grid and non-dynamical processes in the atmosphere. The SCM version of the Met Office Unified Model is used, with the tendencies due to large-scale horizontal and vertical advection prescribed according to observation-derived heat and moisture budgets (based on data from the TOGA-COARE intensive observation campaign over the tropical West Pacific).

Other studies using the same framework have found that errors in the budget-derived large-scale forcings made it difficult to validate models relative to observations. However some departure of the model-state from the one observed in this case is not detrimental to studying sensitivities to high-frequency variability, provided that the model's behaviour remains physically realistic. To examine this point, the meteorology and underlying physics of the case were investigated.

The rainfall variability was found to be primarily forced by large-scale moisture convergence

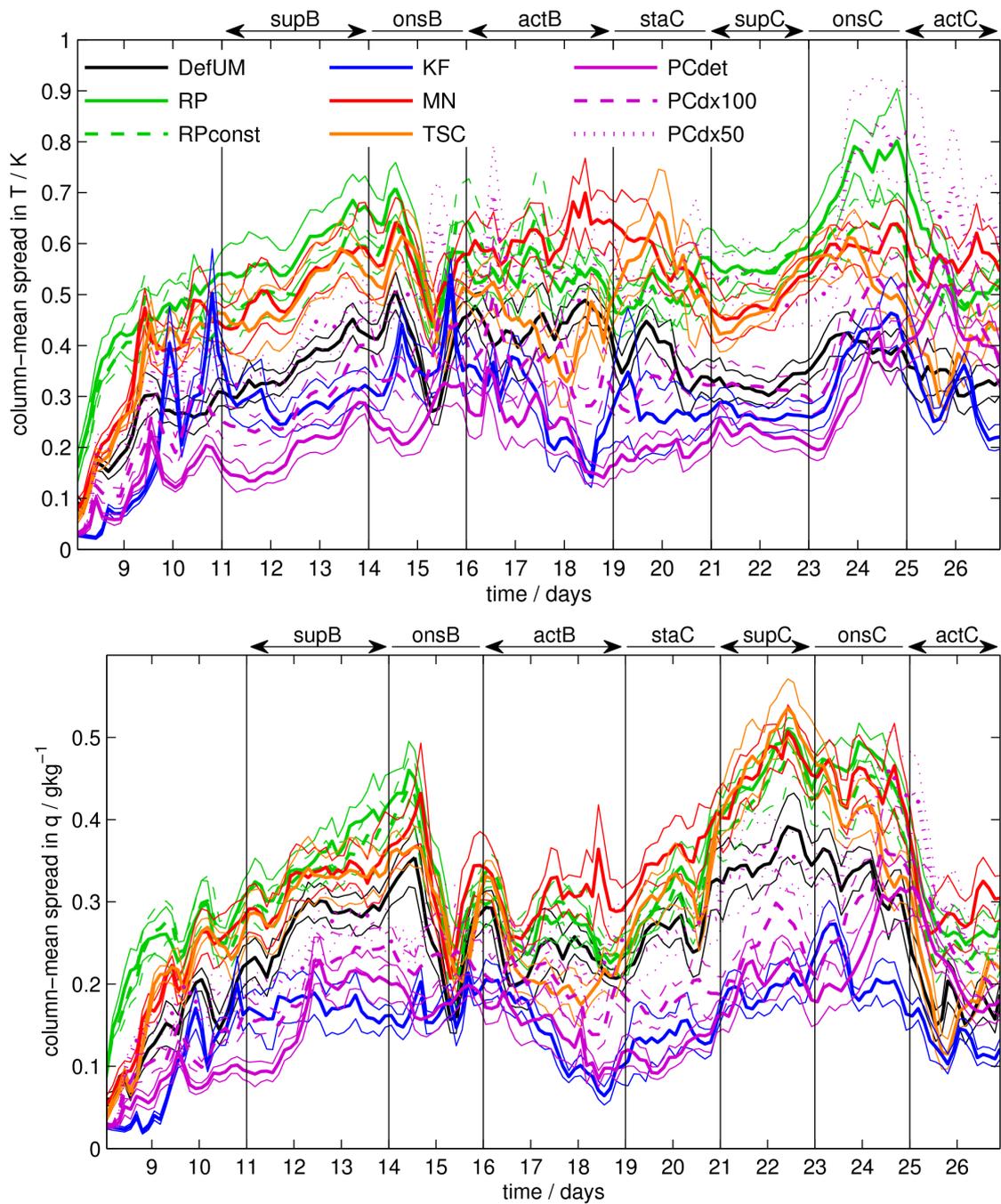


Figure 2.21: Column-mean 3-hourly-mean ensemble spread of (top) temperature T and (bottom) specific humidity q for all the SCM variants described in subsection 2.2.2, for 40-member ensembles including the standard IC perturbations; (black) default UM, (red) multiplicative noise, (solid green) random parameters, (dashed green) constant random parameters, (orange) time-smoothed convection (blue) Kain-Fritsch convection scheme, and (purple) the Plant & Craig convection scheme (solid) in deterministic mode, (dashed) stochastic mode with grid-length 100 km, and (dotted) grid-length 50 km. Thick lines denote the unbiased estimator of ensemble standard deviation, thin lines show its 95% confidence intervals, estimated as described in subsection 2.2.4.

(although in the real world there is a two-way interaction, with moisture convergence promoted by convective heating). The SCM captures the response of convective rainfall to moisture convergence well; due to the effects of entrainment of air from the surrounding environment into convective plumes, deep, rain-bearing convection only occurs when the forcings act to make the mid-troposphere sufficiently moist. Whilst the convection parameterisation is closed on CAPE, it represents this process via its entrainment formulation. The SCM was found to give some systematic drift relative to the observed temperature and specific humidity profiles, but this was not large enough to render the model-states unphysical. Indeed, the SCM seems to produce realistic time-series of rainfall, surface fluxes and cloudiness throughout. It therefore seems reasonable to treat the SCM as a realistic representation of the underlying atmospheric physics in this case, allowing meaningful inferences about the sensitivity of sub-grid and non-dynamical processes to high-frequency variability to be drawn in subsequent results.

In agreement with other studies, the SCM was found to exhibit strongly non-linear sensitivity to its initial conditions. Even when applying very minimal initial condition perturbations, an ensemble of integrations diverged into a set of seemingly independent realisations, with the ensemble spread saturating in just a few days. At a given time, an individual ensemble member may not be representative of the distribution of likely states evolving from a set of almost identical initial states. This supports the point made by Hack & Pedretti (2000) that ensembles of runs need be performed in order to make meaningful comparisons of different parameterisations in an SCM framework.

The saturated level of ensemble spread was found to be independent of the initial condition perturbations used, and instead varies with time as a function of the SCM's variability. An SCM ensemble with an alternative convection parameterisation (the KF scheme described by Kain 2004) was found to give smaller ensemble spreads and diverge much less readily, showing that the SCM's variability and initial-condition sensitivity is highly dependent on the convection scheme used. The KF scheme has a somewhat simpler trigger formulation, with fewer "switches" than the default UM convection scheme. This could explain why it exhibits less strongly non-linear sensitivity.

Further investigation showed that the "on-off" triggering of the convection parameterisation is the primary mechanism driving divergence of ensemble members in the SCM. When the con-

vection triggers, it produces discrete jumps in temperature which dominate the high-frequency variability. The precise timing and magnitude of these jumps has strongly non-linear dependence on the model state, effectively generating substantial uncertainty in the model tendencies at a given time.

Examination of the vertical structure and time-evolution of the spread of the SCM ensemble suggested the spread is closely related to particular uncertainties in the convective behaviour; large spread occurs where the height reached by convective plumes is uncertain, especially near discontinuities such as the melting level or other inversion layers which form during the runs. It was found that once the ensemble saturated, it decorrelated with itself multiple times during the simulation. The timescale for decorrelation of the ensemble members was found to vary enormously with time and height, closely following the pattern of convective heating. These results show that the ensemble spread and decorrelation timescale can be studied as good measures of the amplitude and timescale of convective variability in the SCM.

The inclusion of a simple stochastic multiplicative perturbation method (based on that of Buizza *et al.* 1999, which is designed to account for parameterisation uncertainty in an ensemble forecast) was found to yield consistently larger ensemble spreads than the default deterministic UM SCM. This confirms that the spread of the SCM ensemble can be used to assess the increase in high-frequency variability introduced by the stochastic method. Consistent with its design, the multiplicative perturbation method simply boosts the amplitude of the parameterisation variability, but not the qualitative behaviour of its time-series.

The spread of the stochastic SCM ensemble was found to exhibit almost no sensitivity to the inclusion or not of initial condition perturbations; simply applying different random numbers in the stochastic physics component of each ensemble member gave a rapid divergence of states regardless. This is in contrast to ensemble GCM studies using stochastic parameterisations, but this is unsurprising as the SCM lacks the large-scale dynamical modes which dominate spread-growth in such studies. Such modes are maximally sampled by GCM initial condition perturbations, but only indirectly perturbed by stochastic physics schemes.

The aforementioned relationship between ensemble spread and convective variability illustrates the point that modelling *uncertainties* can be synonymous with the *variability* of param-

eterised processes. For example, in figure 2.16, large uncertainties (as quantified by ensemble spread) are present wherever the parameterised tendencies are highly variable. Because of the noisiness and extreme non-linear sensitivities of the convection parameterisation, its variability between ensemble members is congruent with its variability in time. This lends weight to the paradigm that model *uncertainty* can, at least in part, be accounted for in ensemble forecasts by stochastically representing sources of *variability* associated with parameterised processes.

A range of alternative deterministic and stochastic variants of the UM SCM have been fully described, which include either different deterministic convection parameterisations, stochastic methods based upon those convective formulations, or other deterministic methods relevant to the variability of parameterised processes. When implemented in ensembles with initial condition perturbations, all of these SCM configurations were found to produce different time-series of ensemble spread, characterising their differing variabilities. Each stochastic method produced consistently greater spread than its deterministic counterpart, indicating that the increased variability introduced by those methods is quantifiable from the SCM ensemble statistics.

In the following chapter, comparison of the ensemble spread and decorrelation timescale for all the SCM configurations described should yield useful results regarding the nature and relative importance of the sources of uncertainty / variability they represent.

Single-Column experiments have long been established as a means to study the responses of parameterised processes in isolation from resolved-scale dynamics. This is advantageous, firstly because it allows local physical processes and mechanisms to be identified and studied more easily (the combination of parameterised and resolved feedbacks in a 3D model may be hard to disentangle), and secondly because it is computationally much cheaper than running full GCM integrations, allowing a broader range of model configurations to be explored within the constraints of time, computational resources and disk-space allocations. The results in this chapter show that the high frequency variability of parameterised processes, and their alteration using stochastic methods, can usefully be quantified and analysed in the SCM framework.

The strengths of this SCM approach are also its weaknesses: it can be very helpful to explore the behaviour of parameterisations in a clean arrangement, but the behaviour is not necessarily representative of that in the parent GCM. Therefore, as described at the end of chapter 1, a com-

parison of some of the model configurations described will also be made in a 3D aqua-planet framework in chapter 4.

CHAPTER 3: COMPARISON OF STOCHASTIC PARAMETERISATION APPROACHES IN A SINGLE-COLUMN MODEL

3.1 INTRODUCTION

In the previous chapter 2, a Single-Column Model framework was described and validated for the purpose of investigating the high-frequency variability of parameterised processes, and the sensitivity of those processes to that variability. The simulation, of transitions between suppressed and active convection over the tropical West Pacific, was found to exhibit considerable high-frequency variability. This was primarily associated with the convection parameterisation, and SCM ensemble tests confirmed that the variability (as quantified by the spread of the ensembles) could be altered by changing the convection scheme. In particular, it was shown that its amplitude could be increased by introducing stochastic parameterisations.

Ensembles of runs were performed using the Single-Column version of the Met Office Unified Model, with small initial condition perturbations introduced. A range of different deterministic and stochastic variants of the UM SCM were described in section 2.2.2 of chapter 2, and have been implemented so that a comparison of the SCM's response to the different schemes can be made, for which the results are presented in this chapter.

Some of the model variants include alternative deterministic convection parameterisations; the standard UM bulk scheme based on Gregory & Rowntree (1990), the Kain (2004) individual plume scheme, or the spectral multi-plume deterministic Plant & Craig (2008) scheme. Some include stochastic methods based upon those deterministic configurations; the Random Parameter

method of Bowler *et al.* (2008), the multiplicative perturbation method of Buizza *et al.* (1999), or the stochastic convection scheme of Plant & Craig (2008). And a couple of configurations with other deterministic methods relevant to parameterisation variability are included for comparison; a constant perturbed parameter ensemble method, and a time-smoothed version of the default UM convection scheme.

The purpose of comparing these model configurations in an SCM is to investigate the sensitivity of sub-grid and non-dynamical processes in the atmosphere to differences in the high frequency variability, as described in chapter 1. Firstly, it would be interesting and useful to find out whether the response of the parameterised processes to methods which increase the high frequency variability is similar regardless of the exact method, model-configuration or atmospheric state, or otherwise which of these factors the response is most sensitive to. Secondly, the different stochastic schemes represent different types of uncertainty / variability; information about the relative importance of the corresponding sources of uncertainty / variability can therefore be obtained by comparing the schemes. Thirdly, by comparing stochastic SCMs to their deterministic counterparts, this study aims to ascertain whether sub-grid / non-dynamical processes give any systematic non-linear response to the increase in high-frequency variability, other than the directly forced increase in noise (i.e. changes in mean-state or well-resolved modes). If so, it would be of key interest to elucidate the mechanisms for any such noise-induced effects.

It is also hoped that the results of this comparison will guide the choice of a smaller number of schemes to compare in a 3D GCM framework (an aqua-planet), for which the experimental setup and results will be presented in chapter 4.

Also, this chapter explores a potential criticism of the stochastic methods aimed at representing modelling uncertainties; whilst they aim to sample ambiguities in the formulation of parameterisations, they themselves introduce free parameters (controlling the nature of the added stochastic forcing) whose values are ambiguous. For example, it is not clear *a priori* what amplitude the perturbations should have in a stochastic multiplicative perturbation scheme, such that the range of parameterised tendencies produced matches the genuine uncertainty-ranges in those tendencies.

This point can be examined firstly by comparing the range of model states produced in the

SCM by generic modelling uncertainty schemes to the range of states produced by different deterministic model formulations. If the spread of an SCM ensemble with a generic model uncertainty stochastic scheme (with its default settings as described in the literature) matches the spread of an ensemble of different deterministic schemes, then the range of structural uncertainties included in the latter ensemble is well-sampled by the stochastic scheme. If not, the stochastic scheme's settings should be modified.

Also, the sensitivity of a stochastic SCM's behaviour to the internal parameters and structural details of the included stochastic method can be investigated. Such sensitivity studies were performed as part of this work, but the results of these were considered beyond the scope of this thesis. They are of relevance to the modelling community. For example, results for a stochastic multiplicative perturbation method showed that decorrelating the stochastic perturbations to temperature and moisture gave much greater variability, but did so by violating the conservation of moist static energy during the formation / evaporation of layer cloud. Such decorrelation was in fact applied in the stochastic scheme used operationally in the ECMWF forecast system for some time, but was removed in a subsequent model upgrade due to concerns that the decorrelation of perturbations allowed the model to deviate from the physical attractor (Palmer *et al.* 2009).

The remainder of this chapter is structured as follows: Section 3.2 shows results for a comparison of the ensemble variabilities of the different SCM configurations. Section 3.3 compares the ensemble spread resulting from several stochastic (or otherwise perturbed) methods to the range of model-states produced by different deterministic model configurations. Section 3.4 investigates noise-induced effects on the SCM's ensemble mean-state when the various stochastic methods are applied. Conclusions are drawn in section 3.5.

3.2 COMPARISON OF ENSEMBLE VARIABILITIES

In this section, we investigate changes in the variability of the UM SCM, when various stochastic methods or other convective changes are introduced.

An ensemble was produced for each of the SCM configurations described in section 2.2.2 of chapter 2. In the case of the stochastic convection scheme of Plant & Craig (2008), two separate

ensembles were produced for columns with horizontal scales Δx of 50 km and 100 km. (As explained in subsection 2.2.2.6 of chapter 2, the stochastic fluctuations in that scheme depend on the column size). To ensure a consistent comparison between stochastic physics SCM ensembles and their deterministic analogues, the standard initial condition perturbations described in section 2.2.3 of chapter 2 were included in all ensembles, although as shown therein they make very little difference to the stochastic runs.

It was shown in chapter 2 that the amplitude and timescale of the high frequency variability in the SCM could be quantified well by the ensemble spread and decorrelation timescale respectively. It was found that introducing stochastic methods increased the column-mean spread of the SCM ensemble, but a more rigorous comparison, including consideration of the vertical structure, is given here. Figures 3.1 and 3.2 show vertical profiles of ensemble spread in temperature for each SCM ensemble configuration, time-averaged over each of the suppressed, active and transition phases of convective activity discussed in section 2.3 of chapter 2 (and labelled in figures therein).

It was also shown in section 2.4.2 of chapter 2 that the ensemble decorrelation timescale is very closely related to the convective variability in the SCM. So to gain further insight in to the differences in convective variability between the different ensembles, profiles of ensemble decorrelation timescale are compared. Figures 3.3 and 3.4 show profiles of the decorrelation timescale for temperature.

Equivalent plots for specific humidity (not shown) were also studied, and exhibited similar differences to those presented for temperature in these two figures.

There are distinct differences between the profiles in figures 3.1 and 3.3, which are for configurations using the default UM's convection scheme (default UM, RP, MN, TSC), and figures 3.2 and 3.4, which are for configurations based on the Kain-Fritsch plume model (KF, PC).

The latter grouping exhibits large peaks in ensemble spread in temperature in the upper troposphere and lower stratosphere region, presumably associated with convective overshoots. Such peaks are absent for the first grouping, which tend to have greater spread in the mid-troposphere. As was noted in subsection 2.2.2.5 of chapter 2, the KF scheme explicitly simulates the vertical momentum of the ascending plume and so allows overshoots, whereas the default UM convection

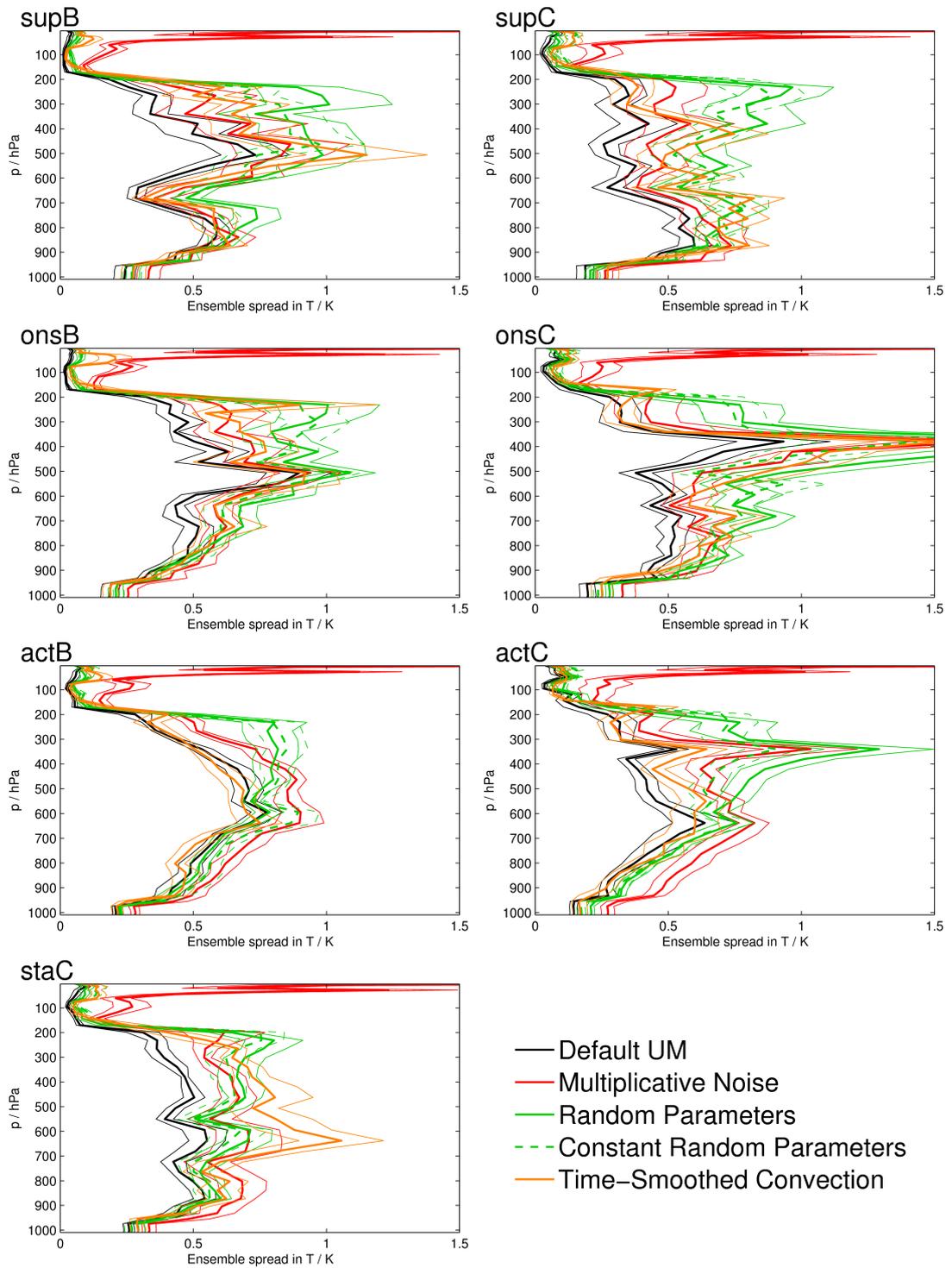


Figure 3.1: Vertical profiles of ensemble spread (standard deviation) in temperature T , averaged over different periods as labelled. Colours denote different SCM ensembles; (black) default UM, (red) MN, (solid green) RP, (dashed green) constant RP, and (orange) TSC. Thick lines denote the unbiased estimator of spread, thin lines denote 95% confidence intervals (estimated by sub-sampling the ensemble, as described in section 2.2.4 of chapter 2).

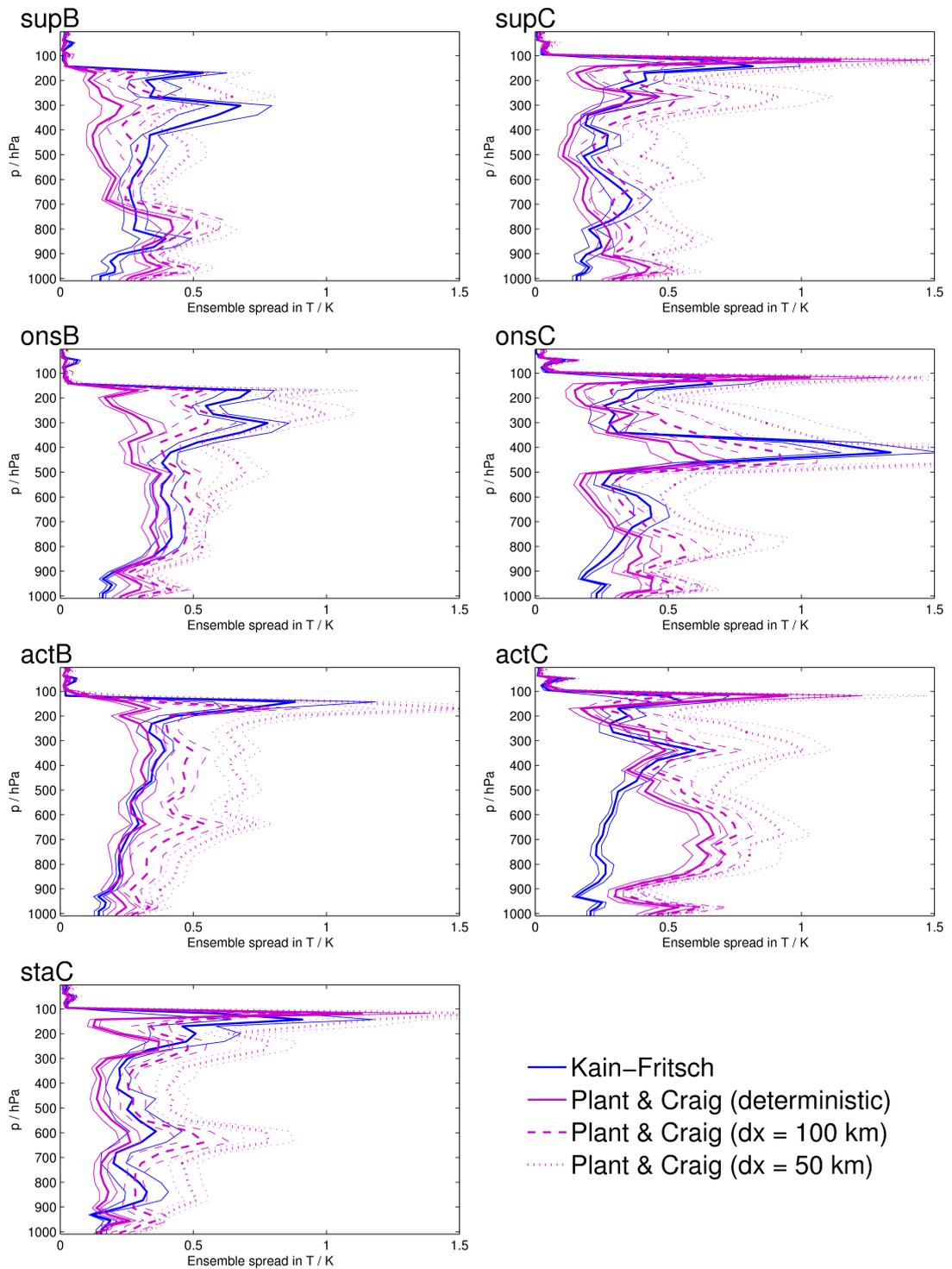


Figure 3.2: As figure 3.1, but for different SCM ensembles; (blue) KF, (solid purple) deterministic PC, (dashed purple) stochastic PC with grid-length 100 km, and (dot-dashed purple) stochastic PC with grid-length 50 km.

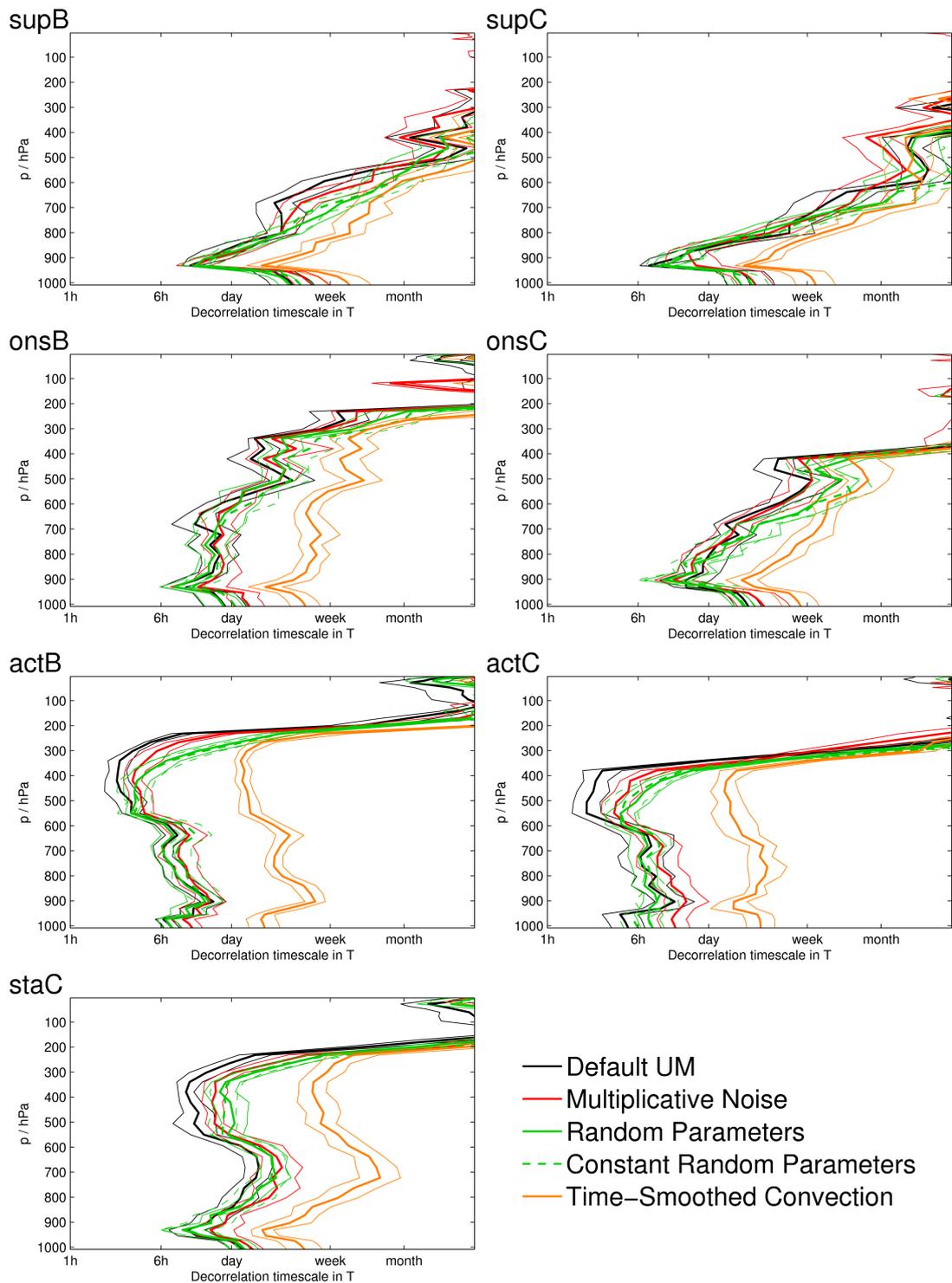


Figure 3.3: As figure 3.1, but for the ensemble decorrelation timescale τ , calculated for temperature as described in section 2.4.2 of chapter 2. The scale on the x-axes is logarithmic.

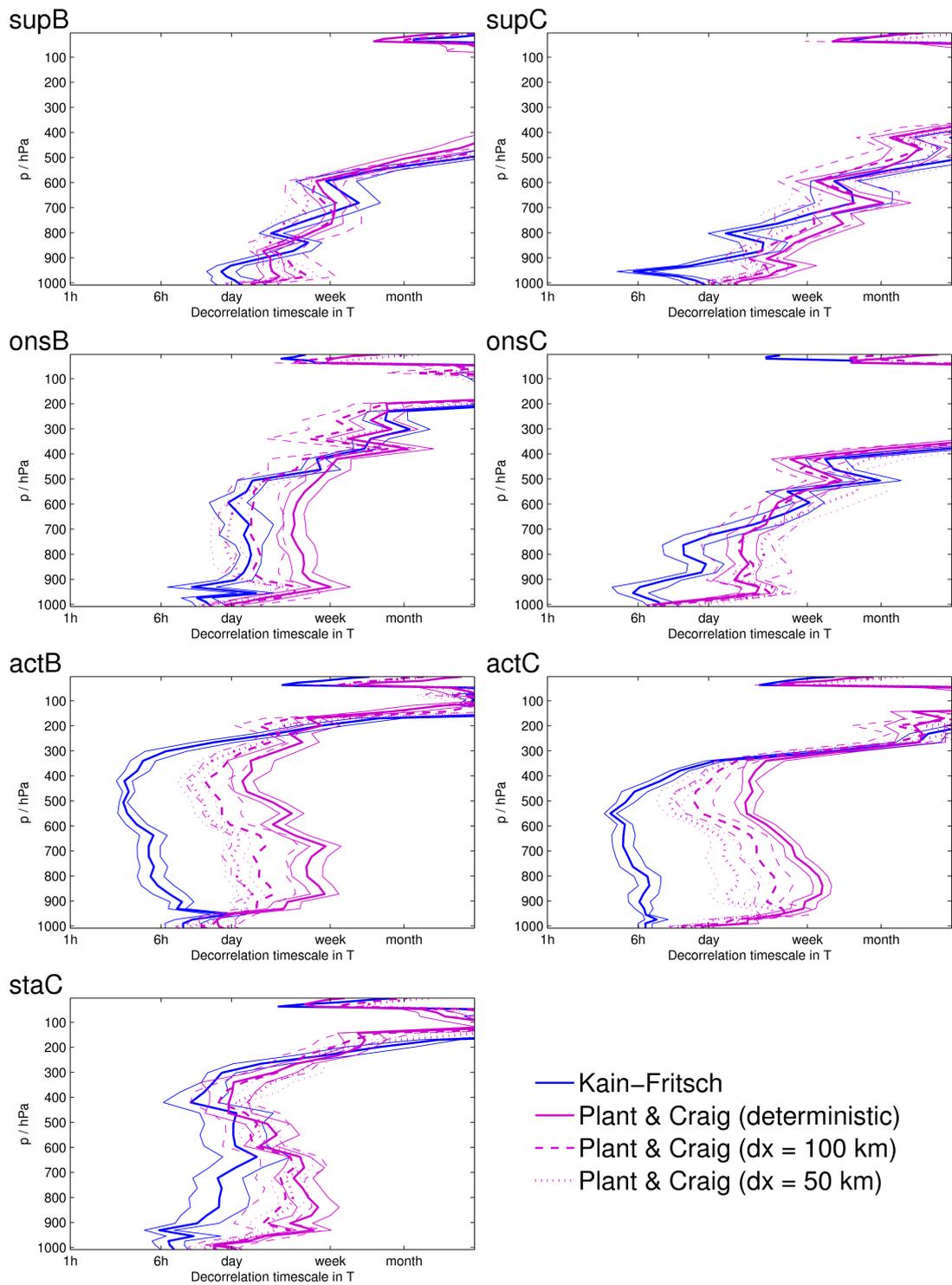


Figure 3.4: As figure 3.2, but for the ensemble decorrelation timescale τ , calculated for temperature as described in section 2.4.2 of chapter 2. The scale on the x-axes is logarithmic.

scheme does not.

The vertical structure of the ensemble spread profile appears to be primarily dependent on the model-state and the convective plume model used, with each stochastic scheme mainly acting to scale-up the spread-profile of the deterministic ensemble which has the same plume model.

There are marked differences between the different phases of convective activity, although again these differ greatly between the two convective plume models. In the ensembles which use the default UM plume, the active phases have peak spread in the mid-troposphere, the suppressed phases have a peak in the lower troposphere, and the onset phases are characterised by a sharp peak where the convection begins to penetrate a mid-level inversion (this was noted in section 2.4.2 of chapter 2). For the KF plume model many of these features are absent, and there is less variation of the spread profile between phases. And in the troposphere, the default UM ensemble generally produces more spread than the Kain-Fritsch ensemble. This confirms that different deterministic convection parameterisations produce rather different amplitudes of variability in the host model. Thus, if the high-frequency variability of a model does have important effects on climate, one should introduce some (stochastic) method to control the high-frequency variability, or at least should investigate the on-off characteristics of the GCM convection parameterisation.

However, the KF scheme does produce very similar variations in ensemble decorrelation timescale to the default UM; both have a rapid increase in timescale with height during suppressed phases (from around 1 day near the surface to greater than a month in the upper troposphere), and timescales of the order a few hours through most of the depth of the troposphere during active phases. So whilst the plume model strongly affects the amplitude of the convective variability, the timescales are less sensitive, at least for these two plume models.

The schemes which represent model uncertainty (multiplicative noise, random parameters and constant random parameters) tend to scale-up the profile of ensemble spread produced by their deterministic analogue (the default UM) in the troposphere, whilst having little systematic effect on the timescales of variability. The RP scheme produces greater increases in spread in the upper troposphere, while the MN scheme produces a more uniform increase in spread throughout the depth of the troposphere.

Note that the MN scheme also produces a large increase in spread in the stratosphere; this is

because it directly perturbs the radiative temperature tendencies, which give large diurnal variations here. The other runs give very little ensemble spread in the stratosphere, because the unperturbed radiative tendencies are actually fairly steady and well-defined. This suggests that applying multiplicative perturbations to the radiation in the stratosphere is not well justified, and the resulting noise is not physically realistic. Indeed, a recent revision to the stochastic multiplicative perturbation method used in the ECMWF ensemble forecast system (described in Palmer *et al.* 2009) includes tapering-off the perturbations in the stratosphere so-as to avoid adding unphysical noise to the radiation. In that revision, the perturbations were also reduced in the boundary layer for numerical stability reasons (note in figure 3.1 that the MN scheme produces greater spreads in the boundary layer than any of the other ensembles using the same convective plume model).

Comparing the deterministic Plant & Craig scheme (a spectral convection scheme) to the KF scheme (single plume, but the same plume model) in figure 3.2, the Plant & Craig scheme in deterministic mode gives ensemble spreads broadly similar to the KF scheme. But it gives less spread than the KF scheme in the upper troposphere during the earlier phases, but notably gives more spread than KF in the lower troposphere at many times. The latter difference is likely a common difference between spectral versus single-plume schemes; the Plant & Craig scheme can simultaneously produce a full spectrum of convective cloud sizes / depths, giving characteristic peaks in spread in the lower, mid and upper troposphere, whereas the KF scheme does not capture the shallower clouds (or the mid-level clouds during actC), and therefore does not reproduce the corresponding spread peaks at these heights.

The deterministic Plant & Craig scheme is designed to behave smoothly and steadily, as it derives the environment profiles and CAPE closure from time-smoothed versions of the grid-point variables, representing the well-resolved state, which is expected to be smoother than the grid-point state (which is affected by numerical “noise”). It is clear in figure 3.4 that the deterministic PC scheme does indeed behave more smoothly than its single-plume counterpart (the KF scheme) wherever convection occurs; the ensemble decorrelation timescale is increased from just a few hours to several days throughout the troposphere during the active phases. However, it does not give any systematic reduction in the amplitude of the variability, as quantified by the ensemble spread (figure 3.2). Indeed, during initial tests of the scheme in the SCM case studied here, it

was found that increasing the time-smoothing of the input profiles did not eliminate convective “on-off” variability; it made the on-off switching somewhat less frequent but equally pronounced. It is unclear whether or not applying spatial smoothing of inputs to the convection scheme (as is done in its 3D implementation) would better reduce this numerical behaviour.

The stochastic form of the Plant & Craig scheme is designed to represent the variability arising from sub-sampling the cumulus ensemble within a finite area, unlike the Multiplicative Noise and Random Parameter schemes which are generic representations of model uncertainty. Like the RP scheme, the stochastic PC scheme tends to scale-up the profile of spread produced by its deterministic mode, with a bias towards the upper troposphere. However, it also significantly reduces the timescale of convective variability at times. As the grid area is reduced, the PC scheme gives a less complete sampling of the ensemble of possible convective clouds, producing greater fluctuations from statistical equilibrium. This reduces the timescale towards that for a single plume (the KF scheme) and increases the amplitude of the variability.

For an area of side $\Delta x = 100\text{km}$, the scheme is certainly more spread than in deterministic mode, but the increase in spread is usually less than that produced by the model-uncertainty schemes. However, with $\Delta x = 50\text{km}$ the scheme yields increases in ensemble spread throughout the troposphere (relative to its deterministic analogue) which are comparable to the increases produced by the model-uncertainty schemes. These results suggest that at scales of around 50 km, local fluctuations about convective equilibrium become as important as the estimates of full parameterisation uncertainty in the MN and RP schemes. This implies that the sub-grid fluctuations dominate the parameterisation uncertainty at this scale, and will exceed those estimates at smaller scales.

The time-smoothed convection scheme increases the timescale of convective variability throughout, consistent with its design. It also produces significant increases in ensemble spread over all the periods except for the active phases. These increases are largely comparable to those produced by the model uncertainty schemes RP and MN. The response of the SCM to the TSC scheme will be explored further in the following sections.

3.3 COMPARISON OF PERTURBED ENSEMBLE SPREADS TO MODEL-UNCERTAINTY

Although the stochastic physics schemes used in this study do produce significant ensemble spread, it remains to determine whether or not the levels of spread are appropriate. For example, it would be useful to know whether the RP and MN scheme's estimates of the magnitude of parameterisation uncertainty match the structural uncertainty between different deterministic schemes. To examine this point, ensembles that are designed to represent model uncertainty can be compared to the range of model states produced by different deterministic structural configurations.

Here, we construct a poor-man's ensemble by combining the 40-member IC ensembles produced by the default UM, the Kain-Fritsch scheme and the deterministic Plant & Craig scheme, each with equal weighting. The spread of this combined 120-member ensemble can be used as a simple measure of the spread of model states associated with model uncertainty. Ensemble percentiles of column-integrated water vapour for the combined ensemble are shown in figure 3.5, along with those for the three constituent ensembles.

The envelope of combined deterministic ensemble states fails to capture the observed evolution of column water vapour (dotted line in figure 3.5) for much of the period. However, it is unclear to what extent this is due to systematic biases in all three of the deterministic SCM configurations, or to errors in the prescribed dynamical forcings used to drive the SCMs (these were discussed in section 2.2.5 of chapter 2). Other SCM and CRM studies using prescribed forcings derived from TOGA-COARE data have largely attributed such discrepancies to errors in the forcing data (e.g. Krueger and Lazarus 1999, Petch *et al.* 2007). The Kain-Fritsch and Plant & Craig schemes (which both use the same plume model, but different closures) produce drier states than the default UM, although well within the range of values seen when comparing various SCMs (Woolnough *et al.* 2007).

The combined deterministic ensemble is significantly more spread than any of the constituent ensembles, which differ in both their means and spreads. This confirms there is a substantial model uncertainty component in the case studied. However, the combined ensemble is also sig-

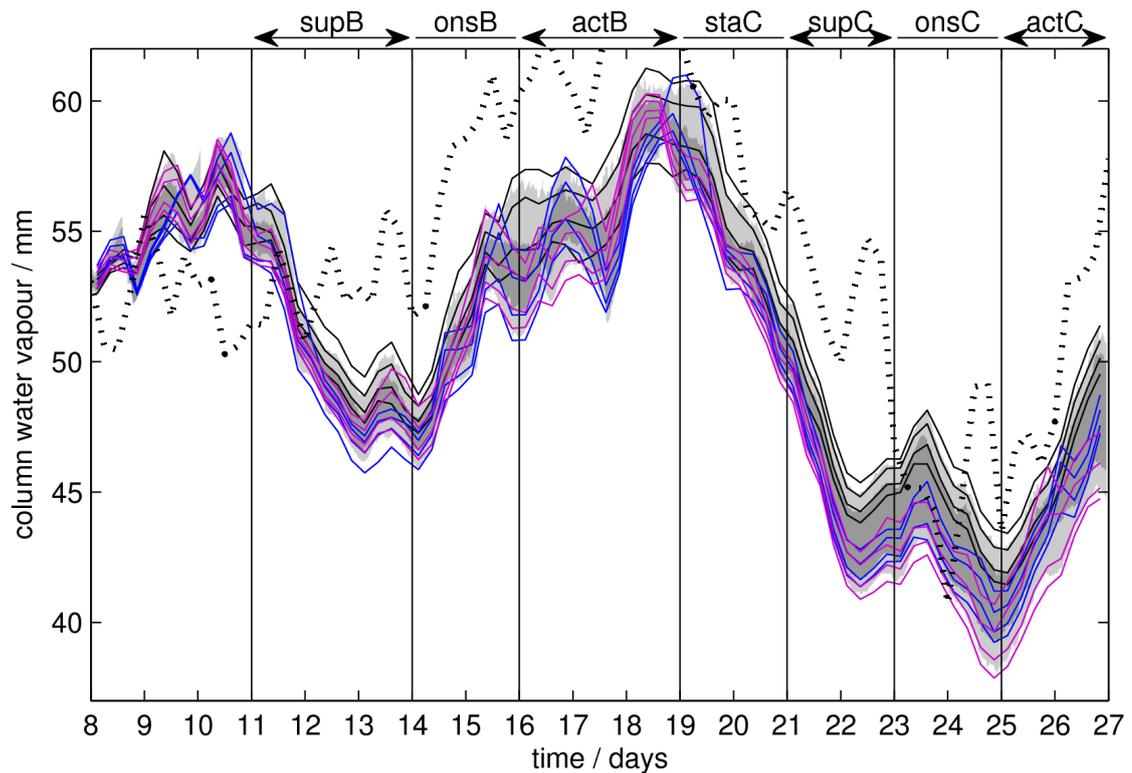


Figure 3.5: 5th, 25th, 75th and 95th ensemble percentiles of column water vapour for three different deterministic SCM ensembles; (black) the default UM, (blue) the KF ensemble, and (purple) the deterministic PC scheme. Shading shows the same percentiles for these three combined in to one ensemble with equal weighting. The dotted black line shows the IFA-mean observed column water vapour.

nificantly more dispersed than the spread of the constituent ensemble mean-states. This shows that the convective variability within each constituent ensemble introduces considerable uncertainty as well. This emphasises the importance of accounting for the uncertainty in the convective realisation in SCMs through the use of ensembles which use initial condition- or stochastic-perturbations; an ensemble of individual SCM runs for different structural model configurations would not capture the full range of associated uncertainties.

Certainly, the representativeness of the combined deterministic ensemble shown here is also questionable, with only three different model configurations represented. This is similar to other “multi-parameterisation” approaches to representing model uncertainty, such as that of Houtekamer *et al.* (1996), who applied only two different convective formulations in their ensemble. Nonetheless, we would suggest that a stochastic scheme which aims to represent model

uncertainty should produce at least comparable levels of spread to the combined poor-man's ensemble in this study. Figure 3.6 shows time-series of several ensemble percentiles of column-integrated water vapour for each of the stochastic schemes and for the constant random parameters scheme and time-smoothed convection scheme, compared to the same percentiles of the combined deterministic ensemble.

It is encouraging to find that the three schemes designed to represent model uncertainty do indeed produce comparable spread in Column Water Vapour to the combined deterministic ensemble. However, these schemes tend simply to broaden the ensemble about the ensemble mean of their deterministic analogue. Thus, they fail to explore regions of phase-space which are accessible to the other deterministic schemes.

The model-uncertainty scheme which looks most promising in this study is the random parameters scheme, in terms of the range of deterministic model states it also covers. As well as producing appropriate ensemble spread, it appears to give a slight systematic drying at times which nudges the distribution of column water vapour towards that of the combined deterministic ensemble. This noise-induced drift will be explored further in section 3.4.

As was noted earlier, the stochastic Plant & Craig scheme produces comparable spread to generic model uncertainty (as quantified by the spread of model uncertainty schemes RP, constant RP and MN, or the combined deterministic ensemble) for a grid-size of 50km. This result for column water vapour backs up those presented for temperature in figures 3.1 and 3.2, suggesting that sub-grid fluctuations about convective statistical equilibrium will dominate parameterisation uncertainty at scales of around 50km or less.

Note that the time-smoothed convection ensemble exhibits a drying trend relative to the other ensembles, drifting into a significantly drier model state than any of the deterministic structural configurations for part of the runs. This is despite it being based on the default UM's convection scheme, which largely gave the moister states in the combined deterministic ensemble. The drift is further explored in section 3.4. As was noted earlier, the TSC scheme also gives an increase in spread relative to the default UM, and this is sometimes comparable to the range of deterministic model states. But it produces a far more disruptive drift in mean-state than the model uncertainty schemes, placing it further from the combined deterministic ensemble.

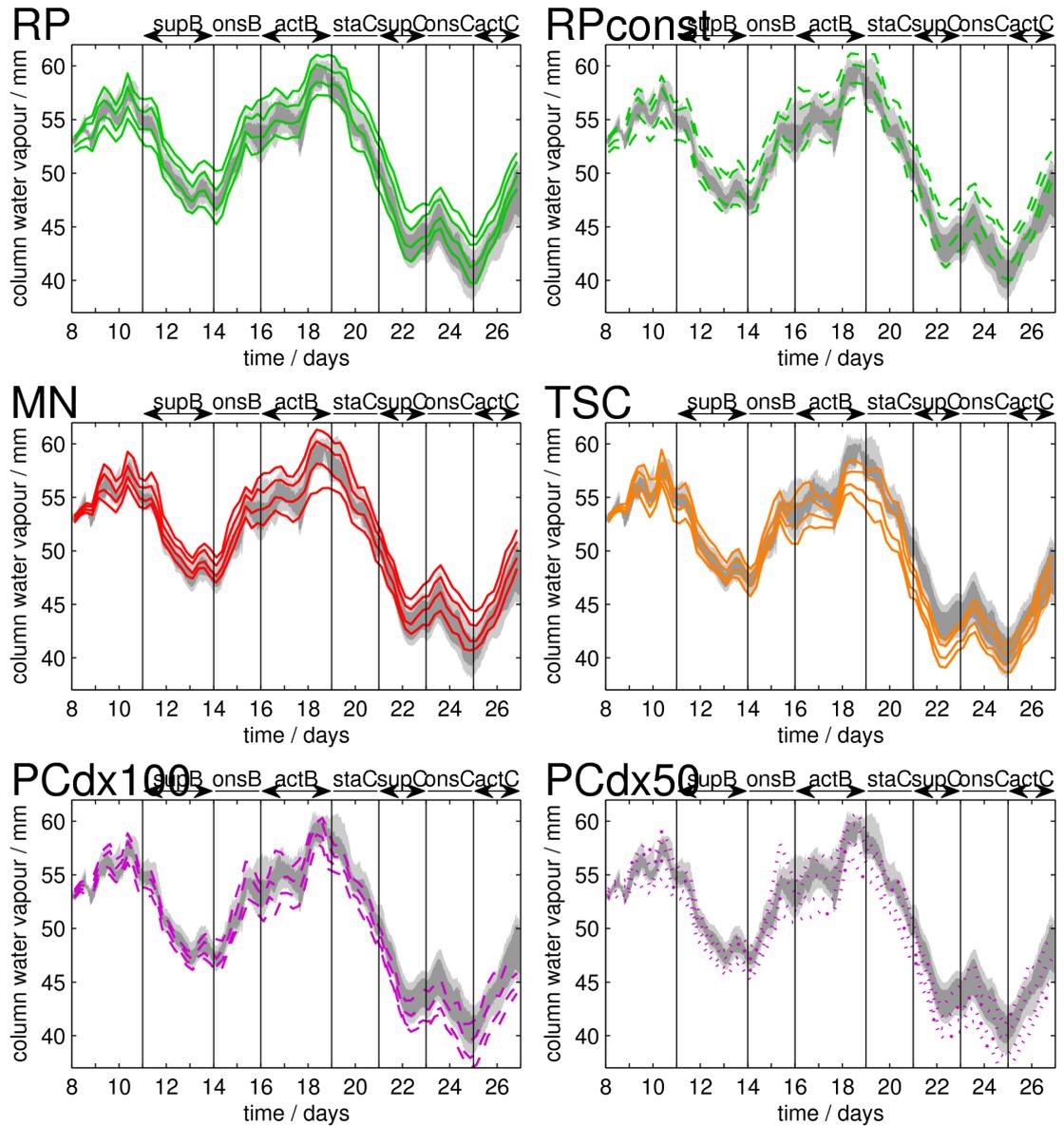


Figure 3.6: 5th, 25th, 75th and 95th ensemble percentiles of column water vapour for six different alternative SCM configurations (solid lines), compared to the same percentiles of the combined deterministic ensemble shown in figure 3.5 (shaded). Colours denote (solid green) RP, (dashed green) constant RP, (red) MN, (orange) TSC, (dashed purple) PC with grid-length 100 km, and (dotted purple) PC with grid-length 50 km.

3.4 COMPARISON OF ENSEMBLE MEAN-STATES

In this section, the sensitivity of the SCM's mean-fields to alteration of the convective variability (through stochastic or other methods) is investigated. First, in subsection 3.4.1 some of the systematic changes in column water vapour (also sometimes referred to as Precipitable Water Content, or PWC) noted in figure 3.6 (section 3.3) are reviewed. The robustness of the subtle moisture drift in the RP scheme is tested using further model runs. In subsection 3.4.2, noise-induced effects on the vertical structure of temperature and moisture are presented and discussed. To further explore the mechanisms behind the differences in ensemble mean moisture profiles between the different SCM configurations, changes in the moisture budget are explored in subsection 3.4.3. Finally, changes in the model cloud fields are investigated in subsection 3.4.4.

3.4.1 NOISE-INDUCED DRIFTS IN COLUMN WATER VAPOUR

Several of the SCM configurations in this study include stochastic physics perturbations but are based on the UM's convection scheme. In figure 3.7 we show the difference in ensemble mean PWC between these configurations and their deterministic analogue, the default UM. Also shown is the difference between the Kain-Fritsch scheme and the default UM. Figure 3.8 shows similar plots for the Plant & Craig stochastic convection scheme.

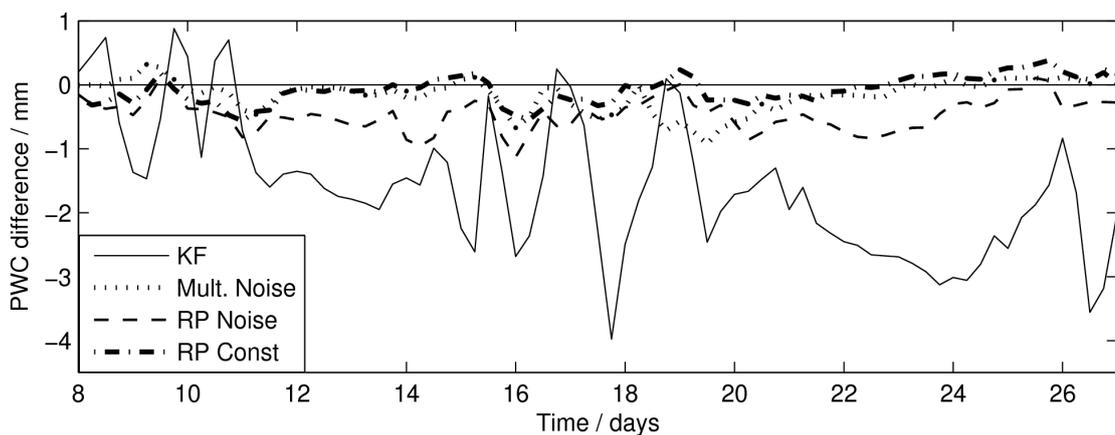


Figure 3.7: Six-hourly mean difference in ensemble mean PWC relative to the default UM, for the multiplicative noise (dotted), time-varying random parameters (dashed), constant random parameters (dash-dotted), and Kain-Fritsch (solid) configurations.

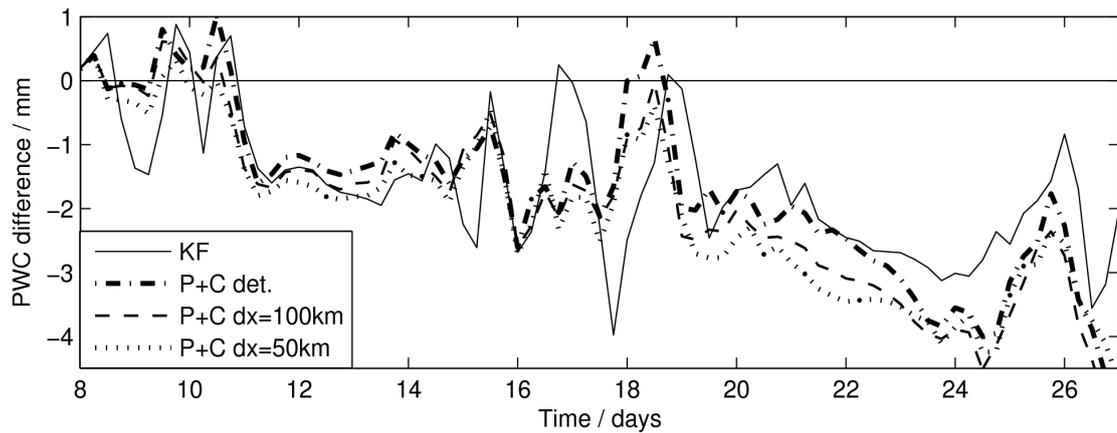


Figure 3.8: As figure 3.7, but for the Plant and Craig scheme with grid-lengths of 50 km (dotted) and 100 km (dashed) and in deterministic mode (dash-dotted), and for the Kain-Fritsch scheme (solid).

The Random Parameters scheme is slightly drier than the default UM during the suppressed and transition phases, by up to nearly 1mm. However the constant random parameters scheme did not produce this deviation despite sampling the same range of values for model parameters (see figure 3.7). This suggests noise induced drift; the stochastic perturbations introduced by the time-variation of the model parameters causes the SCM to explore an asymmetric region of phase-space which is inaccessible to the default UM and constant RP ensembles. The Plant & Craig (2008) stochastic convection scheme also produced a similar drying of the mean-state relative to its deterministic mode. T-tests indicated that these changes were statistically significant at the 5% level.

In terms of ensemble mean PWC, the difference between the two convection parameterisations (default UM and Kain-Fritsch) is several times larger than the changes in mean-state in the stochastic ensembles. Similar remarks apply to other variables (presented in following subsections) and suggest that the ensemble mean fields are far more sensitive to structural differences in the convection scheme than they are to the introduction of stochastic schemes. But the observation that stochastic physics schemes designed to represent model uncertainty or departures from statistical equilibrium can change the mean state of the SCM by even a relatively small amount is nonetheless interesting.

As mentioned in section 2.4.1 of chapter 2, a number of runs were re-integrated to output further diagnostics, but an upgrade to the compiler meant that experiments did not bit-reproduce.

Later, the default UM and RP ensembles were re-integrated again on a different computing platform, this time with a larger ensemble size of 100, to verify the statistical significance of some noise-induced effects. It was at-first assumed that these changes would simply force the ensembles to follow a different sample of realisations from the same attractors. However, some systematic differences between the runs on the different processor standards were later discovered.

After the compiler upgrade, the default UM ensemble had a slightly warmer stratosphere and a cooler boundary layer later in the run (not shown), by of the order 0.1 K (a T-test suggested these slight changes were significant). After the change of platform, the ensemble was found to be significantly cooler and moister during the suppressed phases (by up to 0.5 K / 0.5 gkg⁻¹ at some times / heights, not shown). This was associated with a systematic change in the convective behaviour; previously the shallow convection scheme was inactive (the UM parameterisation always diagnosed deep convection, which seemed reasonable given the deep instability of the profile), whereas on the new platform the shallow convection component triggered during the suppressed phases, yielding large increases in low-level convective cloud, shown in figure 3.9.

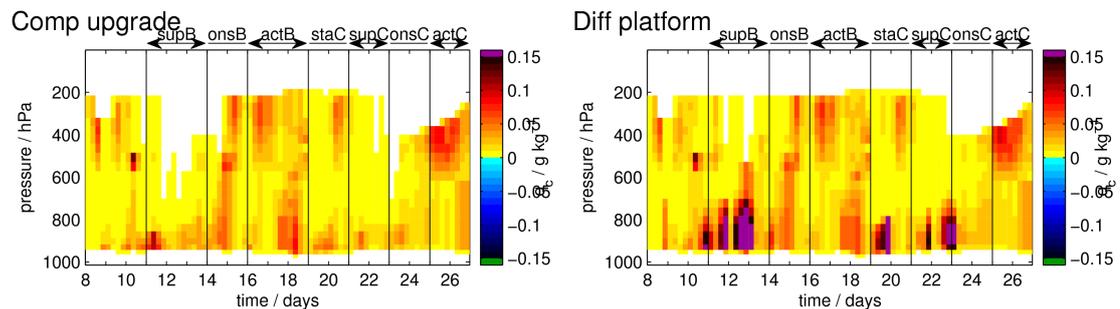


Figure 3.9: Ensemble mean convective cloud water content in the default UM SCM, before (left), and after (right) a change of computing platform.

The inactivity of the shallow convection *scheme* on the old platform doesn't entirely remove the important shallow convection *mode*; instead the deep convection scheme produces plumes which terminate at a shallow depth at times (this is evident in figures 2.16 and 2.18 in section 2.4.2 of chapter 2). But these give smaller cloud amounts than the shallow convection scheme.

The causes of these differences are unclear; no modifications were intentionally made to the model physics. This raises the unsettling possibility that the machine-specific code changes required to run the UM on a different processor standard may interact with the parameterisations

in unforeseen ways. However, it also provides an opportunity to test the robustness of the subtle noise-induced drift in the RP ensemble under changes in model behaviour.

Figure 3.10 shows differences between the RP and default UM column water vapour (similar to figure 3.7) for both runs re-done after the compiler upgrade, and on the alternative new platform. In both cases, a systematic drift of the RP run's mean-state relative to the default UM remains, with no overlap between the confidence intervals of the two mean-states at times during onsB and supC. However, in figure 3.10 the RP scheme is not so clearly separated from the constant RP scheme as it appeared in figure 3.7, as the constant RP ensemble is dryer during suppressed phases in the more recent runs (though this may just be down to sampling uncertainty). However, statistically robust differences are found between the RP and constant RP ensembles if one considers the vertical structure, shown in the following subsection.

3.4.2 NOISE-INDUCED CHANGES IN TEMPERATURE AND HUMIDITY PROFILES

Figures 3.11, 3.12 and 3.13 show time-height plots of differences in specific humidity and temperature between the various perturbed ensembles and their unperturbed analogues. Contours denoting statistical significance are overlaid. These are computed using a point-wise paired-sample T-test for the difference between ensemble means, pairing the members of each ensemble which have the same initial condition perturbations. For the RP and constant RP differences in figure 3.11, the 100-member ensembles performed on the latter computing platform are shown. Plots for the runs on the previous platform (not shown) exhibit the same general patterns, but with more noise and lower significance due to the smaller ensemble size. Figures 3.12 and 3.13 show runs for other schemes on the former platform only.

All of the perturbed or alternative SCM configurations produce complex, state-dependent mean-responses relative to their unperturbed or standard counterparts, with a lot of statistically significant vertical structure not evident in the column water vapour plots. As with column water vapour (figures 3.7 and 3.8), the largest differences are seen between different deterministic structural configurations (KF-DefUM in figure 3.12 and KF-PCdet in figure 3.13), and for the time-smoothed convection scheme (TSC-DefUM in figure 3.12).

Considering the RP ensemble differences in figure 3.11, the SCM's mean response to the RP

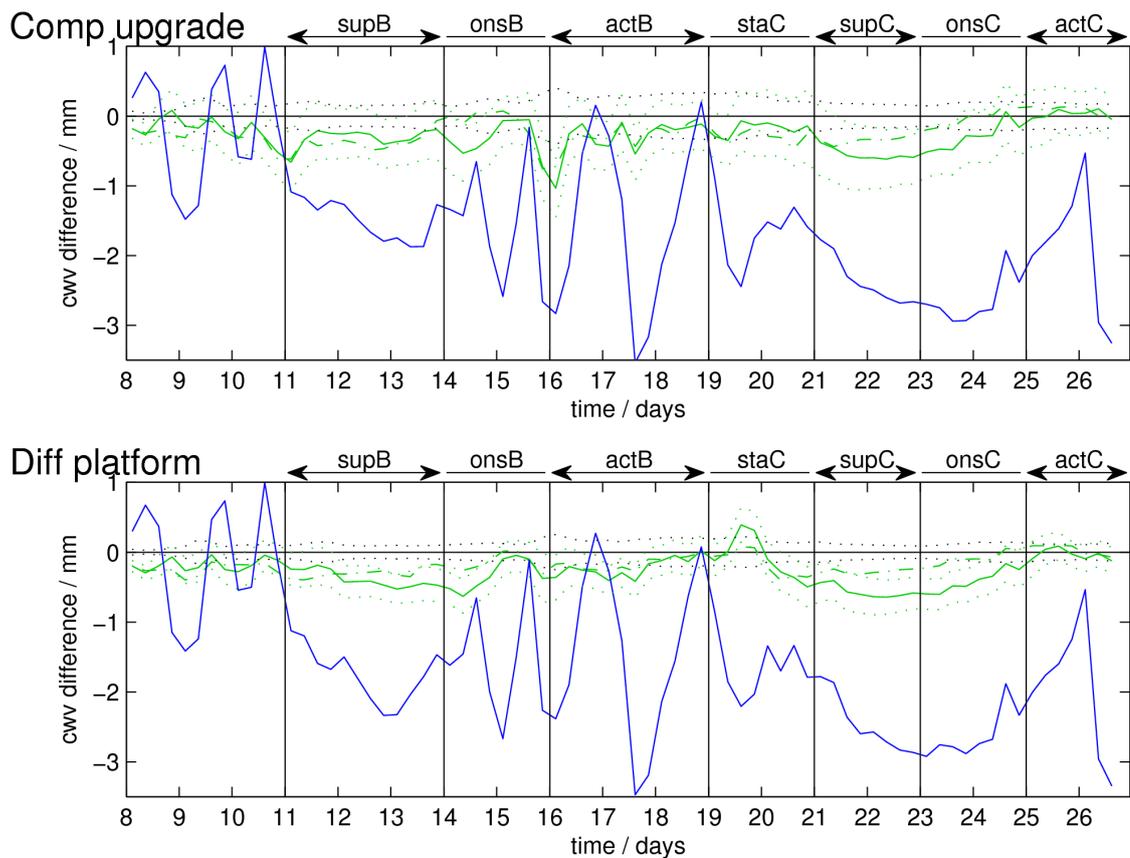


Figure 3.10: Ensemble mean column water vapour differences relative to the default UM as in figure 3.7, but with the RP and default UM simulations rerun; (top) following an upgrade to the compiler, and (bottom) following a shift to another computing platform altogether. Lines show (solid green) the RP scheme, (dot-dashed green) the constant RP scheme, (dotted) estimated 95% confidence intervals for the default UM and RP ensemble means, and (solid blue) the KF scheme for comparison.

scheme (RP-DefUM) can be interpreted as having two components; a deterministic non-linear response to the increased range of parameter values available to the system (RPconst-DefUM), and a stochastic response specific to the time-variation of the parameters (RP-RPconst). The increased range of parameter values alone gives a general moistening of the upper troposphere (accompanied by a warming of around 0.5 K), a drying (and slight warming) of the boundary layer during suppressed and onset phases, and a varying pattern of moistening and drying at different times in the lower free-troposphere. The stochastic response primarily acts to dry parts of the mid-troposphere during suppressed and onset phases by up to 0.5 gkg^{-1} , with a corresponding warming of the atmosphere above inversion layers, and a cooling of the mid-troposphere at other

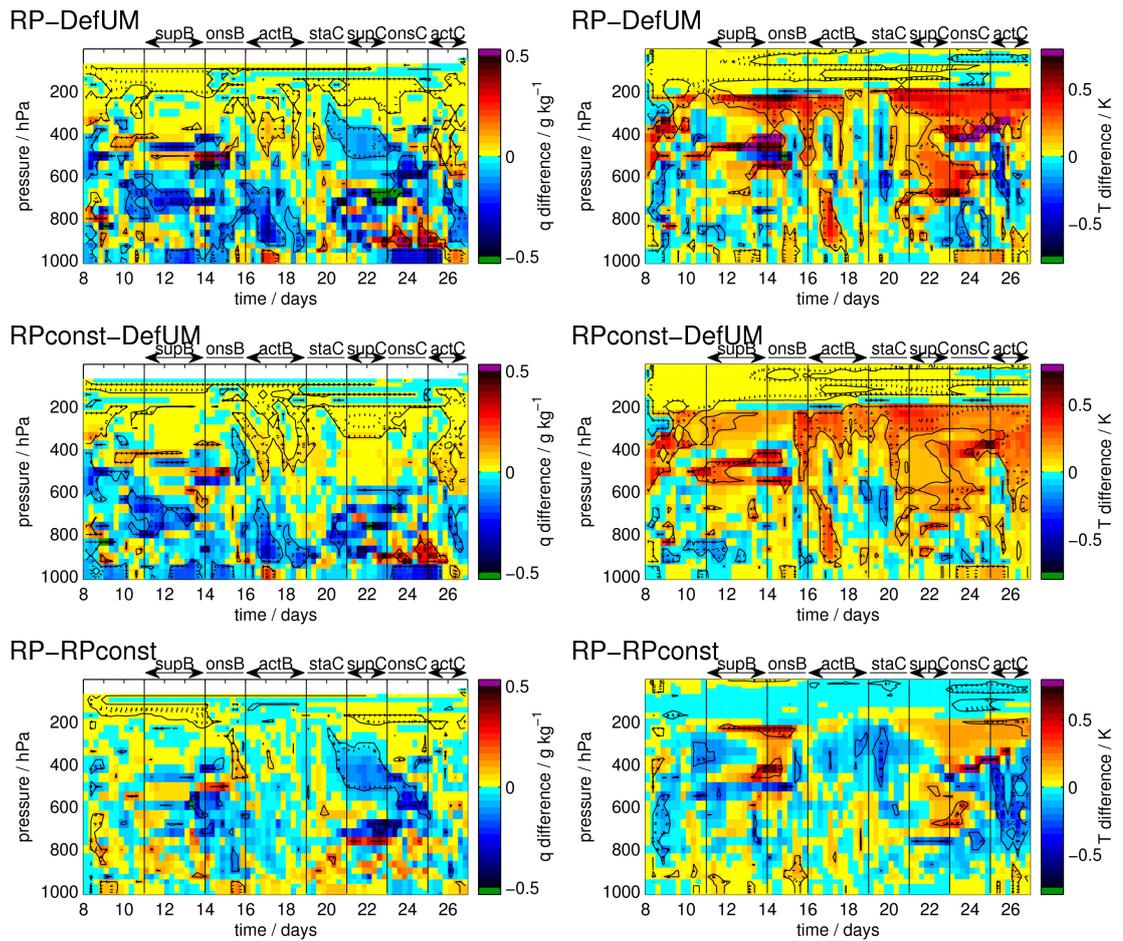


Figure 3.11: Differences in ensemble mean (left) specific humidity q and (right) temperature T , in the 100-member ensembles, for (top) the RP ensemble relative to the default UM, (middle) the constant RP ensemble relative to the default UM, and (bottom) the RP ensemble relative to the constant RP ensemble. Black contours denote statistical significance at the (solid) 5%, and (dotted) 1% levels, using a paired-sample T-test (pairing ensemble members according to their initial conditions).

times. But this noise-induced drift is smaller than the non-linear mean-response to exploring the same parameter ranges deterministically, and far smaller than differences associated with structural uncertainty in the convection parameterisation (e.g. KF-DefUM in figure 3.12).

The stochastic multiplicative perturbations introduced by the MN scheme (middle panels of figure 3.12) give a small but significant cooling of the free troposphere later in the period, and a drying of the column during actB / staC. These differences are far smaller than those which occur when the convection is time-smoothed (bottom panels of figure 3.12). As hypothesised in the literature review on the role of high frequency variability in chapter 1, forcing the convection

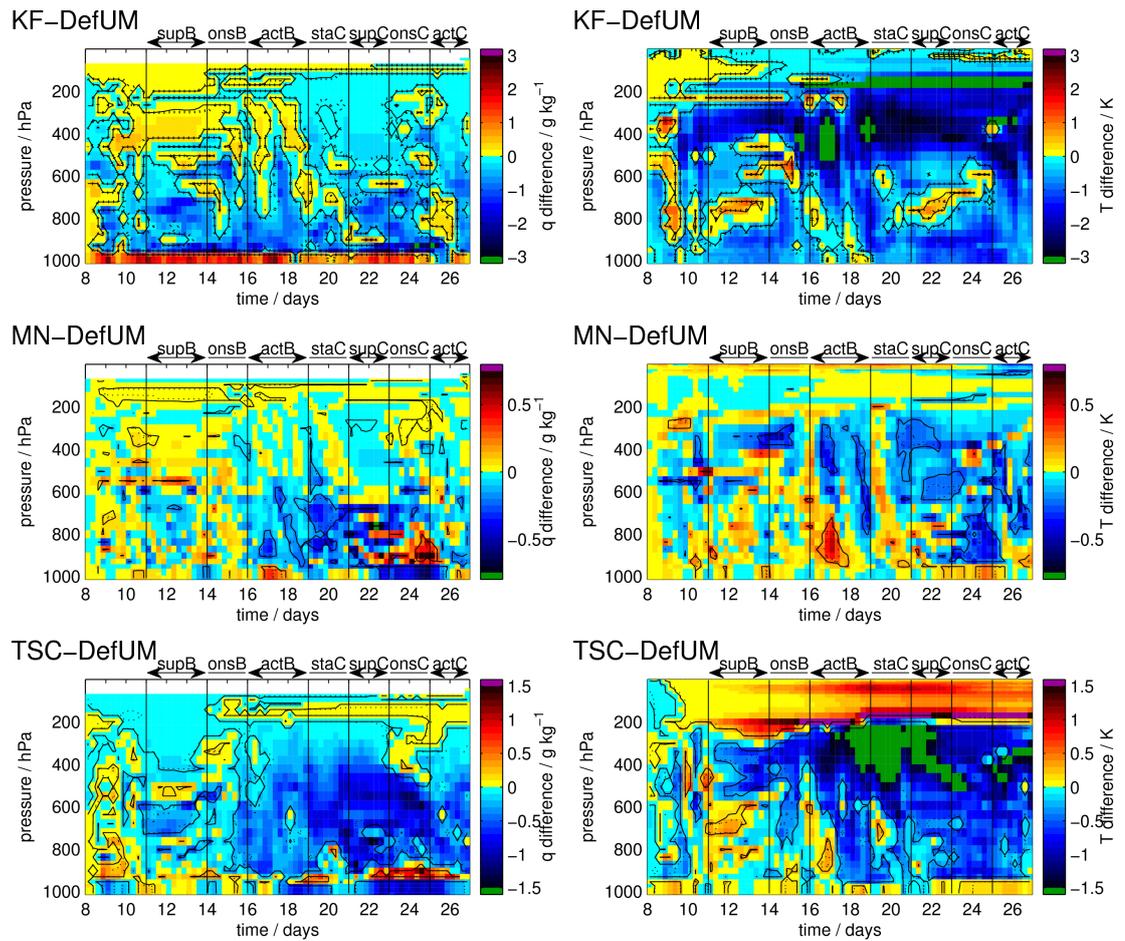


Figure 3.12: As figure 3.11, but for differences of (top) the KF scheme, (middle) the MN stochastic scheme, and (bottom) the TSC scheme, relative to the default UM. Note that the 3 schemes are plotted with different colour scales to show the patterns clearly.

to respond on a longer timescale allows convective instability to build up more before it is removed. As a result, the TSC ensemble drifts into a more unstable profile than the default UM, with temperatures over 1.5 K cooler in the upper troposphere at times. It also becomes significantly drier throughout most of the troposphere; this was clearly seen in figure 3.6 (section 3.3) and is explored further below. The TSC ensemble becomes warmer in the stratosphere; this is presumably a radiatively driven response to the drying of the column below, and also to changes in cloud which are shown in subsection 3.4.4.

The KF scheme (top panels of figure 3.12) also drifts into a state with a cooler upper troposphere than the default UM, but with a different pattern in the lower troposphere. It has a cooler,

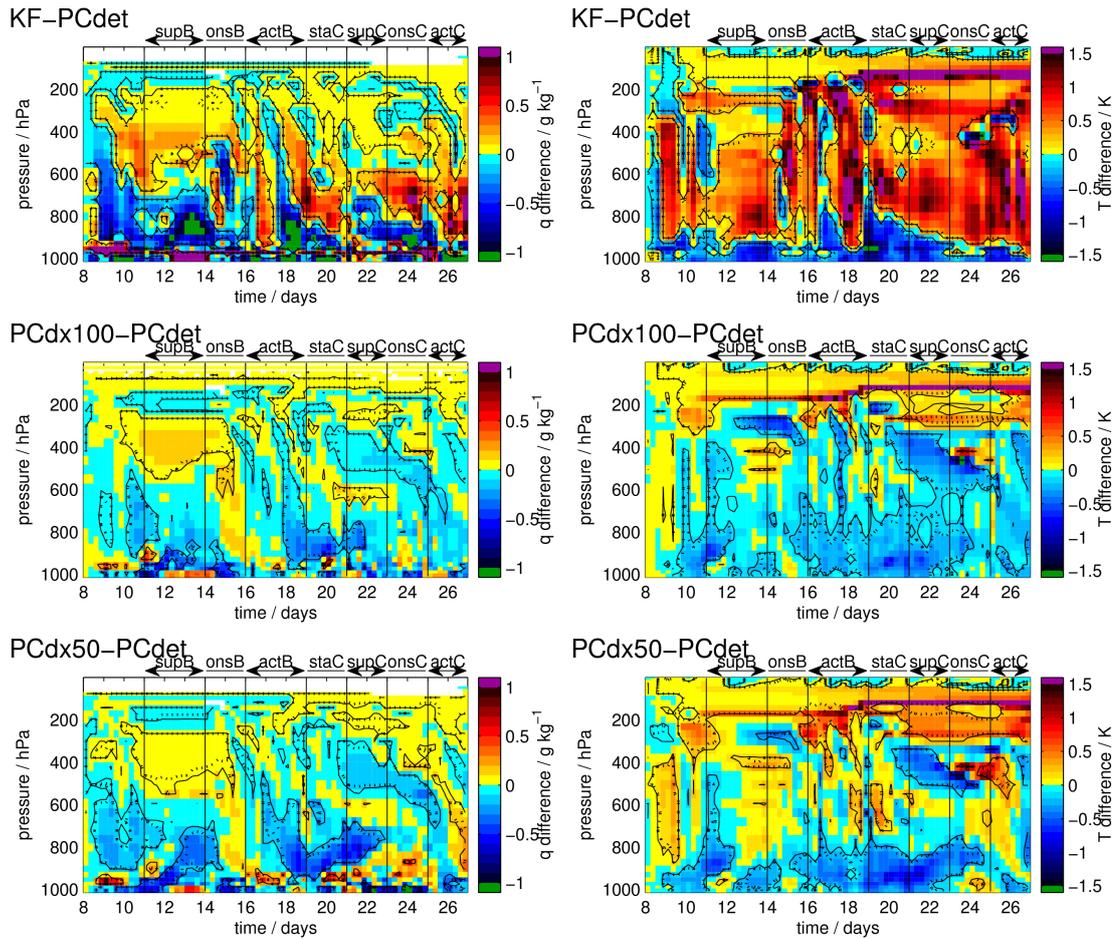


Figure 3.13: As figure 3.11, but for differences of (top) the KF ensemble, and the stochastic PC scheme with grid-sizes of (middle) 100 km and (bottom) 50 km, all relative to the deterministic PC scheme.

moister boundary-layer and dryer, warmer conditions in the lower free troposphere, consistent with its tendency to underestimate the shallower modes of convection in this case (these would transport moisture from the boundary layer to the lower free troposphere, where evaporative cooling would occur). The convective moisture tendencies (not shown) confirm the KF scheme gives less moistening of the lower free troposphere during suppressed phases. It was also noted in section 3.2 that the KF scheme produced less convective variability in the lower troposphere than other schemes (figure 3.2).

The deterministic PC scheme gives a more unstable temperature profile than the KF scheme (top panels of figure 3.13). This is consistent with its spectral design; it simulates the full spectrum of convective cloud sizes, with the total cloud-base mass-flux closed on CAPE. A substantial

fraction of the cloud-base mass-flux is prescribed to occur in small plumes, which do not penetrate high enough to remove deep instability due to their high entrainment rates. This makes the scheme less efficient at removing CAPE than the single-plume KF scheme, which has a fixed entrainment rate sufficient to permit deep plumes for much of the simulation. Investigation of the moisture tendencies from different parameterised processes (not shown) showed that the pattern of humidity differences between KF and PCdet is the residual of large competing changes in the activity of the convection, boundary layer and large-scale precipitation schemes. These changes appear complex and often change sign at different times / heights, making any systematic mechanism hard to pin down, even in this highly simplified framework.

For the Plant & Craig (2008) stochastic convection scheme (figure 3.13), the fluctuations in the simulated sub-grid ensemble of clouds yield a cooling / drying in the lower troposphere for much of the period (relative to the deterministic PC scheme, in which these fluctuations are set to zero) and a warming near the tropopause. These changes follow a very similar pattern for both grid-sizes, with their magnitudes increasing with decreasing grid-size. This is as expected, as larger fluctuations from statistical equilibrium occur for smaller grid areas. However, even at the smaller grid-size of 50km, the changes in mean-state are considerably smaller than the aforementioned differences between the deterministic PC scheme and the KF scheme.

3.4.3 NOISE-INDUCED CHANGES IN THE MOISTURE BUDGET

To investigate the physical mechanisms behind the moisture profile differences shown in figures 3.11, 3.12 and 3.13, the corresponding differences in several components of the accumulated moisture budgets are investigated in this section. Figure 3.14 shows differences in the surface moisture flux and rainfall, figure 3.15 shows the rainfall differences sub-divided in to large-scale and convective components, and figure 3.16 shows differences in the moisture fluxes at 950hPa (sub-divided into terms from the convection and boundary layer parameterisations). The selected height of 950hPa is usually within the boundary layer but near its top. The differences in fluxes at other heights just below the BL top (note shown) were also checked and found to be similar, except for the KF scheme. Therefore, these plots are indicative of differences in fluxes near the boundary layer top, except in the case of the KF scheme, which exhibits greater variations of these fluxes with height near 950hPa than the other schemes.

To give an indication of the relative magnitudes of the moisture budget differences shown in figures 3.14, 3.15 and 3.16, table 3.1 shows the accumulated net 19-day totals for the moisture terms in the default UM ensemble. Time-series of the moisture budget were also presented in section 2.3 of chapter 2.

Additionally, for comparison with the surface moisture fluxes, time-series of ensemble mean near-surface Relative Humidities are shown in figure 3.17. And for comparison with the 950hPa moisture fluxes, the vertical humidity gradients at the same height are shown in figure 3.18.

Table 3.1: Total 19-day accumulated ensemble-mean moisture budget terms in the default UM / mm. The large-scale moisture convergence is prescribed and so is the same for all the runs, whilst other terms differ.

LS moisture convergence	Rainfall		Surface moisture flux	Moisture flux at 950 hPa	
	LS-precip	Convection		BL scheme	Convection
91	1.7	161	67	28	41

Comparing figure 3.14 with figures 3.11, 3.12 and 3.13, the rainfall and surface flux changes underlying the profile changes are apparent.

The constant RP ensemble gives slightly more rainfall than the default UM just before the start of each suppressed phase, slightly less rain at the end of actB, and slightly greater surface moisture flux during the onset phases. The additional noise-induced drying of the mid-troposphere in the stochastic RP ensemble is driven by slight further increases in rainfall during the suppressed phases, and a reduction in the surface moisture flux relative to the constant RP ensemble. Whilst it is unclear in figure 3.14 whether the differences between the RP and constant RP ensembles are significant (their confidence intervals overlap), significant differences are apparent if the convective and large-scale components of the rainfall are considered (top panel of figure 3.15); the stochastic RP scheme gives an increase in rainfall from the large-scale precipitation scheme, whereas the constant RP scheme does not. Whilst this 0.6mm increase is miniscule compared to the total precipitation, it represents a 35% increase in large-scale precipitation, which is largely inactive during these runs (see table 3.1).

The MN scheme gives slight transient deviations in mean rainfall (e.g. increased rains around the 19th give a brief drying of the column just afterwards in figure 3.12), but there are no obvious

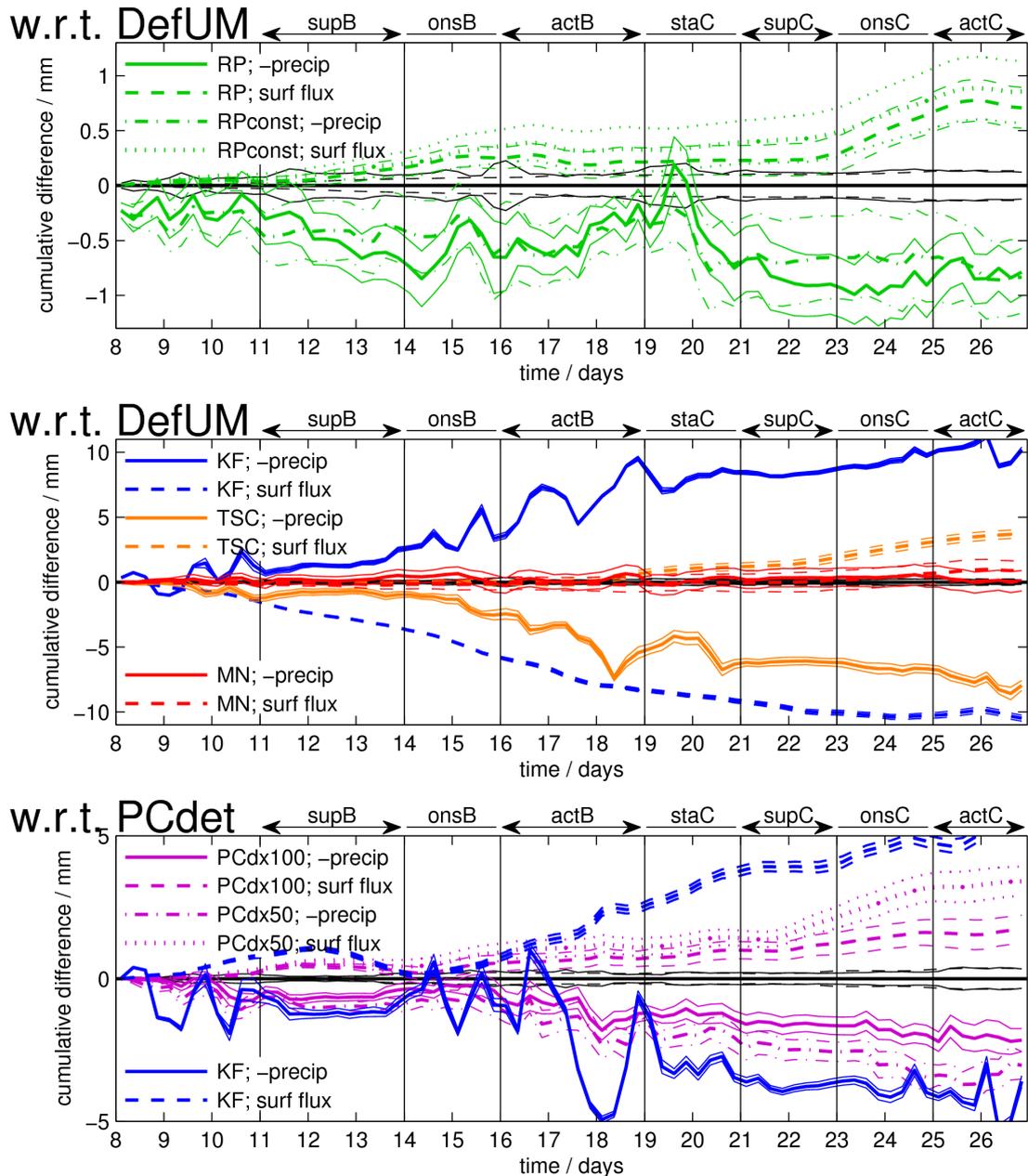


Figure 3.14: Ensemble mean differences in cumulative moisture budget components relative to (top and middle) the default UM and (bottom) the deterministic PC scheme; (dashed, dotted) surface moisture flux, and (solid, dot-dashed) rainfall. Colours denote (green) RP / constant RP, (blue) KF, (red) MN, (orange) TSC, and (purple) PC as labelled. Rainfall has been inverted to give a consistent sign convention for the moisture budget. Thick lines show ensemble means, thin lines show their estimated 95% confidence intervals. Note that each panel has a different y-axis scale, so that the differences can be clearly seen despite their differing magnitudes.

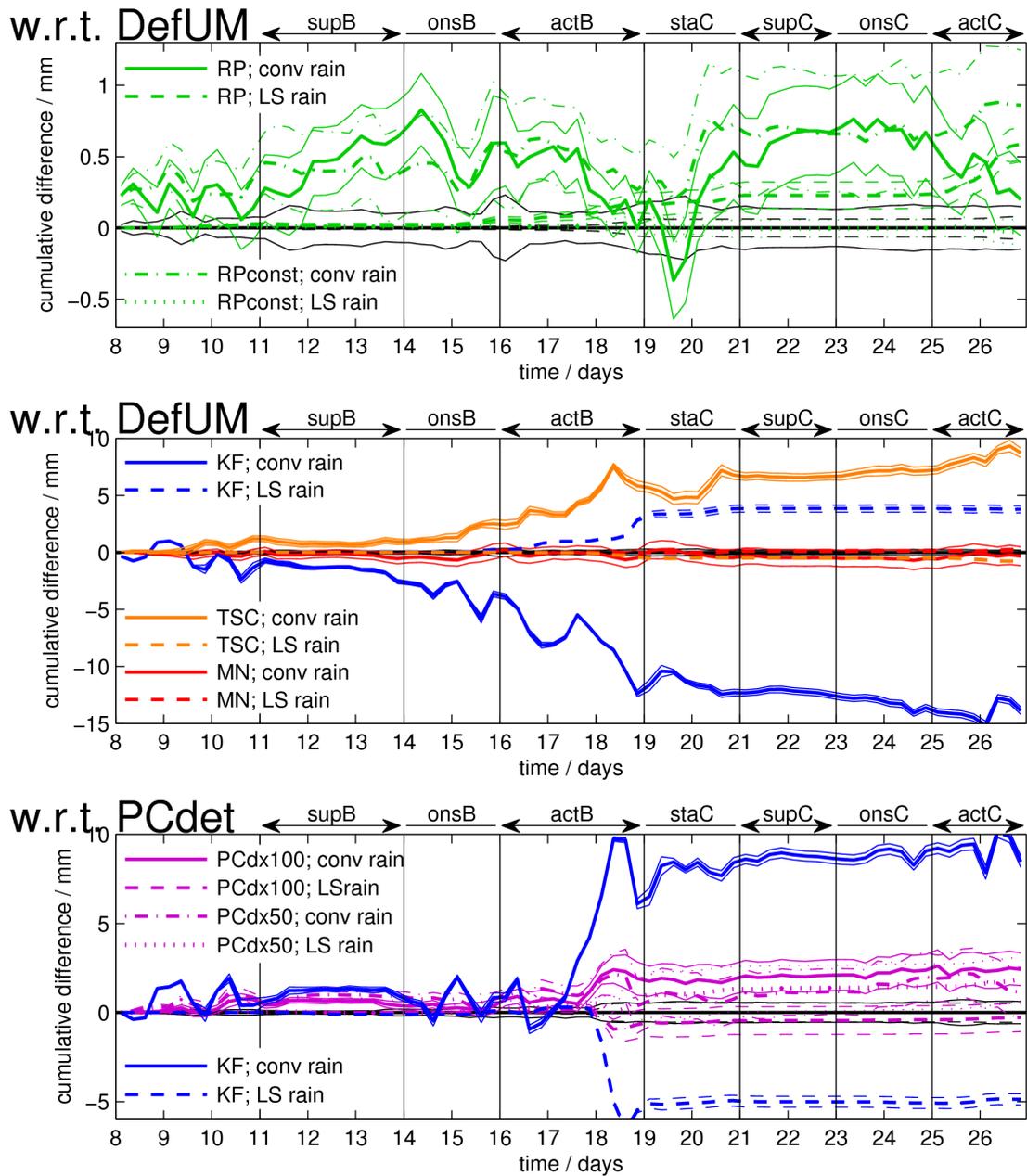


Figure 3.15: As figure 3.14, but for differences in accumulated rainfall from different model components; (dashed, dotted) from the large-scale precipitation parameterisation, and (solid, dot-dashed) from the convection parameterisation. Note that in this figure the rainfall has not been inverted as it was in figure 3.14. As in that figure, note that each panel has a different y-axis range.

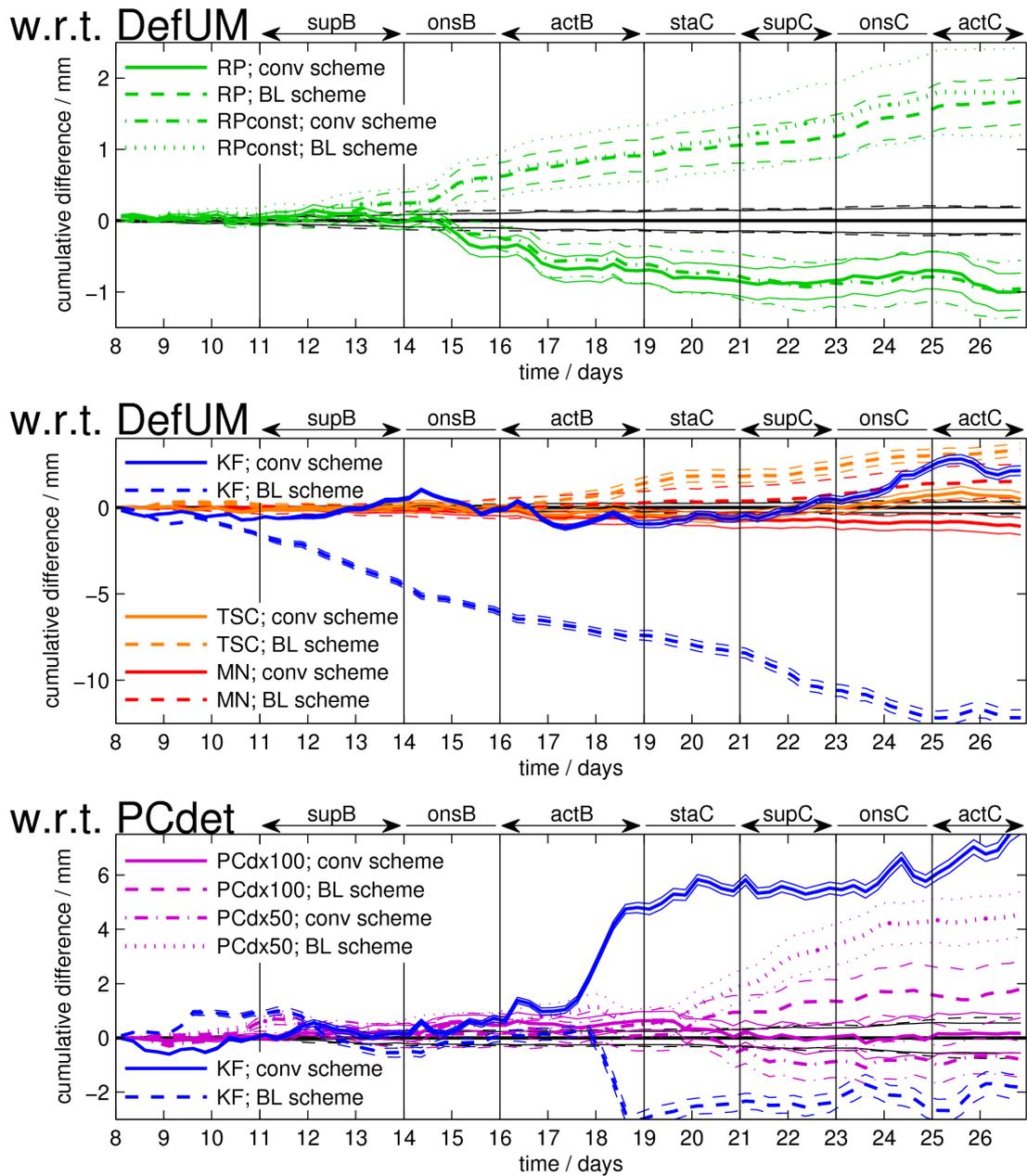


Figure 3.16: As figure 3.14, but for accumulated vertical moisture fluxes at 950hPa, due to (dashed, dotted) the boundary layer parameterisation, and (solid, dot-dashed) the convection parameterisation. As in figures 3.14 and 3.15, note that each panel has a different y-axis range.

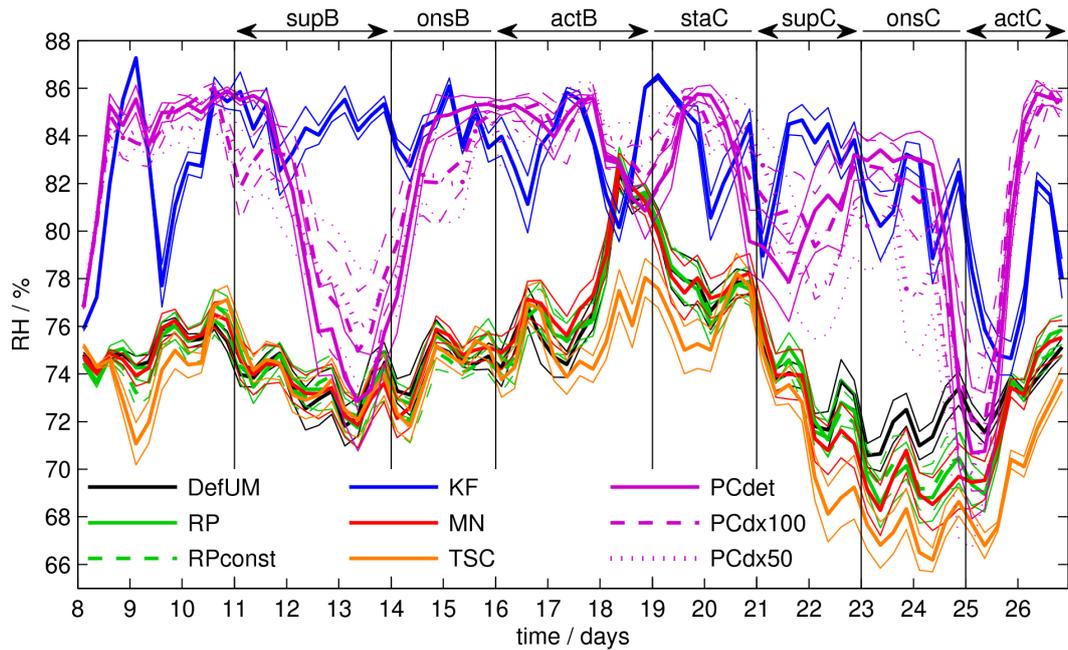


Figure 3.17: Ensemble mean Relative Humidity on the lowest model level (corresponding to a fixed pressure of 1009.4hPa; surface pressure is fixed at 1011.7hPa), for all the SCM ensembles; (black) DefUM, (solid green) RP, (dashed green) constant RP, (red) MN, (orange) TSC, (blue) KF, and (purple) the PC scheme in (solid) deterministic mode, and in stochastic mode with grid-lengths (dashed) 100 km and (dotted) 50 km. Thick lines show the best estimate of the ensemble mean, thin lines show its estimated 95% confidence intervals.

systematic changes. As in the RP ensemble, there is a slight increase in surface moisture flux later in the run. Note that the MN scheme doesn't quite form a closed moisture budget as in the other runs. This is because (following Buizza *et al.* 1999) the total specific humidity tendencies are multiplicatively perturbed, but the condensate tendencies are not. The perturbation scalings have been applied to the rainfall and surface flux diagnostics to make them consistent with the humidity tendencies, but inconsistencies still arise where layer cloud condensate plays a role in the moisture budget.

The strong drying of the troposphere when the time-smoothed convection scheme is applied (figure 3.12) is driven by a significant increase in rainfall, entirely from the convection parameterisation. This is consistent with the more vigorous convection expected, since the temperature profile in the TSC ensemble drifts in to a more unstable state than the default UM. Whilst it was shown in section 2.3 of chapter 2 that the primary factor controlling the rainfall variability in

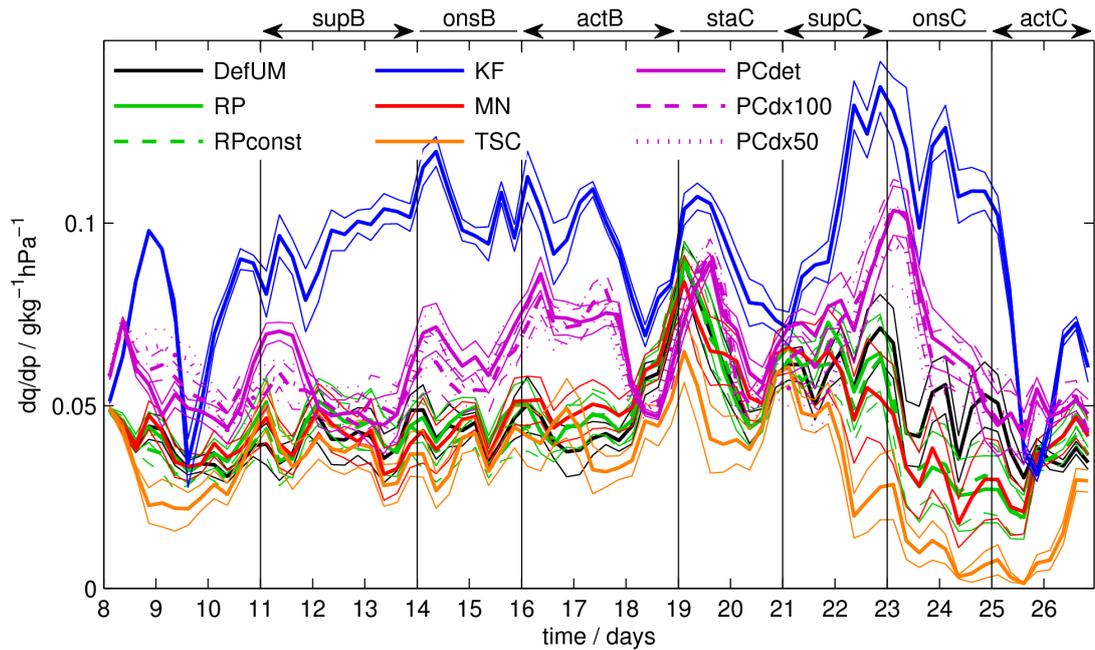


Figure 3.18: As figure 3.17, but for the ensemble mean vertical gradient of specific humidity at 950hPa (corresponding to the height for which vertical moisture fluxes are shown in figure 3.16).

this case is not CAPE but moistening of the troposphere by large-scale convergence, it would be surprising if the systematic increase in profile instability present had no effect.

It is surprising that in the TSC ensemble the upper troposphere remains cooler than in the default UM despite an increase in rainfall (which would give additional latent heating). This is at least partly because the TSC scheme has the effect of delaying most of the convective heating to occur in the couple of hours after each time-step when convection triggers. The greater the rate of cooling by the large-scale forcings and radiation, the greater the temperature difference associated with this lag in the convective heating response. However, the heating-lag cannot explain why tropospheric temperatures don't recover after each convectively active period. Also, the tropospheric drying due to convective rain-out should experience the same lag, but this effect is out-weighted by the increased rainfall. Radiative responses to changes in water vapour and cloud also contribute significantly to the temperature difference pattern.

The KF scheme gives less rainfall and surface moisture flux, by about 10mm over the 19 day simulations. There is a 15 mm decrease in convective rainfall partially offset by a more than 3-fold increase in large-scale precipitation (figure 3.15). Comparing with the default UM moisture

budget in table 3.1, note that the surface moisture flux is reduced by 15%. This would likely give drastic changes in the hydrological cycle if the KF scheme were implemented in a climate simulation. Considering the atmospheric moisture budget, the changes in rainfall and surface fluxes cancel each-other by the end of the run, but the rainfall deficit lags the surface flux deficit at times and exhibits more variations, yielding the pattern of free troposphere drying seen in the top panel of figure 3.12. The cooling of the upper troposphere in the KF ensemble is consistent with the decrease in latent heating associated with the reduction in rainfall.

The deterministic PC scheme gives even less rain and surface moisture fluxes than the KF scheme; it gives almost 10mm less convective rainfall during actB, but this is partially offset as the PC scheme gives twice as much large-scale precipitation as KF. As with KF-DefUM, the reduction in latent heating associated with the reduction in rainfall is consistent with the temperature difference in the troposphere, whilst the effects on the moisture budget broadly cancel, with time-variations of the rainfall deficit relative to the fairly steady surface flux deficit yielding periods of net moistening and drying at different times. The reduction in rainfall / cooling of the free troposphere in deterministic PC relative to KF is consistent with the earlier hypothesis that the PC scheme removes CAPE less efficiently due to its spectral closure.

Considering the stochastic PC scheme, the convective fluctuations give substantial increases in mean rainfall relative to deterministic mode, especially during actB and onsC. These increases are much larger than those seen in the other stochastic schemes, and are largely matched by similar increases in surface moisture flux. Note in figures 3.13 and 3.14 that the larger the stochastic fluctuations (the smaller the grid-size), the more the PC scheme drifts towards the more stable temperature profile and higher rain-rates seen for the KF scheme. This suggests that the fluctuations in the convective ensemble allow the convection to remove CAPE more efficiently, as in the single-plume KF closure.

Comparing the differences in surface moisture flux between all the runs (figure 3.14) with the corresponding differences in near-surface relative humidity (figure 3.17), the surface moisture flux changes appear consistent with simple responses to the humidity changes near the surface. For example, all of the stochastic schemes, the constant RP scheme and the TSC scheme are dryer near the surface during onsC than their unperturbed counterparts, and exhibit corresponding increases in surface flux during that period.

Considering the differences in 950hPa moisture fluxes (figure 3.16), notably all the ensembles which are in some way perturbed (the stochastic schemes, the constant RP scheme and the TSC scheme) give significant increases in the flux produced by the boundary layer parameterisation, relative to their unperturbed counterparts. Comparing the 950hPa boundary layer scheme fluxes to the vertical humidity gradients at the same height (figure 3.18), the times when the perturbed ensembles give the greatest increases in 950hPa flux largely correspond to times when the vertical humidity gradients are smaller in the perturbed runs than in their unperturbed counterparts (e.g. during onsC in the RP, RPconst and MN ensembles relative to DefUM, also days 18–20 for TSC, and days 20, 22 and 24 in the stochastic PC ensembles relative to deterministic mode). At these times the perturbed runs have anomalously dry boundary layers with anomalously moist air just above (see left-hand panels of figures 3.11, 3.12 and 3.13). This is not consistent with a simple response of the 950hPa fluxes to the differences in humidity gradients (one would expect a reduced gradient to yield a reduced net flux). Rather, it suggests the increased 950hPa fluxes are forcing the changes in the moisture profiles. The surface fluxes then increase in response to the drying of the boundary layer, which is forced by the 950hPa flux increases.

In the MN, RP and constant RP ensembles, there are also compensating decreases in the 950hPa moisture fluxes from the convection scheme at times. These changes, combined with the increased surface moisture fluxes, act to balance the boundary layer's moisture budget in the long-run. However, these convective flux reductions only occur at certain convectively active times (e.g. days 17 and 26 in figure 3.16), yielding a characteristic residual time-series of net flux differences between the MN, RP and constant RP runs and the default UM. As a result, these schemes all give similar lower-troposphere profile responses (figures 3.11 and 3.12) with a drying / warming of the boundary-layer and moistening / cooling of the free troposphere just above during onsC, and the reverse of this pattern around day 17.

The RP, constant RP, MN and stochastic PC with grid-length 100 km ensembles all give 5–10% increases in 950hPa boundary layer scheme fluxes, whilst the TSC and stochastic PC with grid-length 50 km runs give larger increases. This response is the same in the RP and constant RP runs, so it is not specifically a noise-induced effect. Rather, it appears the boundary layer scheme is generically affected by an increase in the range of states available to it, whether through stochastic parameterisations or other methods which increase the ensemble spread. This

is likely associated with the strongly non-linear behaviour of the boundary layer top, but more detailed analysis of the boundary layer parameterisation is required in order to elucidate the precise mechanism.

3.4.4 NOISE-INDUCED CHANGES IN CLOUD

Whilst the stochastic schemes studied have been found to produce some drift in the mean temperature and moisture in the SCM, these changes are subtle and generally much smaller than the differences between different deterministic configurations. We might expect the stochastic schemes to have a greater effect on the model cloud amounts, as these are generally more sensitive to small changes in the parameterised physics. Figures 3.19 and 3.20 show differences in convective cloud and layer cloud condensate between the stochastic and deterministic runs.

Interestingly, the RP, constant RP and MN schemes give similar patterns of difference in convective cloud relative to the default UM, with significant decreases in the upper troposphere at various times. Comparing with the default UM ensemble mean convective cloud in figure 3.9, the changes are fractionally quite large, but highly state-dependent. Many of the times of reduced upper-level convective cloud correspond to times when the convective moisture flux at 950hPa (figure 3.16) is reduced (e.g. near days 15, 17 and 26), suggesting an overall reduction in deep convection at those times in the RP, constant RP and MN ensembles.

Reassuringly, the upper-level convective cloud response in the RP and constant RP ensembles remains very similar in the 100-member ensembles run on the new platform, compared to the response seen for the 40-member ensembles on the old platform. There are of course large differences between the responses in the lower troposphere at times when the shallow convection scheme, which was inactive on the old platform, triggered on the new platform (i.e. where the low-level convective cloud markedly increases on the new platform in figure 3.9; during the suppressed phases and on day 20).

The TSC scheme gives a general increase in convective cloud during the active phases, consistent with a general increase in convective activity due to the increased convective instability in the TSC ensemble runs. It also gives significantly less convective cloud near the convective cloud top, indicating a tendency of the TSC scheme to give slightly lower cloud-tops. This is consistent

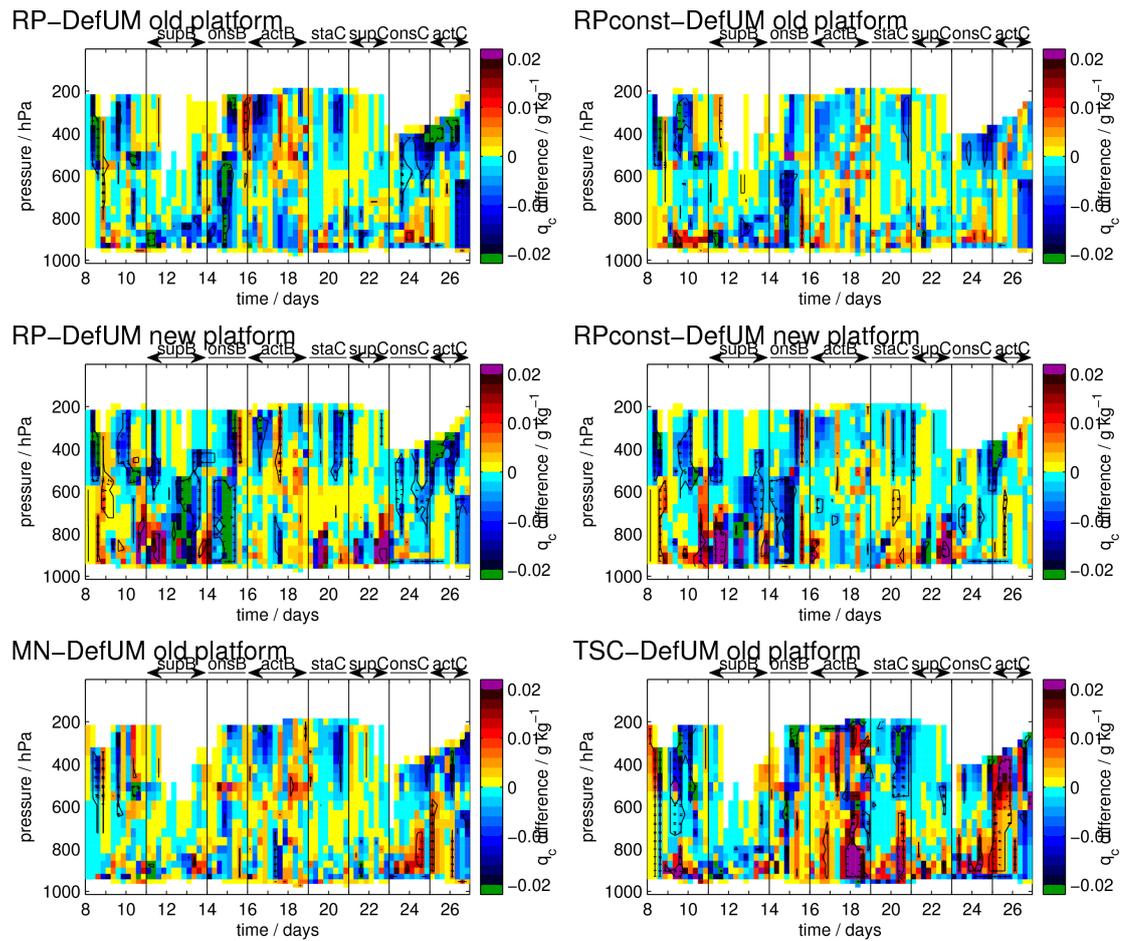


Figure 3.19: Ensemble-mean total specific convective-cloud condensate (liquid + ice) differences between ensembles. Top row shows differences for the RP ensembles relative to default UM on the old computing platform, middle row shows differences for both rerun on the new platform (mean convective cloud for the default UM on both platforms was shown in figure 3.9). The MN and TSC ensembles (bottom row) are shown for the old platform only. Black contours denote statistical significance of the differences at the (solid) 5%, and (dotted) 1% levels, using a T-test as in figure 3.11.

with the warming of the atmosphere at and above the cloud-top in the TSC ensemble (this was shown in figure 3.12). The warming at this height both promotes and is promoted by a slightly lower termination of convective plumes (plumes are less likely to remain buoyant in warmer air, and act to cool the air at the cloud-top through evaporation of detrained condensate). It was also shown in figure 3.12 that much of the troposphere becomes cooler and dryer in the TSC runs. Whilst this gives more convective instability than in the default UM, it would also be expected to make convective plumes cooler and dryer as they rise, via entrainment. Since the default UM

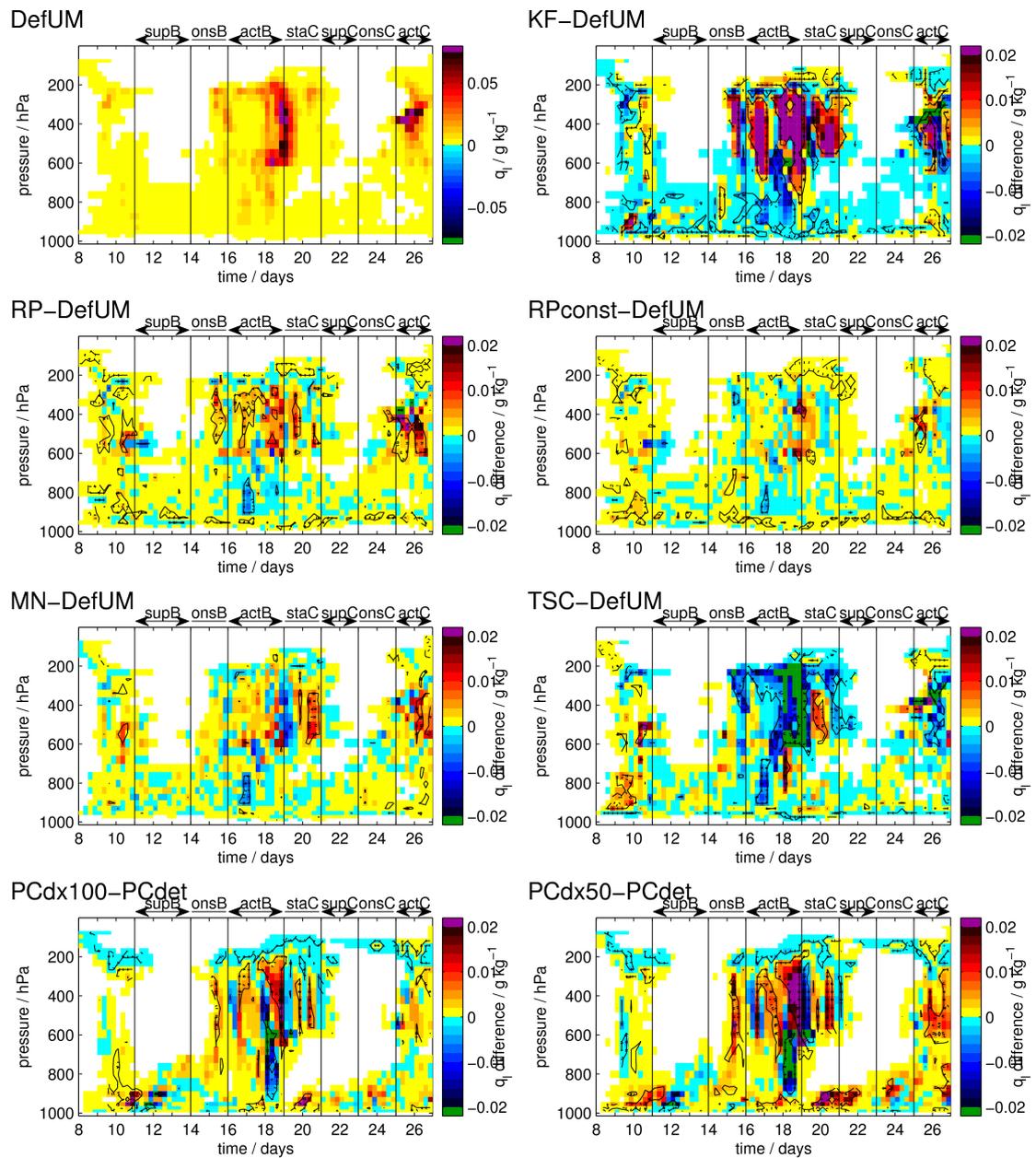


Figure 3.20: Ensemble mean total specific layer-cloud condensate (liquid + ice). Top left plot shows the mean in the default UM ensemble, other plots show differences between ensembles as labelled. Black contours denote statistical significance of the differences at the (solid) 5%, and (dotted) 1% levels, using a T-test as in figure 3.11.

convection scheme does not simulate the vertical momentum budget of plumes, it does not allow convective overshoot, so we expect the increased negative buoyancy of entrained air to cause plumes to terminate lower despite the increased CAPE.

Considering the changes in ensemble mean layer cloud condensate (figure 3.20), all of the stochastic schemes (RP, MN and stochastic PC), and the constant RP scheme, give significant increases in layer cloud at many times. Some of the changes in layer cloud are simply consistent with the noise-induced drifts in humidity (figures 3.11, 3.12 and 3.13); for example, the brief reductions in layer cloud in the lower free-troposphere at day 17 in the RP and MN ensembles, and at day 18 in the stochastic PC ensembles, correspond to times when the stochastic runs are dryer and warmer than their deterministic analogues. The large increases in layer cloud for the KF ensemble relative to the default UM are similarly explained by KF having a mean-state which is substantially cooler, yielding higher relative humidities. However, not all the increases in layer cloud in the stochastic runs line up with obvious humidity increases.

To investigate this further, figure 3.21 shows PDFs of saturation; specific total water content (vapour plus condensate) q_t minus the saturation specific humidity q_{sat} ; comparing all the different SCM ensembles. The PDFs were computed for the two main periods during which extensive layer cloud formed; days 15–21 and days 25–27 (see top-left panel of figure 3.20), and for four different height intervals as labelled.

In general the SCM configurations which use the KF convective plume model (KF and PC) yield saturation PDFs with different characteristic shapes to those based on the default UM's convection parameterisation, with the latter having more spread than the former in the mid and upper troposphere. This is consistent with the differences noted in the profiles of ensemble spread in temperature in section 3.2. The latter also have greater mean saturation, consistent with their increased layer cloud (figure 3.20) and increased large-scale precipitation (figure 3.15).

However, the ensemble mean saturation averaged over each cloudy period remains below zero for all the SCM configurations at all four heights. Whilst the UM's cloud scheme allows some cloud to form at sub-saturation above a critical relative humidity threshold, most of the cloud-formation in these runs must occur when the model states make excursions substantially above the ensemble mean saturation.

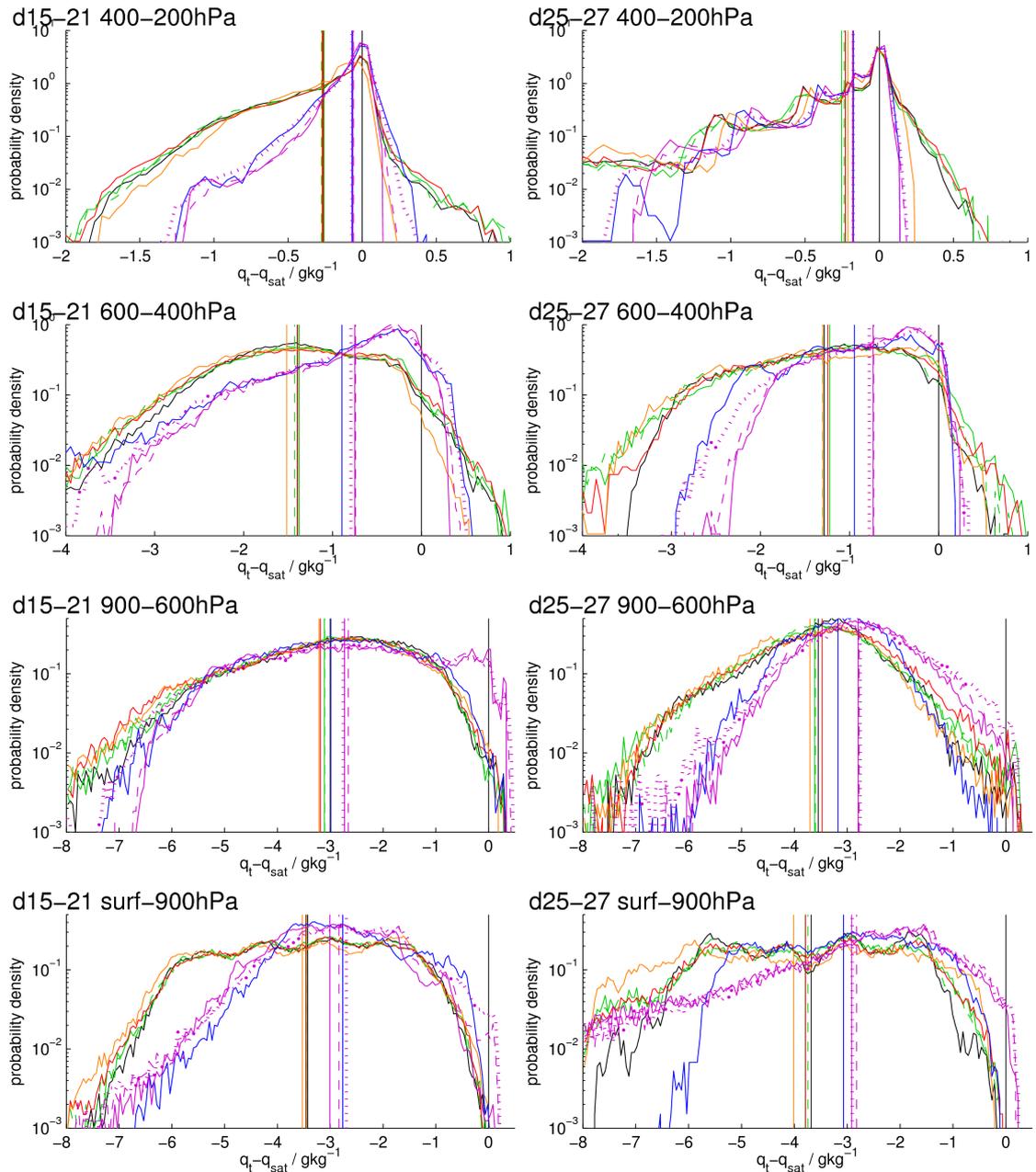


Figure 3.21: Probability densities of specific saturation $q_t - q_{sat}$ for the two cloudy periods (left) $t=15-21$ days, and (right) $t=25-27$ days; for 4 different heights as labelled. Colours denote different ensembles; (black) default UM, (red) MN, (solid green) RP, (dashed green) constant RP, (orange) TSC, (blue) KF, (solid purple) deterministic PC, (dashed purple) stochastic PC with grid-length 100 km, and (dotted purple) stochastic PC with grid-length 50 km. Coloured vertical lines denote the means of the distributions. Note that the plots at different heights have different axis scales, and the y-axis is logarithmic.

The primary effect of the stochastic schemes (RP, MN and stochastic PC) on the saturation PDFs in figure 3.21 appears to be to broaden the tails of the distributions relative to their deterministic analogues (DefUM and deterministic PC). On average, the distributions are broadened equally for both positive and negative excursions, yielding little change in the mean saturation. The increased frequency of excursions to very low saturations has no effect on the cloud condensate, as a moderately low saturation in these distributions is still far enough below zero to yield no cloud condensate. However, the increased frequency of excursions to positive saturations must have a positive impact on the mean cloud condensate. Therefore, we expect the broadening of the saturation PDFs associated with stochastic parameterisations (or indeed any scheme which increases the range of saturation states available to the system) to yield systematic increases in the ensemble mean cloud condensate, aside from the effects of any noise-induced drift in relative humidity. The increases in layer cloud produced by the RP, constant RP, MN and stochastic PC schemes are consistent with this mechanism.

3.5 CONCLUSIONS

The SCM framework described and validated in chapter 2 has been used to compare a number of different configurations for the sub-grid parameterisations, which were chosen so as to explore various sources of uncertainty in the high-frequency variability of parameterised processes. A range of stochastic parameterisations and alternative deterministic formulations were used. Conclusions are drawn below, on the nature of the variability of the parameterised processes studied (section 3.5.1), on the responses of the parameterised processes to changes in their variability (section 3.5.2), and on the relative importance of physical and modelling uncertainties (section 3.5.3).

3.5.1 ON THE NATURE OF PARAMETERISED PHYSICS VARIABILITY

The vertical profile of the ensemble variability of sub-grid and non-dynamical processes was found to be crucially dependent on the convective plume model used. The time-evolution of the ensemble spread profiles followed the changing convective activity, giving peaks in the lower and

mid troposphere corresponding to periods of shallow and deep convection respectively. Notably the KF convection scheme of Kain (2004) tended to give less shallow convection than the other schemes, and correspondingly gave less ensemble variability in the lower troposphere. This is consistent with results from chapter 2 which showed that the ensemble variability is primarily driven by the parameterised convection, and that the KF scheme is less sensitive to initial condition perturbations in the lower troposphere. It also demonstrates that structural uncertainties in the convective plume model manifest as very large uncertainties in the shape of the profile of variability in diabatic heating. We would expect such disparate variability profiles to excite quite different gravity-wave spectra, with likely consequences for the simulation of any tropical atmospheric modes sensitive to high-frequency wave activity.

The characteristic timescales of variability in parameterised processes were found to be highly state-dependent, following the time-evolution of convective activity, but were relatively insensitive to structural differences in the convective plume model or the introduction of stochastic parameterisations. Generic model-uncertainty schemes were able to boost the ensemble spread without imposing their prescribed timescales on the leading modes of variability in the SCM. Rather, the timescales were found to be much more sensitive to alterations in the closure applied in the convection parameterisation. Time-smoothing the input closure profiles or the output tendencies from convection yielded obvious increases in the timescales of the SCM's variability.

In general, stochastic parameterisations were found to increase the amplitude of the variability profile, but mostly had little impact on its shape or internal properties.

A notable exception to this was the tendency of the stochastic multiplicative perturbation method of Buizza *et al.* (1999) to give disproportionate increases in temperature variability in the stratosphere, associated with radiative tendencies. This is physically questionable; in reality stratospheric radiative transfer is fairly steady and well-defined, with error correlation scales which are much larger / slower than those associated with sub-grid processes. A subsequent revision of the stochastic multiplicative scheme in the ECMWF forecast system tapers the perturbations to zero at upper levels to account for this (Palmer *et al.* 2009).

3.5.2 ON THE RESPONSE OF PARAMETERISED PROCESSES TO CHANGES IN THEIR VARIABILITY

In the case studied here, increasing the high-frequency variability of the sub-grid processes primarily gives an increase in the amplitude of the profile of variability in convective heating, even though other processes are also perturbed. This is because convection is the process with the strongest non-linear sensitivity to the atmospheric state, and which generates the largest fluctuations in the moisture budget and diabatic heating. However, if one chose a different case study which was not dominated by deep convection, different results would be likely.

The introduction of stochastic parameterisations also yielded subtle changes in the ensemble mean-states of the SCM, typically of the order $0.5 \text{ K} / 0.5 \text{ gkg}^{-1}$ in terms of tropospheric temperature and humidity. These noise-induced drifts were highly state-dependent, changing sign at different times and heights. They were much smaller than those associated with switching between different deterministic convection scheme formulations, which gave consistent profile changes of the order $2 \text{ K} / 1 \text{ gkg}^{-1}$.

The increase in convective variability seen when stochastic parameterisations (and some deterministic modifications) are introduced increases the range of atmospheric states available to the system. This has some further generic systematic effects.

The simplest of these is the increase in ensemble-mean layer cloud condensate, due to the increased range of total specific moisture contents explored relative to the saturation threshold; increased positive excursions in moisture yield increases in cloud, whereas increased negative excursions can make no difference beyond removing all the cloud. One would expect this to have systematic effects on the radiation budget.

There was also a common systematic response from the boundary layer parameterisation, giving increased moisture fluxes in response to the increased range of states available to it, and forcing changes in the lower tropospheric humidity profile. It is not clear whether this effect is particular to the numerics of the UM boundary layer scheme or whether it corresponds to a more general physical sensitivity of atmospheric boundary layers. This would be an interesting topic for further study.

Many of the stochastic (or otherwise perturbed) SCM configurations also gave occasional compensating decreases in the low-level moisture fluxes from the convection scheme, yielding reductions in convective cloud at upper levels at certain times. The stochastic parameterisations may have served to highlight the highly nonlinear sensitivity of interactions between parameterised boundary layer fluxes and convection. The distinction between the two is somewhat arbitrary in models (there is no defining line between buoyancy-driven boundary layer over-turning and shallow convection in reality), but the parameterisations for each behave differently at different times. As a result, the changes in the relative activity of these two parameterisations promoted by increased variability yield highly state-dependent changes in lower troposphere temperature and moisture profiles, and in model cloud fields.

Deterministically forcing the convection to respond on longer timescales (in this case by applying a simple time-smoother to the output tendencies) was found to increase the ensemble spread, giving a range of model states comparable with estimates of parameterisation uncertainty. The time-smoother acts to delay part of the convective heating response to profile instabilities, yielding a cooler, more unstable profile, which promotes more vigorous, coherent convective responses overall. This is consistent with the results of Scinocca & McFarlane (2004), who found that applying a prognostic closure to a convection parameterisation increased the variability of the convective rainfall. Prognostic closures, which relate the rate of change of dissipation of CAPE to the amount of CAPE, will similarly act to delay the convective response to instabilities as the convective mass flux will take a finite time to ramp up to the equilibrium response rate. It therefore seems reasonable to conclude that they yield increased activity and variability of the convection parameterisation through a similar mechanism to the simple convective time-smoother in the present study.

These results show that strong sensitivity of convective variability to the convective closure can arise from parameterised physical feedbacks alone, although this could be further amplified by excitement of dynamical modes in the atmosphere in 3D simulations, such as those of Scinocca & McFarlane (2004).

3.5.3 ON THE RELATIVE IMPORTANCE OF PHYSICAL AND MODELLING UNCERTAINTIES

Stochastic parameterisations designed to represent generic parameterisation uncertainty were found to produce similar ensemble ranges of states to that associated with a somewhat limited sampling of different structural configurations for the convection parameterisation. This suggests that either the poor-man's ensemble (of three deterministic convection schemes) used in this study turned out to be fortuitously representative (i.e. that the particular uncertainties in the convective plume model and closure sampled by that ensemble actually dominate overall parameterisation uncertainty), or else the stochastic model uncertainty schemes only partially sample parameterisation uncertainty. The latter is clearly the case when considering the vertical profiles of variability; applying the stochastic schemes to the default UM convection scheme never yields profiles of heating variability which look like the other deterministic convection schemes.

Encouragingly however, the stochastic Random Parameter scheme (Bowler *et al.* 2009) gave spreads very similar to a constant perturbed parameter ensemble analogue, suggesting that locally the stochastic method can access the full range of states available to the perturbed-parameter ensemble, whilst gaining the desirable quality that the ensemble members all have the same attractor.

The effects of sampling the uncertainty in the sub-grid state of convection were quantified by running the stochastic convection scheme of Plant & Craig (2008), which simulates a realistic population of convective plumes within a given grid area, drawn at random from a theoretical distribution of cloud sizes (and corresponding entrainment rates). This scheme was found to explore a similar range of model states to that associated with the modelling uncertainties investigated in this study (as quantified by schemes designed to represent generic model uncertainty, and by the aforementioned sample of three different deterministic convection scheme configurations), for a grid-size of 50 km. For a larger grid-size of 100 km, the ensemble of possible sub-grid convective states is more fully sampled, giving more well-defined grid-mean convective tendencies, and therefore exploring a smaller range of model states.

These results suggest that structural uncertainties in the convection parameterisation are more important than convective sub-grid uncertainty at scales characteristic of climate GCM grid resolutions (100km or more), but that in current weather forecast models, which often have grid-sizes

well under 50km, sub-grid uncertainties in convection may well dominate the overall uncertainty in the grid-mean tendencies.

There is a significant caveat to the result for larger grid-sizes; the Plant & Craig (2008) stochastic convection scheme used to quantify sub-grid convective uncertainty simulates departures from statistical equilibrium in an ensemble of *non-interacting* plumes. It is therefore unlikely to be representative of fluctuations associated with organised convective systems born of the neglected interactions between plumes. In current weather forecast models, one would hope that such systems can be resolved by the model dynamics, making this a moot point, but in climate GCMs they are certainly not resolved. Therefore, sub-grid uncertainties associated with convection at the larger grid-size tested in this study might be underestimated; organised convective systems are the most energetic sub-grid phenomena at these scales.

In the next chapter, the sensitivity of the atmosphere to high frequency variability in an aqua-planet GCM is investigated, with a horizontal resolution consistent with a global climate simulation. The results in the SCM suggest that, for the large grid-sizes in the GCM, the sub-grid convective fluctuations generated by the Plant & Craig (2008) scheme will be dwarfed by generic parameterisation uncertainties. Therefore, in the aqua-planet climate GCM experiments, the model uncertainty schemes (random parameters and multiplicative noise) are more likely to have a significant effect and yield relevant results than the Plant & Craig scheme. The latter scheme would be a far more valuable tool for investigating the sensitivity to sub-grid variability in higher-resolution simulations consistent with weather forecast models.

CHAPTER 4:

**SENSITIVITY OF THE ATMOSPHERE
TO HIGH-FREQUENCY VARIABILITY
IN AN AQUA-PLANET FRAMEWORK**

4.1 OVERVIEW

This chapter gives a preliminary investigation into the sensitivity of the 3-dimensional global atmosphere, with resolved-scale dynamical interactions, to high-frequency variability associated with sub-grid and non-dynamical atmospheric processes. As in the previous chapter, the method will be to alter the high-frequency variability in a host model by introducing stochastic parameterisations from the literature, and investigate the occurrence (or not) of noise-induced changes to the mean-state or to modes of variability not directly perturbed by the stochastic scheme.

In the previous chapter, the response of parameterised physical and sub-grid processes to high-frequency variability consistent with generic parameterisation uncertainty was studied in a Single-Column Model (SCM) simulation of convection over the Tropical West Pacific, using two different “stochastic physics” parameterisations used in ensemble prediction systems. Both the schemes substantially increased the variability of the SCM. The resulting increase in the range of states available to the model parameterisations also yielded a systematic increase in mean layer cloud condensate, an increase in the moisture fluxes produced by the boundary layer scheme, and a highly state-dependent response from the convection parameterisation, which reduced its moisture flux in the lower troposphere. One of the stochastic schemes (the Random Parameters scheme described in Bowler *et al.* 2009) also produced a small systematic drying of the model column due to increased large-scale precipitation, whereas the multiplicative perturbation scheme of Buizza *et al.* (1999) did not. Whilst these changes were on the whole fairly subtle, it would be

interesting to investigate the additional effects of resolved-scale dynamical interactions.

In this chapter, the same two stochastic methods are implemented in a 3-dimensional model, and any systematic changes in mean-state or variability are investigated. It will be interesting to check whether the full dynamical atmosphere's responses to the stochastic physics schemes mirror those of the Single-Column Model, or whether they are muted by a compensating "dynamical stabilisation" response, or amplified, or otherwise altered by dynamical feedbacks. Or it is possible that the atmosphere's response to high-frequency variability is heavily state-dependent; thus switching to a different modelling framework, which dynamically follows its own attractor rather than being constrained to observed tendencies, would yield very different responses.

The SCM experiments in the previous chapter also used the stochastic convection scheme of Plant & Craig (2008) to investigate the tropical atmosphere's response to sub-grid variability arising in deep moist convection due to the likely variations in the size and number of convective plumes present in a model grid area. It was found that these fluctuations about convective equilibrium produce more variability than that consistent with estimates of generic parameterisation uncertainty for grid-sizes less than 50km, but dwindle in magnitude at larger grid-scales.

In this chapter, multi-year integrations consistent with climate simulations will be studied, using an appropriate horizontal resolution with grid-sizes in the hundreds of kilometres. The results from the previous chapter show that the Plant & Craig (2008) stochastic convection scheme would produce only very small fluctuations at these scales; any noise-induced effects would likely be subtle and hard to detect. Whilst fluctuations associated with mesoscale organised convective systems are likely to be an important component of sub-grid uncertainty / variability at climate GCM resolutions, their thermodynamic effects are not explicitly represented by any convection parameterisation known to the author. Therefore fluctuations about convective statistical equilibrium will not be studied in this context.

To aid interpretability, it will be desirable to use the simplest experimental framework available that is able to simulate the general behaviour of the global atmosphere. For this, the Aqua-Planet framework, in which the earth's surface is assumed to consist entirely of open ocean, will be used. This is an ideal framework to use because it includes full interaction of dynamics and physics and the basic features of the global circulation, but removes the complications of land sur-

face interactions, topography and land/sea configuration. The framework used here also fixes the solar date at equinox so that there is no annual cycle. Further, it is zonally- and hemispherically-symmetric. Thus in terms of its climate it can be zonally and temporally averaged without losing any of the representativeness of a single point. All these properties allow robust results to be obtained more easily and with much shorter runs than would be required in a realistic earth-configuration.

The Aqua-Planet simulations used here are based on those described by Neale & Hoskins (2001a,b). They investigated the sensitivity of the global circulation to the Sea Surface Temperature distribution, and established a benchmark set of simulations which have been widely used for testing models and investigating the sensitivity of the global atmospheric circulation to various factors. In this chapter, preliminary results for aqua-planet simulations employing two different stochastic parameterisations are presented, compared against a control simulation with no stochastic element.

The experiments, model configuration and statistical methods used are fully described in section 4.2. Results are presented in section 4.3. Conclusions are drawn in section 4.4, along with further discussion in the light of the conclusions from the Single-Column Model experiments described in the previous chapter.

4.2 EXPERIMENTAL SETUP

Firstly, the host model is described in subsection 4.2.1. The stochastic perturbed runs are then described in subsection 4.2.2. Before comparing the stochastically perturbed runs against the unperturbed control run, it is essential to find appropriate statistical tests to determine which differences are likely to be systematic effects, rather than chance variations between the runs. The statistical methods used are described in subsection 4.2.3.

4.2.1 MODEL DESCRIPTION

The model used is the UK Met Office Unified Model (UM). The particular version used is the global model used in their forecasting system at Parallel Suite 13 (PS13). The system is described by Bell (2008); see the table of global model cycles in Annex A of Bell (2008) for detail of which modifications to the UM system were in effect at PS13.

The model used in the present study differs from the Met Office global forecast model at PS13 in that (a) a much lower horizontal resolution is used (N48 rather than N320), and (b) it is run in aqua-planet configuration. To run the model in aqua-planet configuration the following modifications are included:

- All orography is set to zero and all surface points are set to ocean.
- A fixed zonally and hemispherically symmetric Sea Surface Temperature (SST) field is specified. This is the “control” case described by Neale & Hoskins (2001a). The surface temperature T_s in °C is specified as a function of latitude ϕ using the formula

$$T_s = \begin{cases} 27(1 - \sin^2(\frac{3\phi}{2})) & : -\frac{\pi}{3} < \phi < \frac{\pi}{3} \\ 0 & : otherwise \end{cases}$$

- The time specification in the radiation scheme is modified so that the date is fixed at 21st March (spring equinox), but the time of day progresses as usual. This removes the seasonality, but preserves the diurnal cycle, in the solar flux calculated by the radiation scheme.
- An idealised calendar is used with months and years set to fixed lengths of 30 days and 360 days respectively. Since there is no seasonality, this has no consequence for the model evolution, but allows time to be easily sub-divided consistently.
- radiatively important trace gases such as ozone are set to zonally and hemispherically averaged climatological fields.

The resolved dynamics are simulated using the dynamical core described by Davies *et al.* (2005). This is a grid-point based, non-hydrostatic, non-shallow-atmosphere, mass-conserving, semi-implicit fluid solver. The scheme uses a predictor-corrector method in which a first estimate of the next time-step model fields is made using semi-Lagrangian advection, and dynamical

adjustment is then calculated using an iterative algorithm to solve the 3D Helmholtz equation. The model grid uses Charney-Phillips staggering in the vertical (pressure and horizontal winds on model-levels, with thermodynamic variables and vertical wind on half-levels above and below), and the Arakawa C-grid staggering in the horizontal (pressure at main grid-points, zonal winds at half-points to the E and W, meridional winds at half-points to the N and S).

The parameterisation schemes are essentially the same as those used in the SCM described in the previous chapter, but with some minor differences in accordance with the modifications present in the global model at PS13 (Bell 2008). The most substantial of these differences is the inclusion of a scheme to make convective cloud decay gradually rather than assuming it all disappears by the next time-step; Convective Cloud Amount (CCA) and Convective Cloud Condensed Water Path (CCCWP) decay with a timescale of 2 hours after the convection scheme outputs them. A bug fix was also introduced before PS13 which corrected an error in the formula for partitioning convective cloud condensate between liquid and ice based on cloud-top temperature. This fix did not affect the convection scheme itself, but applied to the diagnostic calculation of convective cloud profiles used by the radiation scheme. This bug-fix was not used in the SCM experiments described in the preceding chapters, but was later implemented in the SCM and ensemble tests indicated that it had a detectable but very small effect on the simulations.

Another key difference between the aqua-planet studied here and the SCM studied in the preceding chapters is the number of vertical levels. The aqua-planet has 50 rather than the SCM's 38. However, the two level-sets have exactly the same vertical spacing throughout the troposphere, with all the aqua-planet's extra levels being added in the stratosphere.

The aqua-planet has a horizontal resolution of N48, which is 96 grid-points in the longitudinal direction and 73 in the latitudinal. This gives a grid-size of 417-by-278 km at the equator.

As in the PS13 global model, a time-step of 15 minutes is used, with a 7.5 minute sub-step for the convection scheme and the radiation scheme called once every 3 hours. For comparison, the single-column model used in the previous chapter had a time-step of 30 minutes with no convective sub-stepping, but the same 3-hour radiation time-step. 15 minutes may seem an unnecessarily short time-step to use in a model with such a coarse horizontal resolution (presumably this is the conventional setting for the N320-resolution PS13 forecast model). It would be tempting to in-

crease the time-step to 30 minutes to make the runs more consistent with the SCM experiments and with climate simulations. However, it was felt it would be wise to preserve the model settings “as received” from the UK Met Office, so that the model can be trusted to behave as documented. It is also likely to have been tuned for optimal stability with these settings.

4.2.2 EXPERIMENTS

Three Aqua-Planet simulations were initialised using the same model dump from a pre-existing spun-up aqua-planet simulation. Note that although the experimental *framework* is entirely zonally and hemispherically symmetric, the *initial state* used for these experiments is not. Therefore the runs evolve in a realistically asymmetric manner, permitting irregularly spaced eddies and flow across the equator as in the real world. But the symmetries of the experimental framework should eventually yield a zonally and hemispherically symmetric climatology if the model is run for a long-enough period of time.

Three different experiments were performed, with:

- *DefUM*: an unmodified control simulation of the aqua-planet described above.
- *RP*: a version with the Random Parameters stochastic physics scheme based on Arribas (2004), also described by Bowler *et al.* (2009).
- *MN*: a version with the Multiplicative Noise stochastic scheme (MN) following Buizza *et al.* (1999).

The implementation of both of the stochastic schemes used are described in detail in section 2.2.2 of chapter 2, and the internal parameters of the stochastic schemes are set to the default values specified therein. The motivation and context for these stochastic schemes was discussed in section 1.5 of chapter 1.

In the case of the MN run, in addition to the time-variation of perturbation scalings described in chapter 2 for the Single-Column Model, there is a spatial component to the variability in the aqua-planet. This is implemented following Buizza *et al.* (1999), by drawing random numbers for the scheme independently at different locations. Horizontal autocorrelation of perturbations is

applied very simply by drawing the same random number for grid-points falling within each 2D tile of neighbouring points. Buizza *et al.* (1999) used tiles of size 10° longitude by 10° latitude. In the implementation studied in this chapter, this is approximated using tiles spanning 3-by-4 grid-points, equivalent to 11.25° longitude by 10° latitude. In the case of the RP scheme, following Arribas (2004), there is no spatial component to the scheme and the same time-series of model parameter values is applied at all grid-points.

In each of the runs, time-averaged zonal mean data were output for various diagnostics on 15 selected model levels spread throughout the depth of the atmosphere. Also, to investigate the variability in the runs, 6-hourly mean data for basic model fields, precipitation, surface fluxes and top-of-atmosphere radiative fluxes were output for the full length of each run. The 3D fields were calculated as mass-weighted vertical averages over 6 vertical intervals in pressure-coordinates.

Initially, each of these three model configurations was run for 18 months, treating the first 6 months as spin-up and analysing the remaining 1-year of data, following Neale & Hoskins (2001a). However, it was difficult to gain statistically robust results from these runs, as all three simulations exhibited interesting long-timescale modes of variability. These led to apparent differences between the 1-year average zonal mean fields in the runs which were in fact the result of the slow modes following a different evolution in each run. The non-robustness of these differences was immediately apparent by their possessing large hemispheric asymmetry despite the model framework being entirely symmetric about the equator.

In order to gain more robust results, longer runs were then performed with a target length of 4 years and 4 months, treating the initial 4 month period as spin-up. The DefUM and MN runs completed successfully, each giving 4 years of 6-hourly data and time-averaged zonal mean fields. The climate of the 4 year DefUM run is described in section 4.3.1.

However, the RP run crashed after 2 years, 4 months and 29 days in to the run. Analysis revealed that a numerical instability over the South polar region had blown-up in the RP run. Considering that this may just have been an unlucky numerical misfortune, the integration was repeated, but this time with the RP scheme set not to apply poleward of $\pm 80^\circ$ latitude, in case the interaction of the stochastic scheme with the polar singularity and any related instabilities had caused the crash. This slight change would also force the new integration to diverge from the old

one, and so avoid following the precise trajectory known to have led to a crash. Frustratingly, the new run crashed again in much the same way, after the slightly longer period of 2 years, 11 months and 29 days, even though analysis confirmed the run had indeed followed a very different trajectory.

There are two possible explanations for this. Firstly the model used might generally have a polar instability problem which infrequently erupts, and this happened to occur by chance in both of the RP runs but neither the control run nor the MN run. Or secondly, the RP scheme may be triggering the crash through a systematic noise-induced effect absent from the other simulations. Taking the former possibility as a null hypothesis, and assuming that a 4-year integration of the model ending in a crash is a binary event with a probability of 0.5 (since half of the runs crashed), the likelihood of the observed outcome can be estimated. The outcome that in a sample of 4 runs, the 2 RP runs crash and the 2 runs without the RP scheme do not, has a statistical weight of 1 out of 2^4 possible permutations, yielding a probability of $1/2^4 = 0.0625$. Therefore, it cannot quite be rejected at the 5% level, but is fairly improbable.

The mechanism for the crash was studied in more detail; it was found that for both RP runs, a deep area of low pressure moved over the South polar region, coinciding with the formation of unusually cold air in the upper troposphere and a strengthening of stratospheric zonal winds around the pole. In both cases, the low then remained slow-moving near the pole for around 20 days, steadily intensifying to a central pressure below 900hPa, before a rapid further deepening which precipitated the model crash. The intensification of the low was accompanied by large surface fluxes of heat and moisture, heavy convective snowfall and unrealistically large vertical velocities. It kept intensifying despite slack horizontal temperature gradients, suggesting that it was driven by surface heat exchange and convective instability, rather than baroclinic instability. However, its horizontal extent spanned many hundreds of kilometres; much larger than scales typical of polar lows or even tropical cyclones in the real world.

The non-realism of this synoptic-scale feature is presumably associated with the prescribed sea surface temperatures at the poles, which at 0°C are unrealistically warm and ice-free; the polar sea in these runs acted as an inexhaustible supply of heat and moisture for the developing polar cyclone. Coupling the aqua-planet to any kind of interactive ocean component would likely suppress this unphysical development, as the sea surface would cool in response to the surface

fluxes associated with the cyclone.

Analysis using a simple feature-tracking algorithm found that intense polar cyclones only occurred in the run-up to the crash in each RP integration, not earlier in the runs, or at all in the control or MN simulations. Further work is needed to ascertain whether this is just down to chance occurrence of an inherently infrequent event. Otherwise, if the RP scheme was to blame for the occurrence of the polar cyclone, it must be via an indirect mechanism, as the cyclone still appeared when the RP scheme was switched off over the polar regions. The appearance of colder air aloft (and the strengthening of stratospheric zonal winds) just as the cyclone became established over the pole hint that stratospheric dynamics may have played a role in triggering the cyclone. Unfortunately, insufficient data was output from the stratosphere in these runs to investigate this in detail.

Meanwhile, by discarding the 4-month spin-up period at the start of each of the two RP integrations, and a similar period in the run-up to the ill-fated end of each run, the two RP runs could be compared against the control run to investigate any systematic differences. The results of this comparison will not be as statistically robust as was hoped, due to the truncated nature of the runs. Further, the two RP runs cannot be treated as two realisations of an identical model, in case the altering of the stochastic scheme to suppress perturbations near the poles in the second RP run had some systematic effect. The two RP runs are therefore compared to the control run separately. Nonetheless, some significant effects were still detected.

4.2.3 STATISTICAL METHODS

In general, the textbook statistical tests for assessing the significance of differences between samples, such as T-tests, make the assumption that the data within each sample are independent. However, the 6-hourly grid-point atmospheric variables from the simulations presented here exhibit considerable autocorrelation in both time and space, associated with variability on a broad range of scales. So the sample of data for a given model run are clearly not independent.

One approach in common use in the literature, for performing statistical tests on samples which contain autocorrelation, is to carry out standard tests but account for the autocorrelation by reducing the degrees of freedom in the tests. There is assumed to be some number of truly

independent variables underlying the data, this number being some degree less than the sample size, depending on the degree of autocorrelation. This approach, involving the estimation of the underlying “effective sample size”, is used in this study. In this subsection, the statistical tests used are described, followed by the methods used to estimate the effective sample size.

To test the significance of differences in the mean-states of the model runs, a T-test can be used. Since we anticipate that the stochastic schemes may alter the variability of the model, a test which does not assume equality of variance between the runs is required. The “Welch T-test” (Welch 1947) is therefore used here. The implementation of the test used is described in section B.1 of appendix B.

To test the significance of differences in variance between the simulations, the test of Levene (1960) is used. This test has been chosen as it is robust even when the distributions of the variables being tested are substantially non-normal. We would indeed expect some of the atmospheric variables tested in this study, such as cloud water content, to have heavily skewed distributions. The implementation of the test used here, and the method for incorporating reduced effective sample sizes, are described in section B.2 of appendix B.

Next, a robust method of estimating the effective sample size must be found. Perhaps the most widespread method is that used by Dawdy & Matalas (1964). According to their method, for samples which consist of time-series containing some autocorrelation, the effective sample size n_{eff} is given by:

$$n_{eff} = \frac{1 - r}{1 + r} n \quad (4.1)$$

where n is the number of values in the time-series, and r is the lag-1 autocorrelation (the correlation coefficient of the data with itself lagged by one step in the time-series). This formula is derived by modelling the sample data as a first-order autoregressive process (i.e. it assumes that the autocorrelation in the data can be entirely described by iteration of the lag-1 relationship).

For the aqua-planet data, we expect that the effective sample size will vary with height and latitude, depending on the character of the variability in different regions. But we will assume that it remains constant in time and longitude when the spin-up period is excluded. At each

height/latitude there is a corresponding 2-D field in time/longitude from which to calculate n_{eff} . Equation 4.1 can be generalised in to 2-D:

$$n_{eff} = \left(\frac{1 - r_t}{1 + r_t} \right) \left(\frac{1 - r_x}{1 + r_x} \right) n_t n_x \quad (4.2)$$

where r_t and r_x are the lag-1 autocorrelations in time and longitude respectively, n_t is the number of time-points in the data-set, and n_x is the number of grid-points in the longitudinal direction.

To test the statistical methods in this study, null comparison tests can be made using two samples from the same simulation. To verify the statistical tests, the first and last year of the control run have been compared. As long as there is no systematic drift in the model between these two periods, the null hypotheses should be true, and the test statistics should follow their theoretical distributions. To assess this, probability values have been computed from the tests for differences in mean-state and variance described earlier, with effective sample sizes calculated using equation 4.2. This calculation is done separately at each height/latitude in the data, yielding a sample of p-values. This sample should follow a flat distribution between 0 and 1, if the null hypothesis is true *and* the effective sample sizes used are appropriate. The cumulative distributions of these p-values are shown in figure 4.1, for the test statistics and effective sample sizes calculated separately for three different model variables.

In the null case, the p-values in figure 4.1 produced by the tests should be approximately evenly spread between 0 and 1, corresponding to the flat diagonals. However, they are systematically under-spread for the zonal winds u and temperature T , whilst somewhat over-spread for convective cloud water content q_c , especially for the test of difference in variance. This suggests these statistical tests will tend to erroneously fail to reject the null hypothesis when testing on u and T , and erroneously reject it for q_c . Correspondingly, the effective degrees of freedom used have been under-estimated for u and T , but over-estimated for q_c .

This suggests that the lag-1 autocorrelation method based on Dawdy & Matalas (1964) is not appropriate here and would yield misleading results in statistical tests for differences between the runs. Subsequent analysis of the data indicated that its autocorrelation behaviour departs substan-

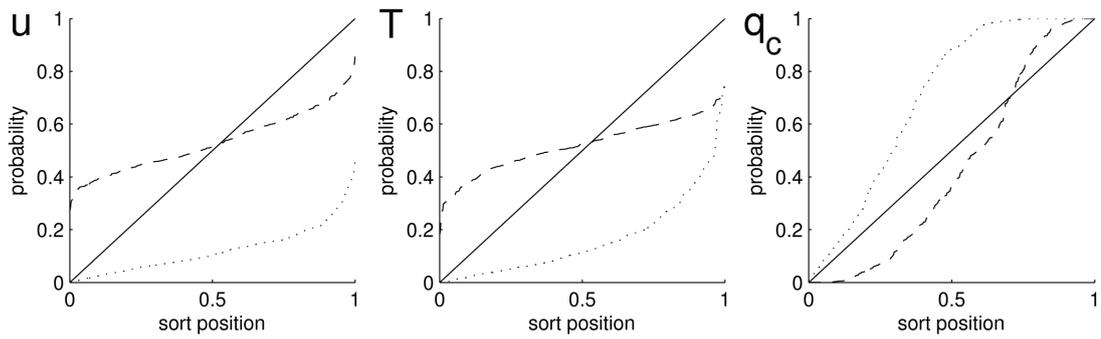


Figure 4.1: Cumulative probabilities produced by (dashed) the Welch T-test for difference in mean-states, and (dotted) Levene's test for difference in variance, for a comparison of the first and last years of the 4 year control run. Both tests use the effective sample size based on Dawdy & Matalas (1964), as described in the text. The solid diagonal corresponds to a flat distribution of probabilities. These have been computed for (left) zonal wind u , (middle) temperature T , and (right) convective cloud water content q_c

tially from the first-order autoregressive model on which the Dawdy & Matalas (1964) effective sample size estimate (equation 4.1) is based. Substantial autocorrelations occur at lags of many days, independent of the lag-1 autocorrelation, presumably associated with the atmosphere's numerous and complex dynamical modes.

An alternative method commonly used for estimating the effective sample size uses the properties of the known sampling distributions of sample statistics. Most useful statistics have a well defined theoretical sampling distribution whose variance is a known function of the number of degrees of freedom used to estimate the statistic. If the statistic in question is calculated for a sample of independent sub-sets of the available data, the variance of the statistic can be computed. By inverting the expression for the variance of that statistic's sampling distribution, the degrees of freedom in each sub-set can be recovered.

For example, Lorenz (1969) estimated the degrees of freedom in the observed northern-hemisphere extra-tropical circulation from the sampling distribution of mean-squared differences between all possible pairs of observed fields drawn from the data. Similarly, Van Den Dool & Chervin (1986) estimated the degrees of freedom in anomaly maps of model fields based on the known sampling distribution of the correlation coefficient (Panofsky & Brier 1968). They estimated the variance of the anomaly pattern by correlating the fields at 239 different pairs of

time-points in their data, and derived the degrees of freedom from this estimate.

One problem with using an estimated effective sample size or degrees of freedom to perform statistical tests is that we have no reason to assume that the same persistence or autocorrelation scales will apply to different statistics being compared (e.g. means, variances or correlation coefficients) in a given data-set. Indeed, they appear not to for q_c in figure 4.1, where the test for a difference in variance is considerably more over-spread than that for a difference in mean-state.

It therefore seems necessary to estimate the effective sample size separately for each type of statistic being tested, in a way that specifically considers the sampling distribution of the given statistic. Here, such methods are described and verified, for use in tests on sample means and variances.

- *Effective Sample Size for the Sample Variance.* As was utilised by Lorenz (1969), mean-squared-differences between sub-sets of a data-set should approximately follow a χ^2 distribution with known mean and variance for a given d.o.f. This is also true for mean-squared differences of sub-sets of the data from their mean, thus the sampling distribution of the variances of sub-samples of a data-set can be written in terms of a χ^2 distribution. From this, the theoretical mean and variance of the sub-sample variances can be expressed in terms of the number of degrees of freedom, and the effective sample size can be recovered by calculating these quantities and inverting the expressions for them. The resulting formula for effective sample size and its derivation are given in section B.3.1 of appendix B.
- *Effective Sample Size for the Sample Mean.* Similarly, the sampling distribution of the means of sub-samples of the data has known mean and variance, in this case given by the central limit theorem. Therefore, if the variance of sub-sample means is empirically calculated from the data-set, the theoretical expression for the variance of means can be inverted to recover the degrees of freedom within each sub-sample, and hence the effective sample size. The formula used and its derivation are given in section B.3.2 of appendix B.

In the remainder of this chapter, statistical tests are performed with effective sample sizes estimated from the sampling distributions of the relevant statistics as described above, using sub-

samples which span all longitudes (96 grid-points) and time-intervals of 15 days (60 output time-steps). First, a verification of the statistical tests using this revised method for estimating effective sample size is presented in figure 4.2.

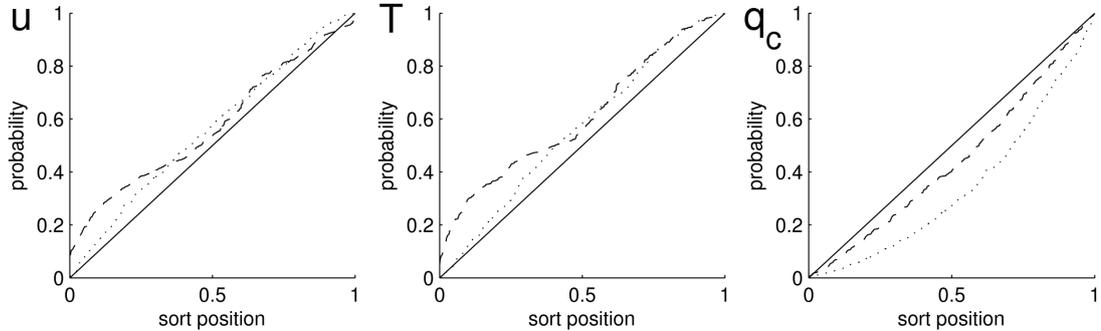


Figure 4.2: As figure 4.1, but for statistical tests using effective sample sizes estimated using the sampling distribution methods described in the text.

Comparing figure 4.2 with figure 4.1, it is clear that the statistic sampling-distribution method for estimating effective sample size performs far better than the lag-1 autocorrelation method. This time, the distributions of p-values for the null comparison (between the first and last year of the control simulation) are close to the theoretical even spread between 0 and 1, suggesting that the effective sample sizes used are (at least approximately) appropriate.

4.3 RESULTS

In this section the differences in mean-state and variance between the stochastic aqua-planet simulations and the control run are presented, to investigate any systematic effects resulting from the increase in high-frequency variability associated with the stochastic schemes. Firstly in subsection 4.3.1 the control run is analysed in order to elucidate the nature of the simulated climate. The statistical tests described in section 4.2.3 are then used to compare the mean-states and variances of the stochastic simulations with the control run, considering the moisture budgets and rainfall variability in subsection 4.3.2, the radiative fluxes and heat budgets in subsection 4.3.3, and the large-scale circulation in subsection 4.3.4.

4.3.1 ANALYSIS OF THE CONTROL RUN

To give an overview of the general features of the aqua-planet's climate, figure 4.3 shows a representative snapshot of surface pressure, thickness and precipitation rate projected on to the sphere. Note the intense rainfall along the narrow Inter-Tropical Convergence Zone (ITCZ), the sequence of dry subtropical anticyclones either side, the active mid-latitude storm track at around 40°N with rain-bearing frontal systems, and the broad polar region characterised by slack temperature gradients and scattered light precipitation.

The specified Sea Surface Temperature (SST) in the aqua-planet differs substantially from the zonal mean of observed SSTs (not shown) in that the warm tropical belt is narrower, with the largest gradients in surface temperature occurring substantially closer to the equator. As a result, the tropical region and the sub-tropical anticyclones which bound it are more contracted towards the equator. Also there is no sea-ice, and the SST is held at a constant 0°C poleward of 60° latitude. This gives polar surface temperatures which are much warmer than in the real world, and uniform over the broad polar region. Due to the lack of temperature gradients, baroclinicity vanishes polewards of 60° , so there is relatively little eddy activity in the polar region. But with warmer than real-world surface temperatures across the region, one would expect the atmosphere immediately above to be similarly anomalously warm and become unstable through longwave radiative cooling. The presence of convective showers over the pole (as evident in figure 4.3) is consistent with this source of instability.

Figure 4.4 shows the zonal-mean, time-mean precipitation broken down in to its large-scale and convective, rain and snow components. Nearly all the rain in the ITCZ is produced by the convection scheme, and there are also substantial regions of convective activity at $\pm 30^{\circ}$, on the poleward flanks of the sub-tropical anticyclones. These appear to be associated with troughs that extend from trailing cold fronts in the mid-latitude storm track in to the tropics (2 such features are visible in figure 4.3, one near the centre of the image at 0° longitude, the other near 60° longitude. Tropical / extra-tropical interactions of this kind seldom happen in reality, but seem to be quite common in the aqua-planet, presumably because the storm-tracks are much closer to the equator than in the real world. Peaks in storm-track precipitation occur at $\pm 40^{\circ}$. Poleward of about 50° most precipitation falls as snow, with snowfall associated with convective showers

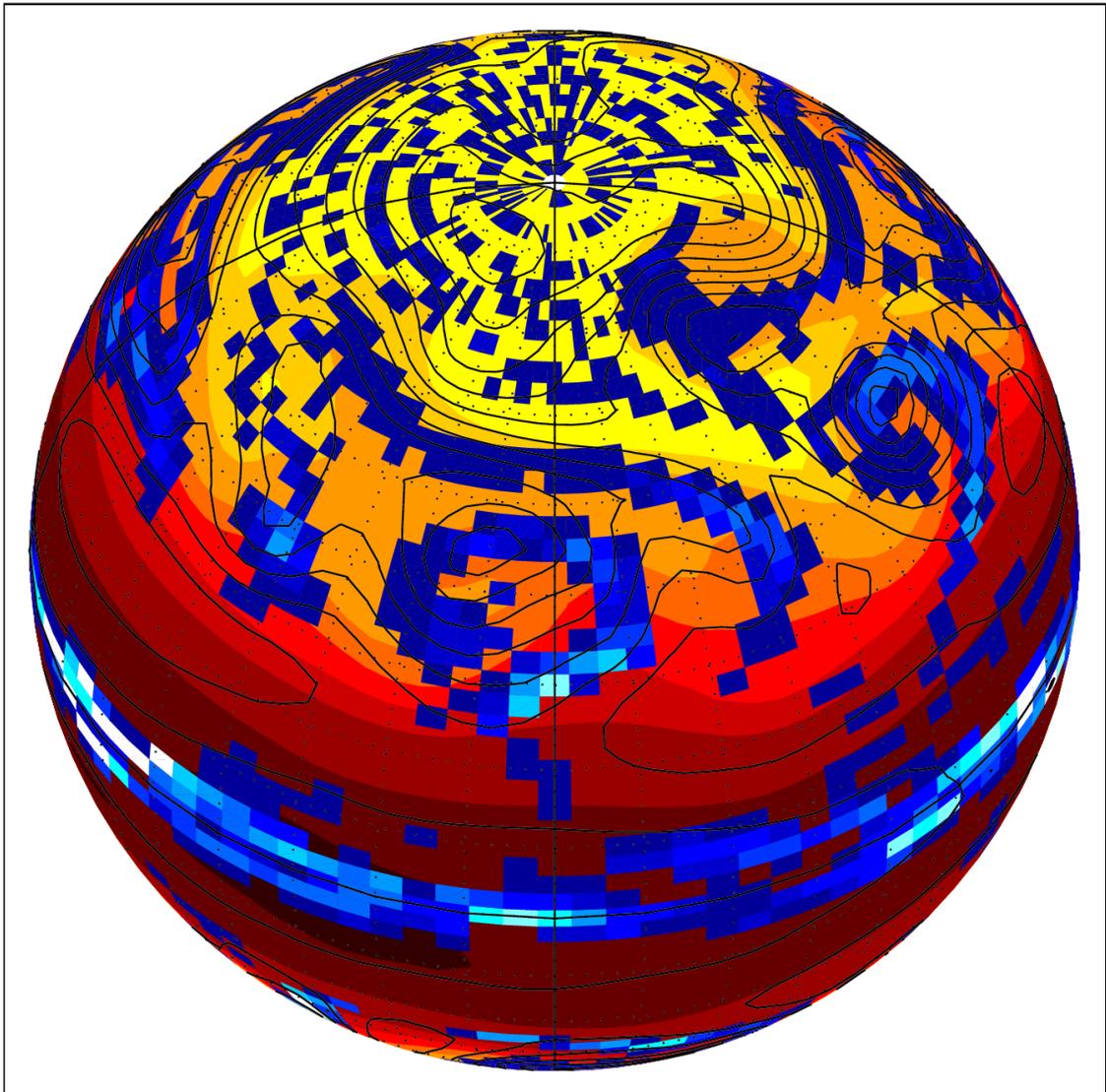


Figure 4.3: Snapshot of 6-hourly mean data from the Default UM aqua-planet control run; (solid contours) surface pressure with a 5 hPa contour interval, (red/yellow) 500-850 hPa thickness with a 60m contour interval, and (blue/white) precipitation rate on a linear colour scale. Rates less than 1 mm day^{-1} are not coloured, and rates greater than 40 mm day^{-1} are coloured white.

steadily increasing towards the poles to just over 1 mm day^{-1} .

In figure 4.5, tephigrams are presented showing the control run's 4-year mean-state and variability at 4 different latitudes, which characterise the ITCZ, sub-tropical highs, mid-latitude storm track, and polar regions.

As one would expect, the tropical profile is moist and conditionally unstable up to around 200

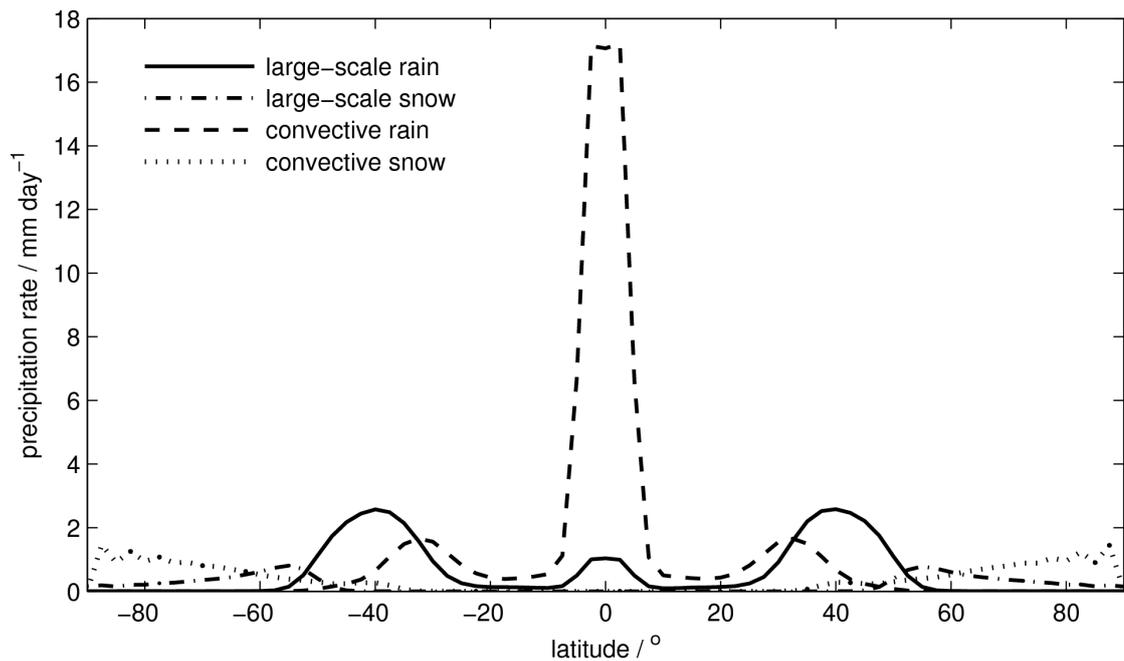


Figure 4.4: Zonal mean 4-year mean precipitation rates for the Default UM aqua-planet; (solid) large-scale rain, (dashed) convective rain, (dot-dashed) large-scale snow, and (dotted) convective snow.

hPa, consistent with the deep convection which the model produces in this region. It has very little temperature variability, consistent with Weak Temperature Gradient Theory (e.g. Charney 1963). The sub-tropical profile has a strongly stable layer from about 850hPa up to 700 hPa, above which the sounding is very dry, consistent with slow radiative subsidence of air which has ascended in the ITCZ convection. The mid-latitudes exhibit the greatest variability in tropospheric temperatures, consistent with the frequent passage of mid-latitude cyclones and their associated fronts, warm-sectors and cold air outbreaks. As inferred from the presence of convective showers over the unrealistically warm polar region, the polar tephigram has slight conditional instability; for the 25th percentile of the temperature profile, the non-dilute ascent from near surface (dotted line) reaches 500 hPa.

Figure 4.6 shows the zonal-mean precipitation, surface moisture flux, and net surface moisture budget. On average, moisture is net extracted from the sea-surface in the sub-tropical regions between 5° and 30° latitude by up to 5 mm day^{-1} , with a large net rain-out of 15 mm day^{-1} in the ITCZ. Most of the moisture added to the atmosphere in the sub-tropics is clearly advected in to the tropical rain-belt, but some is advected in to the storm-track where it is rained out in mid-latitude

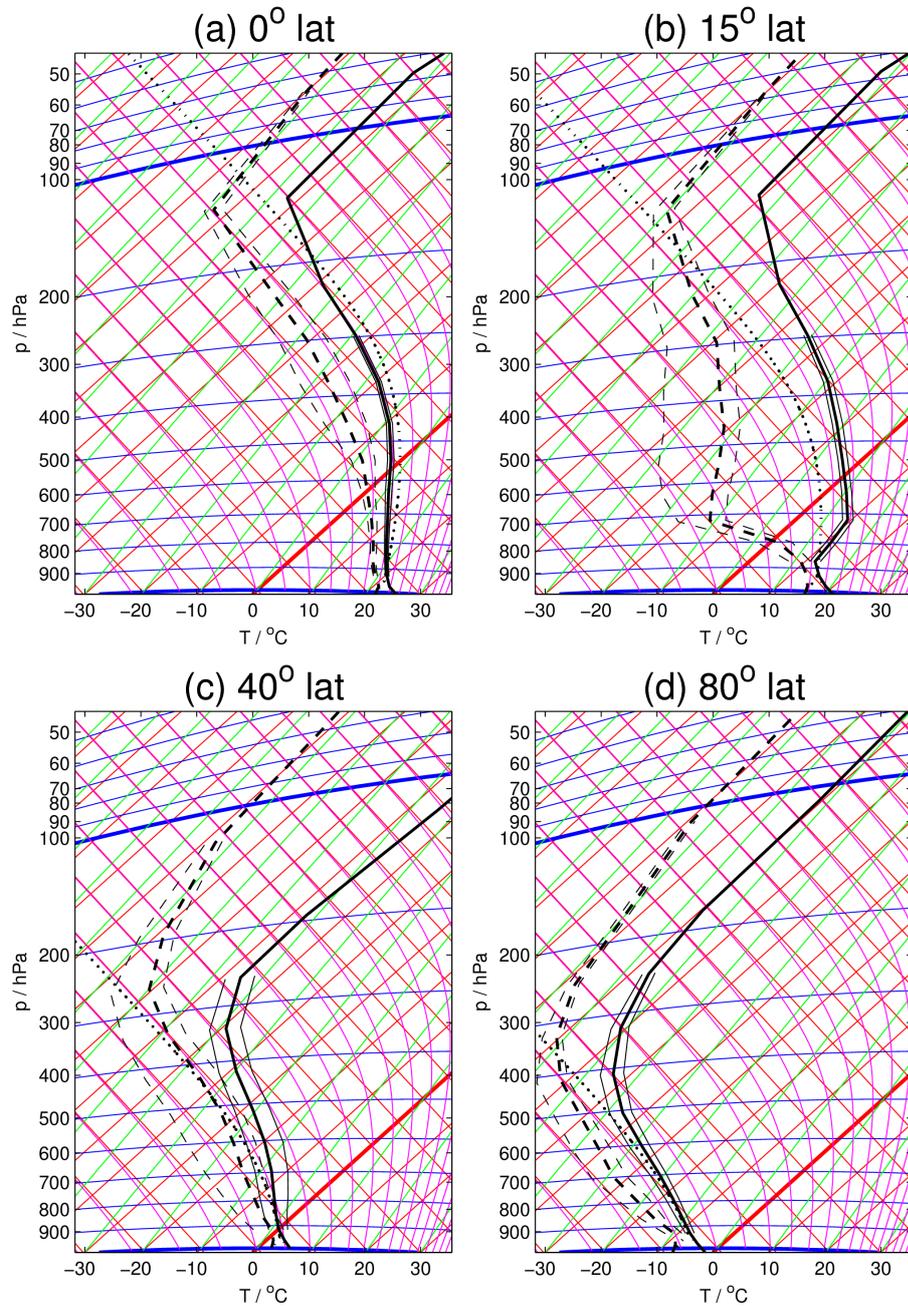


Figure 4.5: Tephigrams for the Default UM aqua-planet, at four representative latitudes; (a) tropics, (b) sub-tropics, (c) mid-latitude storm track, and (d) polar. Lines represent (solid) dry-bulb temperature, (dashed) dew-point temperature, and (dotted) a non-dilute moist parcel ascent from the lowest model level. Thick lines are the 4-year-averaged zonal-mean, thin lines show the 25th and 75th percentiles of 6-hourly data throughout the run.

cyclones. Poleward of 60° , the mean surface moisture flux and precipitation rate nearly cancel, so that the broad polar regions are on average in local moist equilibrium. The surface moisture flux has its greatest variability just on the equatorward side of the storm-tracks. However, the variability in the net moisture budget is dominated by that in precipitation everywhere except in the sub-tropics, where little precipitation occurs.

The area-weighted mean net surface moisture budget shown in figure 4.6 should in principal integrate to zero for a long run with no systematic drift in atmospheric moisture. However, for this simulation it integrates to $+0.116 \text{ mm day}^{-1}$. This global mean precipitation excess would rain out all of the moisture in the atmosphere in well under a year. But over the 4 year run this has clearly not happened, and analysis confirmed that no systematic drift occurred in the global-mean atmospheric moisture content. Evidently the model dynamics don't quite conserve moisture, generating a slight excess through numerical error in the advection scheme. However, this is of little concern in the aqua-planet as universal contact with the sea surface has prevented any model drift in moisture from occurring.

The net heat budgets at the top-of-atmosphere and the surface are shown in figure 4.7, along with the difference between the two, which quantifies the net heat budget for the atmosphere. Contributions from SW and LW radiation and surface fluxes are plotted, as well as their variability.

At top-of-atmosphere, the difference between incoming SW and outgoing LW radiation yields a net flux of heat in to the planet everywhere between $\pm 45^\circ$ latitude), whilst the polar regions strongly lose heat to space. There are considerable variations in both incoming SW and outgoing LW in the tropics, associated with variations in cloud. Both fluxes reduce towards the ITCZ, and become more variable in time. But the net radiation at top-of-atmosphere is flat across the tropics with relatively little time-variability, suggesting that the modelled effects of the tropical clouds on SW (increasing reflected radiation) and LW (reducing outgoing radiation) nearly cancel each-other out on average.

At the surface, the balance of SW and LW radiative fluxes, and sensible and latent fluxes of heat from the underlying global ocean, yields a net flux of heat in to the sea across most of the planet. Exceptions are the near-neutral sub-tropics and a strong net loss of heat by the polar seas. In the tropics, the surface primarily loses heat by imparting moisture (latent heat) to

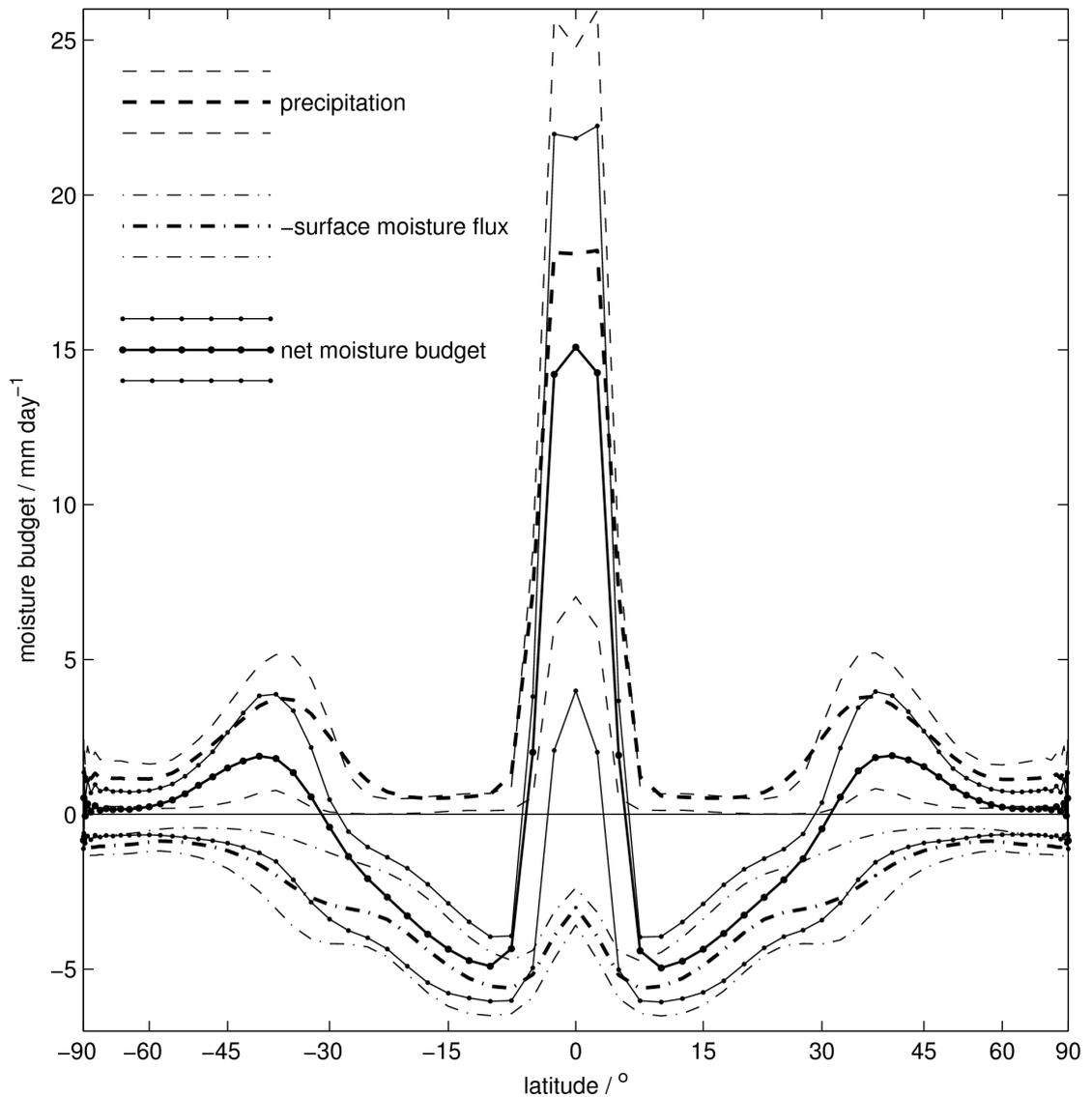


Figure 4.6: 4-year zonal mean moisture budget components in the control run; (dashed) precipitation, (dot-dashed) surface moisture flux, and (solid with markers) net moisture budget. Thick lines denote the mean, thin lines show 25th and 75th percentiles based on daily mean data. The latitude axis has been scaled by an area weighting so that the area enclosed by the curves is representative of surface-integrated flux, as in later budget figures. The sign convention followed is to plot fluxes of moisture *out of* the atmosphere positive.

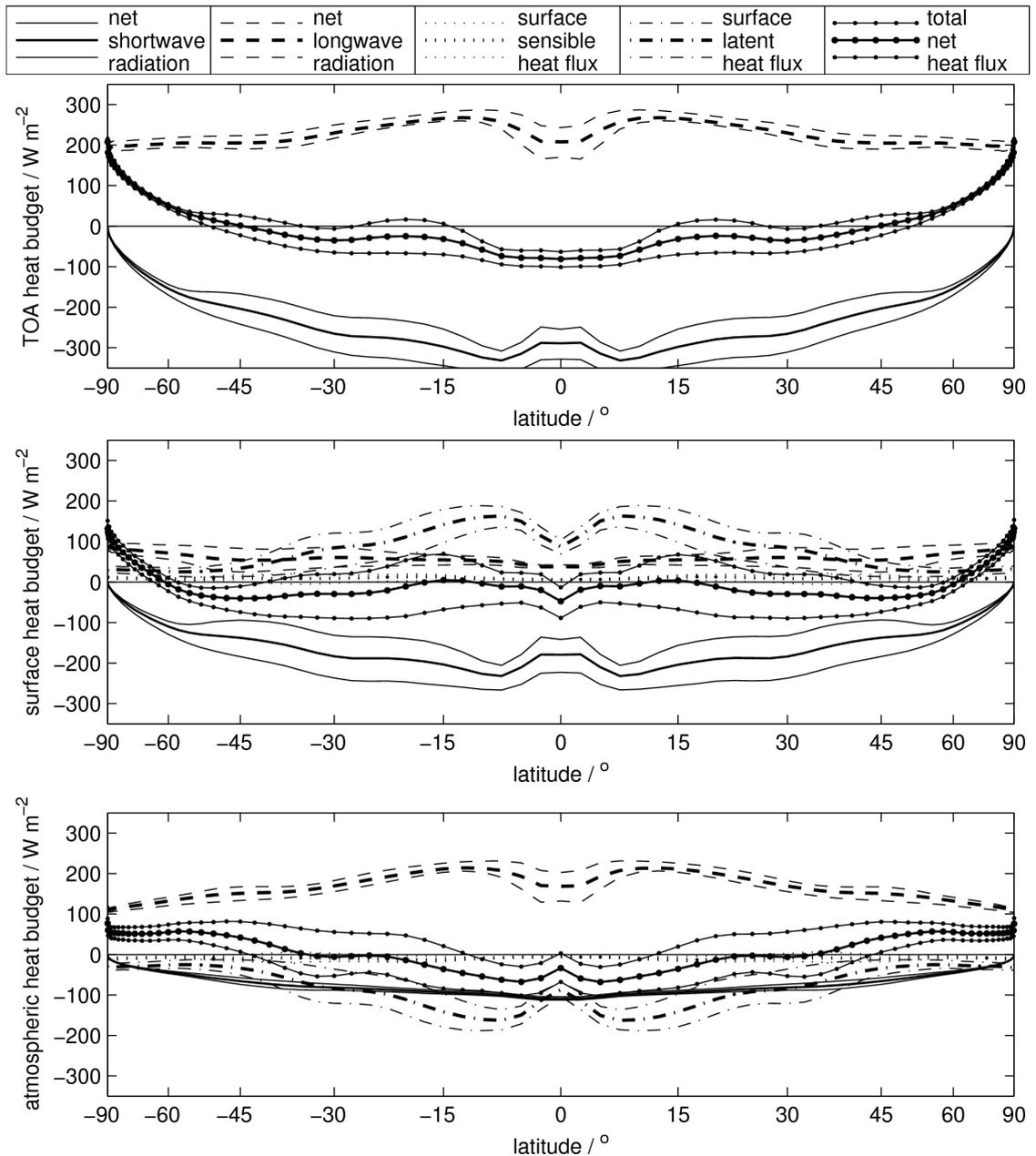


Figure 4.7: 4-year zonal mean heat budgets; (top) top-of-atmosphere radiation, (middle) surface heat fluxes, and (bottom) atmospheric net budget, equal to the difference of the upper two plots. Lines show (solid) net SW radiation, (dashed) net LW radiation, (dot-dashed) latent heat flux, (dotted) sensible heat flux, and (solid with markers) the total net heat flux, as labelled in the legend at the top of the figure. Thick lines denote the mean, thin lines show 25th and 75th percentiles based on daily mean data. The sign convention followed is to plot *upward* fluxes positive for the top-of-atmosphere and surface budgets, and plot fluxes *out of* the atmosphere positive for the atmospheric budget.

the atmosphere, whilst the polar seas mainly lose heat by LW radiation. There is a pronounced minimum in the latent heat flux on the equator, due to the lighter winds and higher humidities in the ITCZ. This dip more-than offsets the reduction in surface SW associated with ITCZ cloud, so there is a considerable net flux of heat in to the surface at the equator.

The difference between the top-of-atmosphere and surface heat budgets (bottom panel of figure 4.7) yields a net flux of heat in to the atmosphere of around 50 W m^{-2} equatorwards of 20° , and a net flux of similar magnitude out of the atmosphere polewards of about 35° . Interestingly, there is very little time-variation in the amount of SW radiation absorbed by the atmosphere, despite SW being the biggest source of variability in the top-of-atmosphere and surface heat budgets. This suggests that the variations in cloud (which drive the large variations in the amount of SW reflected out to space or absorbed by the surface) have little effect on the modelled SW absorption by the atmosphere. The large time-variations in the atmosphere's net heat budget are mainly driven by variations in surface latent heat flux and LW emission.

Whilst the atmosphere over the polar regions was found to be near local moist equilibrium, it is clearly not in local thermodynamic equilibrium, with a net heat-loss of around 50 W m^{-2} . Interestingly, this implies a strong meridional heat transport by the atmosphere in to the polar regions, despite there being rather flat temperature gradients and very little implied meridional moisture transport there. As is evident in later plots, the model is in fact transporting heat to the poles via convergence and subsidence at mid-levels.

Assuming the model to be thermodynamically balanced, the zonal-mean net heat budgets should be balanced by meridional heat transports at each latitude, and integrate to zero globally. However, the area-weighted global means of the heat budgets were computed and found not to be in global balance during the simulation. At top-of-atmosphere there was a global mean net radiative forcing of $+11.1 \text{ W m}^{-2}$, whilst at the surface there is a net mean global downwards heat flux of $+12.4 \text{ W m}^{-2}$.

It is unsurprising that the surface heat budget remains globally out of balance during the simulation, as the Sea Surface Temperatures are prescribed. If the simulation included an interactive ocean component instead, a net warming of the ocean would occur until the surface heat budget equilibrated. The model also produced a difference between the global mean top-of-atmosphere

radiative forcing and the global mean surface heat flux, implying that the atmosphere lost heat during the simulation at a mean rate of 1.2 W m^{-2} . Analysis found no corresponding drift in temperature in the model, so the atmosphere must contain a small erroneous heat source. As was noted earlier, the model atmosphere also rained out slightly more moisture than was evaporated in to it during the run, at a global mean rate of $0.116 \text{ mm day}^{-1}$, due to numerical error. We would expect a spurious latent heating term associated with this, but the latent heat equivalent of this numerical moisture excess is 3.3 W m^{-2} , somewhat greater than the heat excess of 1.2 W m^{-2} . So the model clearly doesn't quite conserve heat. But as with the moisture budget imbalance, this is of little consequence in a fixed SST aqua-planet framework.

In summary, the aqua-planet control simulation appears to capture the basic features of the earth's atmospheric global circulation. However, it differs substantially from the real world in that the tropical circulation is more closely confined towards the equator, the mid-latitude storm track is closer to the equator and interacts more strongly with the tropical circulation, and the broad polar regions are characterised by slack gradients and convective showers (due to the somewhat unrealistic ice-free surface with prescribed 0°C flat temperature distribution there). The moisture budget reflects the equatorially confined nature of the model circulation; in the sub-tropics, evaporation exceeds precipitation polewards of 5° latitude, whilst in the polar regions evaporation and precipitation are in local balance polewards of 60° . The heat budgets during the simulation give insights in to the radiative impacts of cloud in the model, and the implied meridional heat transports. Variations in tropical cloud appear to have little net effect on the top-of-atmosphere radiation balance or on SW absorption by the atmosphere. There were found to be small discrepancies in the moisture and heat budgets, as the model dynamics don't quite conserve these. However, this should be of little concern, as the prescribed SST pattern prevents any substantial model drift.

4.3.2 COMPARISON OF MOISTURE BUDGETS AND RAINFALL VARIABILITY

In this section, the effects of the two stochastic parameterisations on the hydrology of the model atmosphere are investigated.

First, we consider differences in the moisture budgets of the stochastic simulations relative

to the control run (the moisture budget for the control run was presented in figure 4.6). Differences in precipitation, surface moisture flux, and net moisture budget (precipitation minus surface moisture flux) are shown in figure 4.8. The primary differences between the stochastic runs and the control are hemispherically asymmetric shifts in rainfall near the equator. The T-test suggests that only an isolated point in the tropics in each simulation differs significantly from the control run. Given the large number of points being considered, such a sparse scattering of “significant” points is consistent with the expected type I error-rate of the test. Also, given the symmetry of the modelling framework, any systematic biases induced by the stochastic schemes ought to possess hemispheric symmetry.

The apparent differences in rainfall in the tropics are the residual of long-timescale variability in the ITCZ, which happens to follow a different evolution in each run. Hovmoller plots of tropical rainfall variability over the first year of the control run are shown in figure 4.9. Variability in the total amount of rainfall in the ITCZ is dominated by equatorially trapped Kelvin waves, which travel around the globe from West to East in about 30 days. However, the North-South asymmetry of the ITCZ rainfall is dominated by much slower modes. Northward or Southward excursions of the ITCZ sometimes persist over broad sectors of the tropics for up to 40 days, exhibiting a gradual Westward propagation, circling the globe in 60–80 days. This is in the opposite sense to the Kelvin waves, and with a much lower phase speed.

Given the long timescales of the ITCZ’s North-South mode, the hemispherically asymmetric differences between runs of up to 0.5 mm day^{-1} are not surprising. The effective sample size used in the T-test for these differences has identified the long timescales, and accordingly assigned low statistical significance to most of these differences.

An investigation in to the Kelvin wave and North-South modes which dominate the model’s tropical rainfall variability was performed, but did not detect any systematic differences in these modes between the stochastic aqua-planet simulations and the control run. Correspondingly, no significant difference in the overall variance of tropical precipitation was found. This is evident in the power spectra of tropical rainfall, shown in figure 4.10. The spectral peak at frequencies corresponding to around 10 days is due to the Kelvin wave activity evident in figure 4.9. The only systematic difference between the tropical rainfall power spectra is the slight increase in variance at higher frequencies in the Multiplicative Noise run. The precipitation diagnostics were

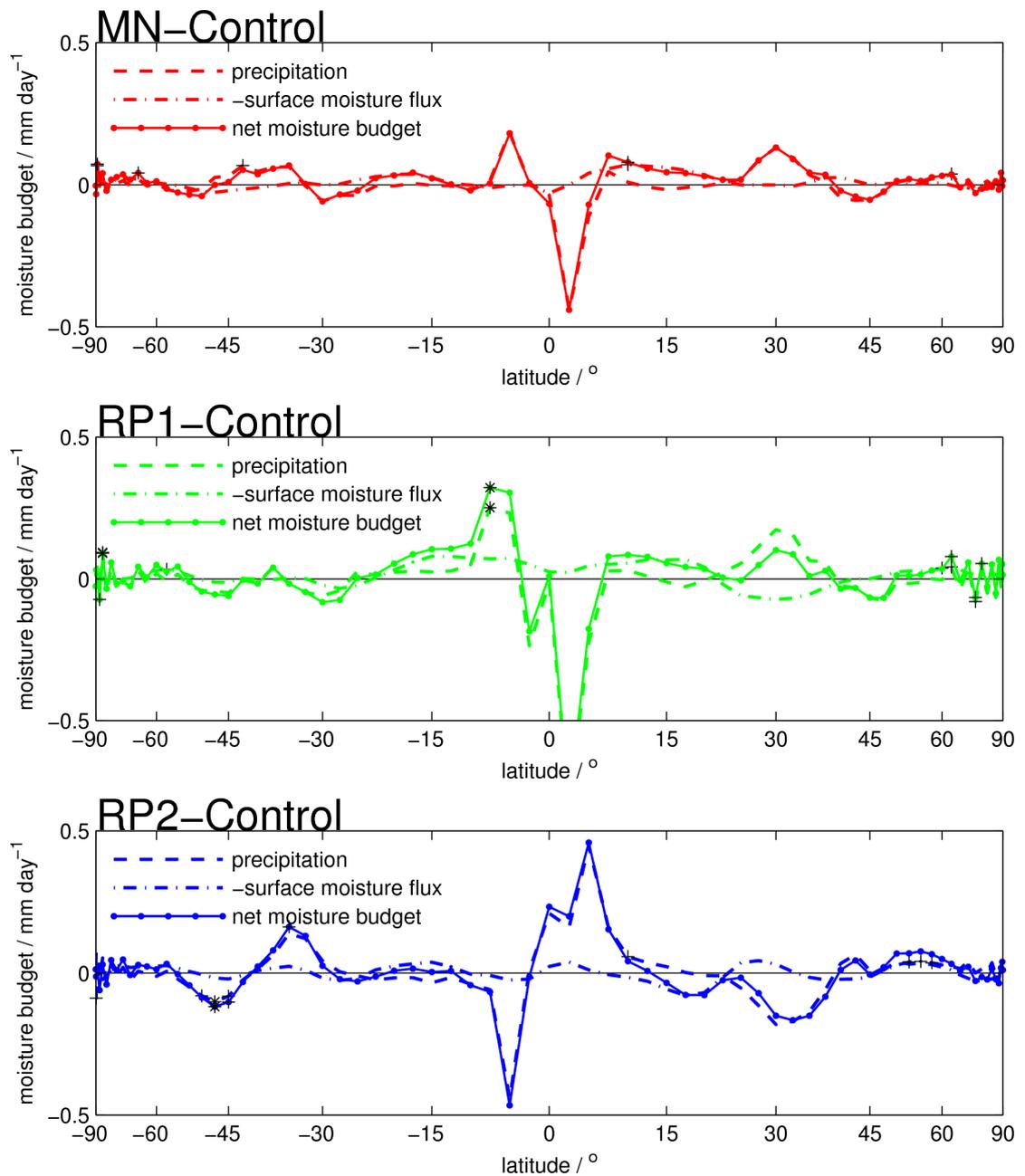


Figure 4.8: Time-averaged zonal mean moisture budget differences of stochastic runs relative to the control run; (red) the simulation with Multiplicative Noise, and (green, blue) the two runs with Random Parameters; (dashed) precipitation, (dot-dashed) surface moisture flux, and (solid with markers) net moisture budget. Black cross-hairs denote statistical significance of the differences, at the 5% level, whilst stars indicate significance at the 1% level. The sign convention followed is as for the control run in figure 4.6; increased precipitation is positive.

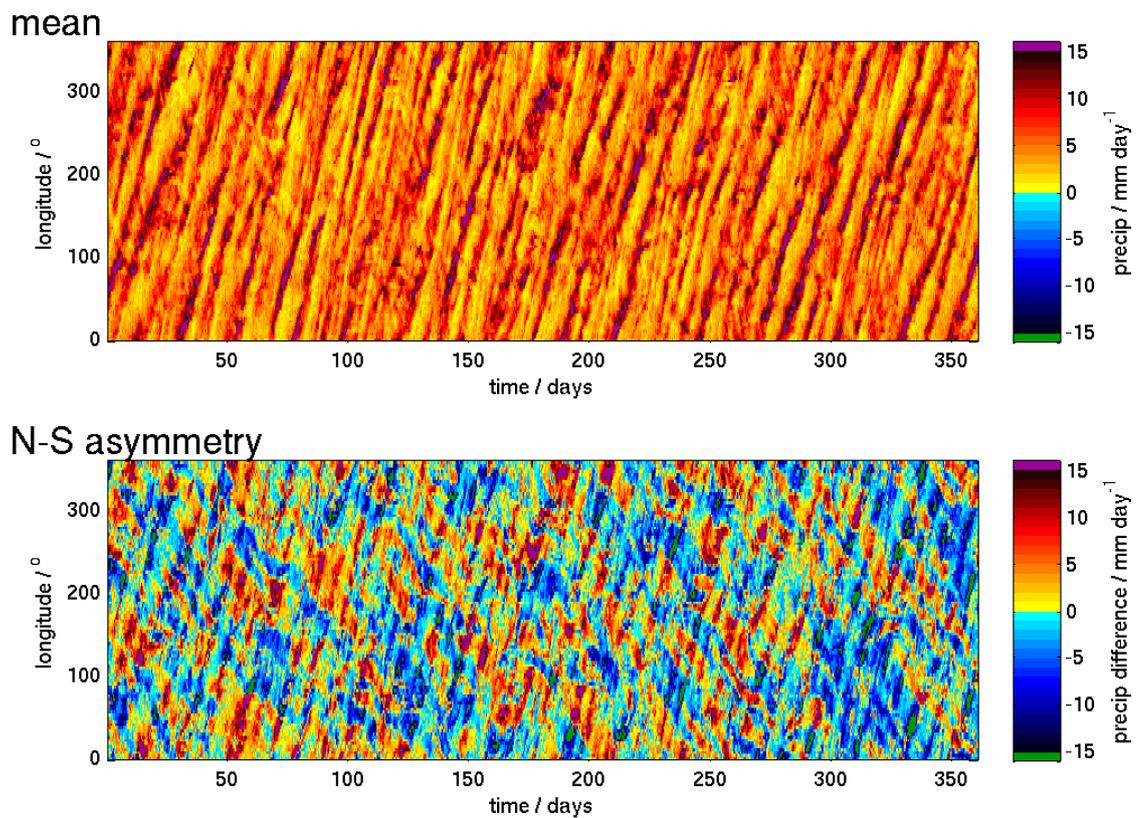


Figure 4.9: Hovmoller plots of rainfall between $\pm 15^\circ$ latitude; (top) total mean rainfall, and (bottom) mean rainfall North of the equator minus mean rainfall South of the equator.

directly perturbed in this run, to ensure a consistent moisture budget when the stochastic scheme perturbed the parameterised tendencies in moisture. The increase in spectral power in the MN run is consistent with this direct white-noise-like forcing and so does not indicate any dynamical or physical response of the model to the multiplicative noise scheme.

These results are in contrast to Lin & Neelin (2000), who found that introducing a simple stochastic parameterisation scheme substantially increased tropical rainfall variability, including its slower modes. They introduced stochastic variability, modelled as an autoregressive process, to the CAPE input to the convection parameterisation. In their study, the response of tropical rainfall variability was highly sensitive to the autocorrelation timescale of the applied stochastic forcing; major changes in the precipitation power spectra occurred for their experiment with the stochastic timescale set to 1 day, but the response was much smaller for timescales of 2 hours or 20 minutes. A similar stochastic element is applied in the Random Parameters scheme studied

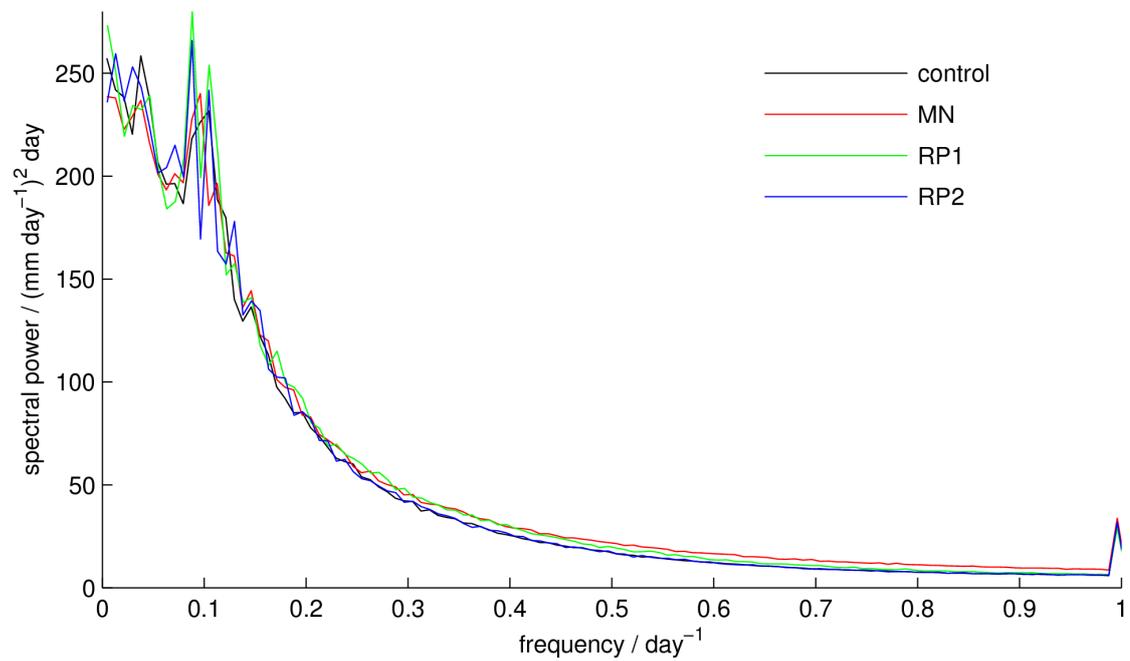


Figure 4.10: Mean power spectrum for rainfall between $\pm 15^\circ$ latitude, for (black) the aqua-planet control simulation, (red) the Multiplicative Noise run, and (green, blue) the two Random Parameters runs. Spectral power has been normalised so that its integral over the frequency domain is equal to the total variance.

here, through variation of the UM convection scheme's CAPE removal timescale parameter. The RP scheme updates parameter values once every 3 hours, but with considerable autocorrelation between subsequent updates (see the description of the scheme in chapter 2), giving it a timescale most consistent with Lin & Neelin's 1-day stochastic timescale run, which gave dramatic changes in rainfall variance.

It is interesting that the Lin & Neelin (2000) stochastic convection scheme induced such a strong response in their experiment, whilst the RP scheme, with a similar stochastic timescale, and extensive perturbations to convection and other sub-grid processes, induces no systematic response in the rainfall variability. This maybe due to differences in the host models used; Lin & Neelin's control simulation had very little rainfall variability at low frequencies and required the stochastic scheme to activate the slow modes, whilst the aqua-planet simulation in the present study exhibits slow modes with considerable amplitude without a stochastic scheme.

Scinocca & McFarlane (2004) showed that tropical rainfall variability in a 3D simulation was

strongly connected to the partition of rainfall production between the convective and layer cloud parameterisations. The precipitation differences between the stochastic runs and the control run in this study are shown divided in to their components from the UM convection and large-scale precipitation schemes in figure 4.11.

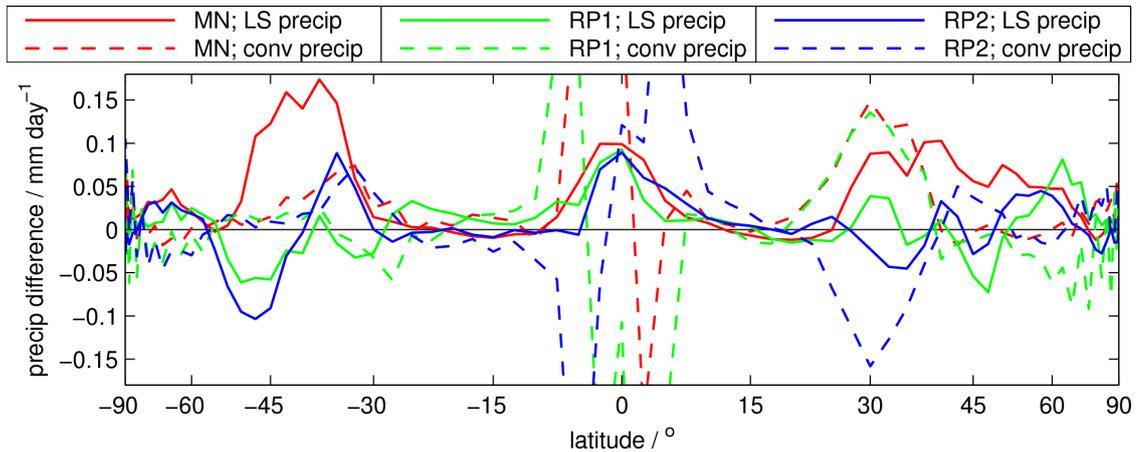


Figure 4.11: Differences in precipitation components; (solid) from the large-scale precipitation scheme, and (dashed) from the convection scheme, for (red) MN-Control, (green, blue) RP-Control, as labelled in the legend above the figure.

Whilst the convective rainfall differences are dominated by the large hemispherically asymmetric shifts noted in figure 4.8, there is a consistent increase in large-scale precipitation in the tropics for all three stochastic runs. Note that as the separate large-scale and convective components of precipitation were only output as time-averaged zonal means, the statistical tests used in other figures could not be applied here. Comparison with the control run's precipitation components shown earlier (figure 4.4) shows that the 0.1 mm day^{-1} additional large-scale rainfall in the stochastic runs amounts to about a 10% increase in large-scale precipitation. This, and the consistency of the increase between all three stochastic runs, suggests this is likely to be a significant, systematic effect of increased high frequency variability. But it makes a negligible contribution to the total rainfall in the tropics.

4.3.3 COMPARISON OF RADIATION, HEAT BUDGETS AND CLOUD

In this section, differences between the heat budgets in radiative and surface fluxes in the aqua-planet simulations are presented, with the main differences associated with changes in cloud

which are also presented. The control run's heat budgets were shown in figure 4.7, and differences of the stochastic runs relative to the control run are shown in figures 4.12, 4.13 and 4.14 below.

In the Multiplicative Noise run (figure 4.12) there is a significant increase in the net downwelling SW radiation by around 3 Wm^{-2} in the sub-tropics, at top-of-atmosphere and at the surface. There is also a significant increase in surface upwelling LW (by about 1 Wm^{-2}). But these systematic changes evident in the sub-tropical radiation budget do not have a significant effect on the atmosphere's net heat budget. Note that the control run's heat budget showed remarkably little variability in net SW absorption by the atmosphere, suggesting that variations in cloud have relatively little effect on it, despite dominating the variability in the surface and top-of-atmosphere radiation budgets. The increases in net downwelling SW and net upwelling LW are consistent with a systematic decrease in sub-tropical cloud in the MN run. They also give a net positive radiative forcing on the surface relative to the control run; globally averaged, this amounts to $+0.9 \text{ Wm}^{-2}$.

Differences in the atmosphere's heat budget appear to be dominated by differences in the latent heat flux, which the statistical tests indicate can be accounted for by internal variability. There is however a significant decrease in the sensible heat flux in the sub-tropics. No corresponding significant differences in near-surface temperature or winds (not shown) were found between the runs, suggesting that the sensible heat flux has changed as a direct response to the multiplicative noise scheme.

In the Random Parameters runs (figures 4.13 and 4.14), there are significant differences in the heat budgets relative to the control run in the extra-tropics. Both RP runs give a significant decrease in the net downwelling SW at the surface and top-of-atmosphere at latitudes around $\pm 50 - 75^\circ$, by around $1-2 \text{ Wm}^{-2}$ (there are substantial differences in the extent and magnitude of this effect between the two RP runs). At the surface, these decreases in SW are more than compensated for by significant decreases in net upwelling LW and surface fluxes in the same regions, yielding a net positive radiative forcing on the surface in those regions (relative to the control run).

With the significant SW-change forcing in the upward sense at top-of-atmosphere, and significant LW and surface flux-change forcings in the downward sense dominating the surface budget,

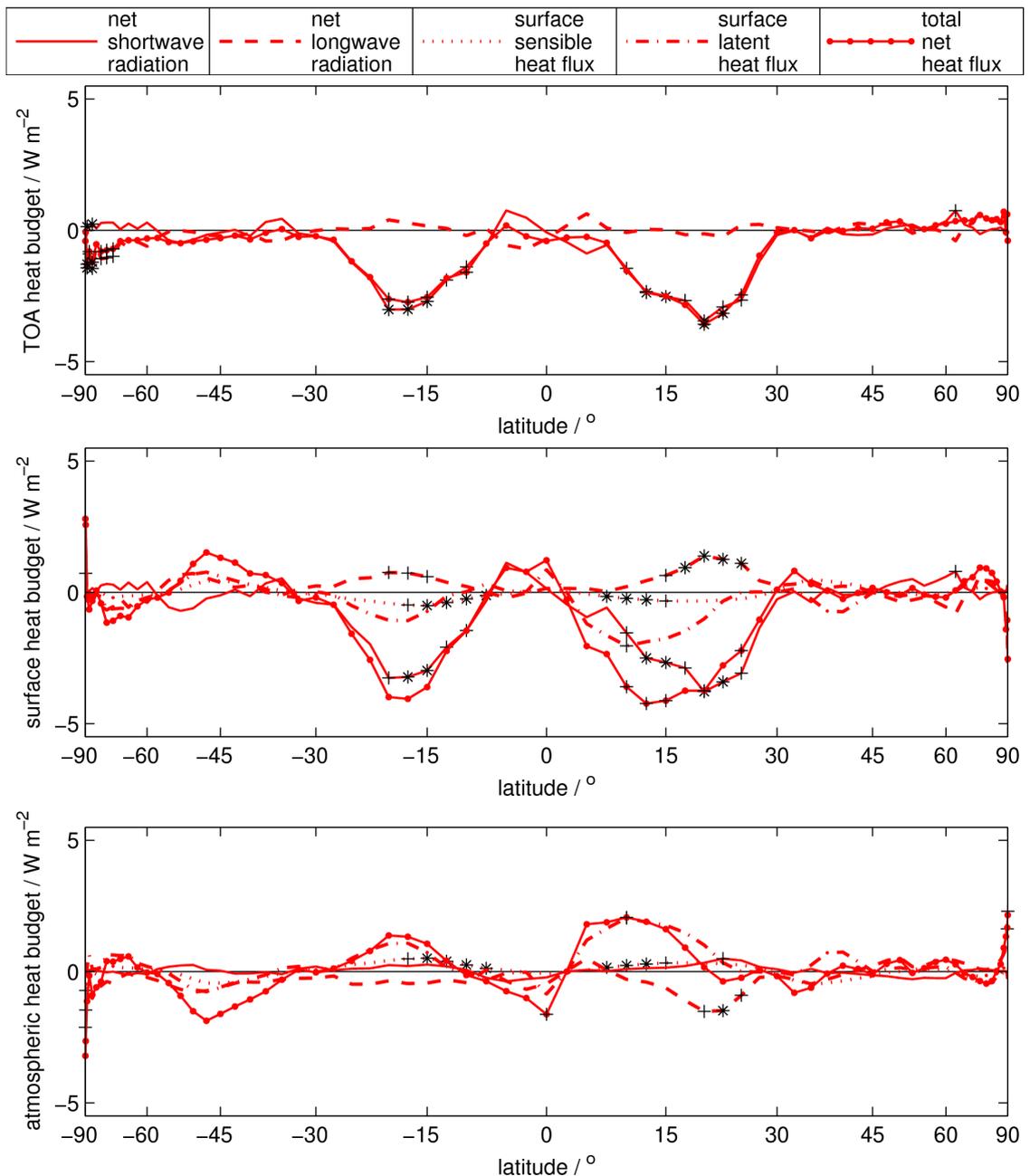


Figure 4.12: Time-averaged zonal mean heat budget differences of the Multiplicative Noise run relative to the control run; (top) top-of-atmosphere radiation, (middle) surface heat fluxes, and (bottom) atmospheric net budget, equal to the difference of the upper two plots. Lines show (solid) net SW radiation, (dashed) net LW radiation, (dot-dashed) latent heat flux, (dotted) sensible heat flux, and (solid with markers) the total net heat flux. The sign convention followed is to plot *upward* fluxes positive for the top-of-atmosphere and surface budgets, and plot fluxes *out of* the atmosphere positive for the atmospheric budget. Black cross-hairs denote statistical significance of the differences at the 5% level, whilst stars indicate significance at the 1% level.

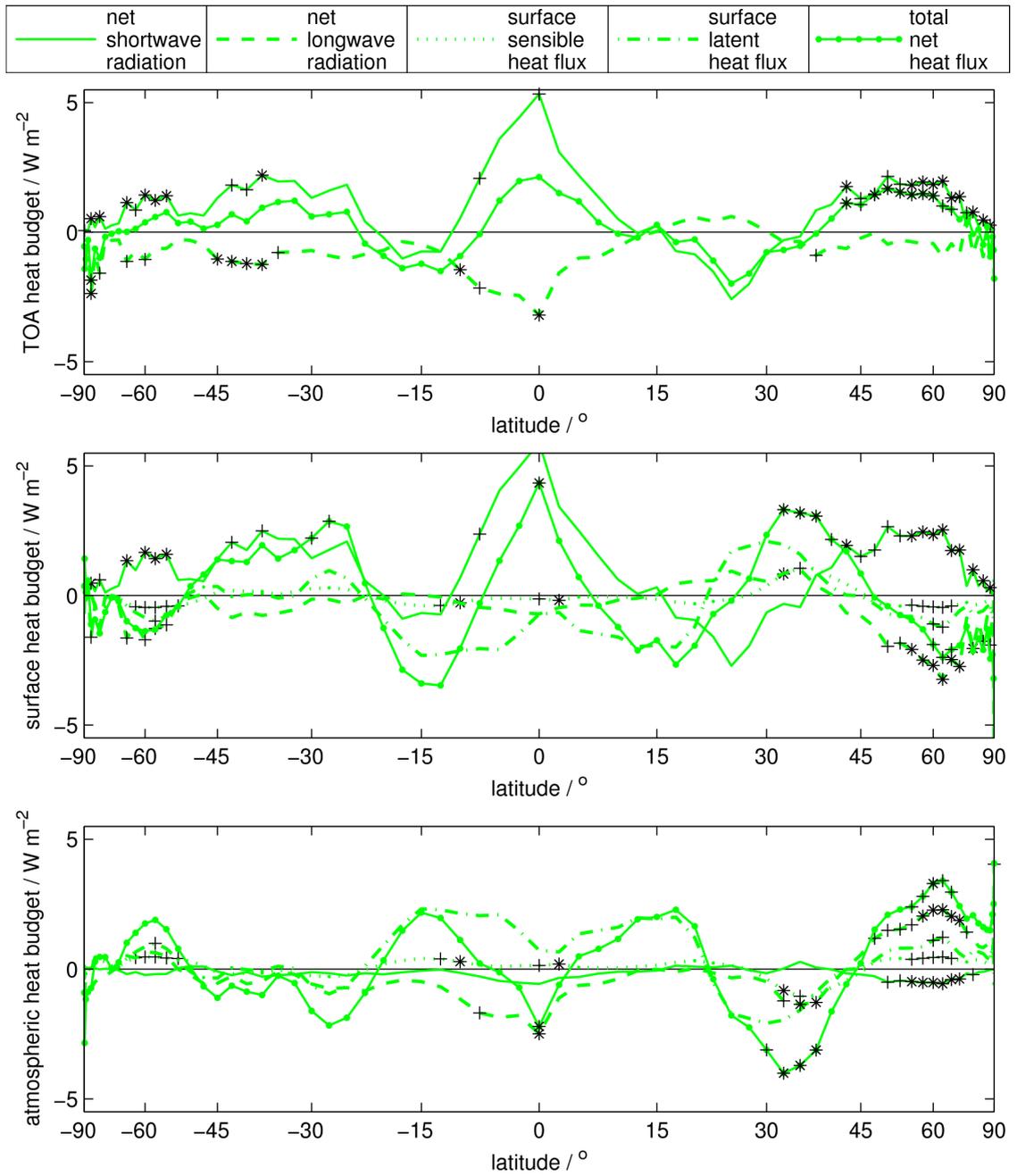


Figure 4.13: As figure 4.12, but for differences of the first Random Parameter run relative to the control run.

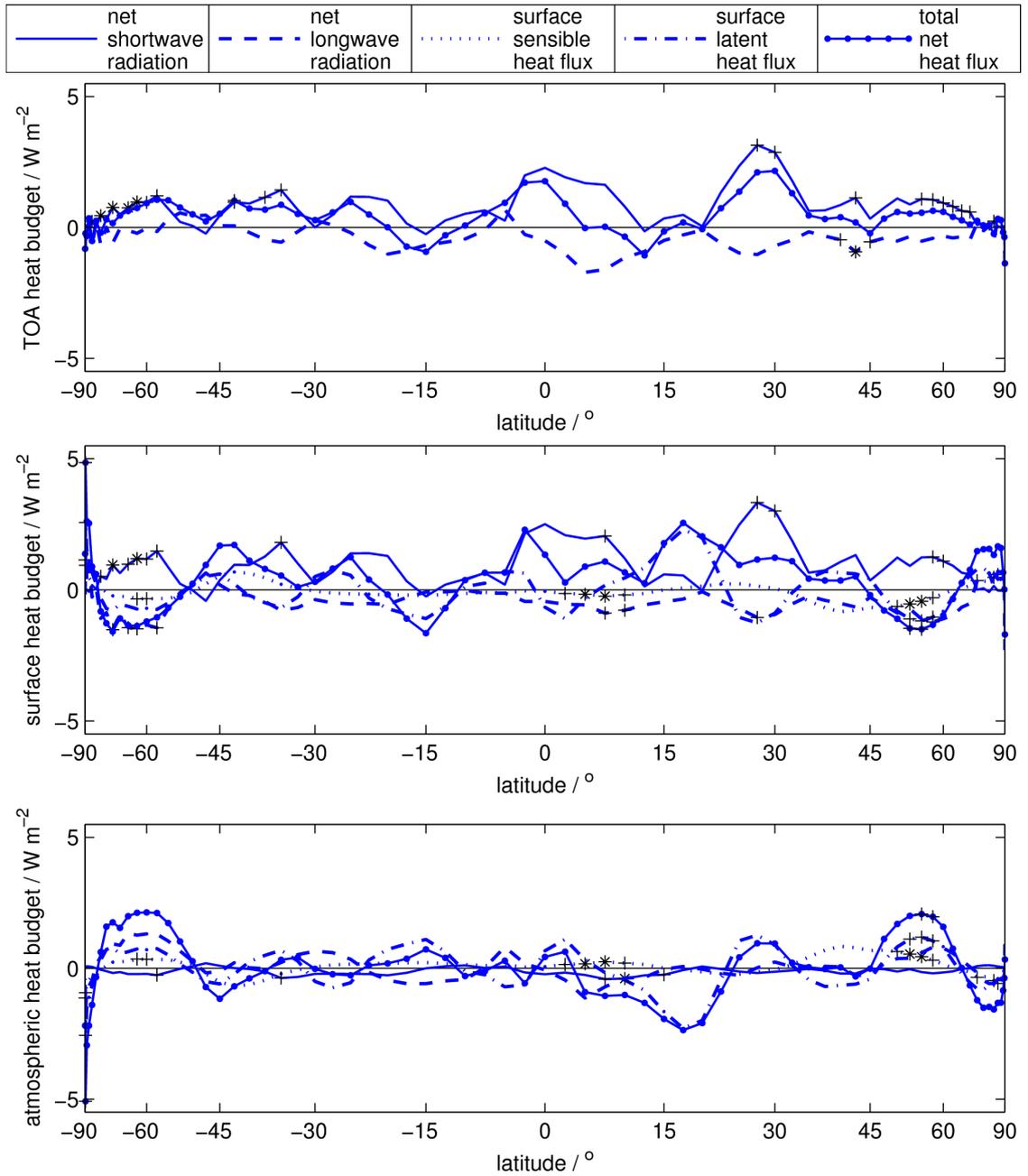


Figure 4.14: As figure 4.12, but for differences of the second Random Parameters run relative to the control run.

a significant loss of heat from the atmosphere at the corresponding latitudes would be expected. The atmospheric heat budgets do show net outward flux differences of about 2 Wm^{-2} at around $\pm 60^\circ$ latitude in both RP runs, though they are not consistently statistically significant. The Random Parameter simulations also appear to give a reduction in net down SW over the equator, though the two RP runs disagree over the magnitude and significance of this effect. More robust statistics (i.e. longer runs, or more of them) are needed to verify whether these differences are systematic effects of the RP scheme.

The changes seen in the radiation budgets when the stochastic schemes are applied are most likely driven by changes in cloud. To investigate this in more detail, model cloud variables are presented. Figure 4.15 shows total layer-cloud water content (the sum of the liquid and frozen components), and figure 4.16 shows the total convective cloud water content. 4-year zonal means and variances for the control run are shown, along with differences of the stochastic runs relative to the control.

All three stochastic runs give substantial increases in layer cloud-water, by of the order 5%, across large areas of the model domain. The increase has been identified as significant in all three over the ITCZ, in the storm-tracks at around $\pm 40^\circ$ latitude, and at around $\pm 60^\circ$. The increase in layer cloud condensate over the tropics is consistent with the increase in large-scale precipitation seen in all three stochastic runs. The Multiplicative Noise run also gives a significant decrease in layer cloud of a similar magnitude in the sub-tropical boundary layer.

The MN scheme also has a dramatic effect on the convective cloud, decreasing it by around 5% throughout the sub-tropics, and in the boundary layer further poleward. The convective cloud response in the RP runs is less clear; there are significant increases in the tropical upper troposphere and in the mid-troposphere at around $\pm 40^\circ$ and $\pm 70^\circ$, but the two RP runs give very different extents, magnitudes and levels of significance for these increases.

Note from the colour-scales of the figures that in the control run the mean convective cloud water content is typically about 4 times the layer cloud content, and the changes occurring in convective cloud have typical magnitudes about 3 times greater than those in layer cloud. So the changes in convective cloud will likely have greater radiative impact than those in layer cloud.

Comparing the cloud changes shown in figures 4.15 and 4.16 with the changes in radiation

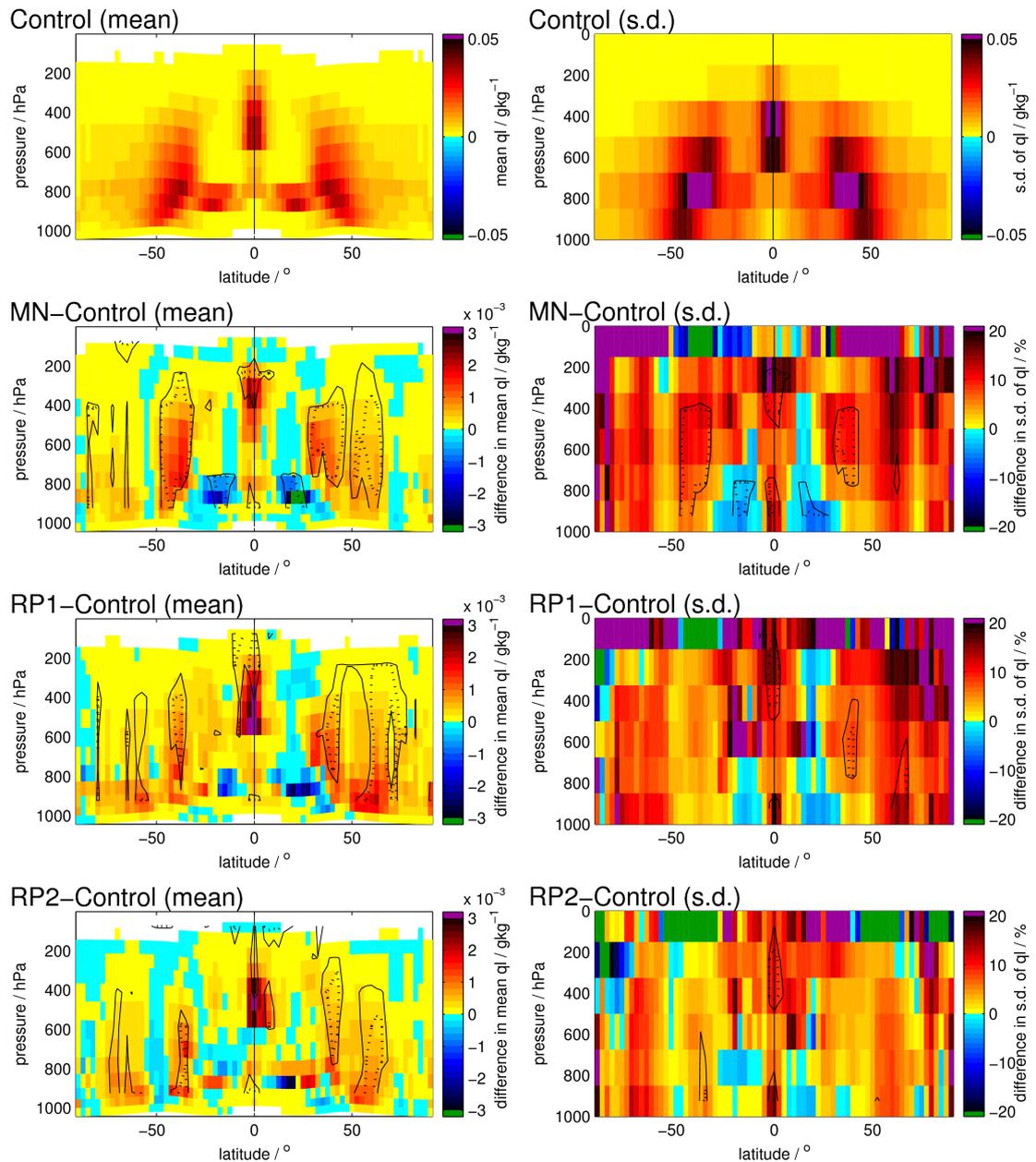


Figure 4.15: Layer Cloud water content q_l ; (left) time-averaged zonal means and (right) standard deviations; for the control run, and differences of the stochastic runs from the control run, as labelled. Black contours denote statistical significance of the differences, at the (solid) 5% and (dotted) 1% levels.

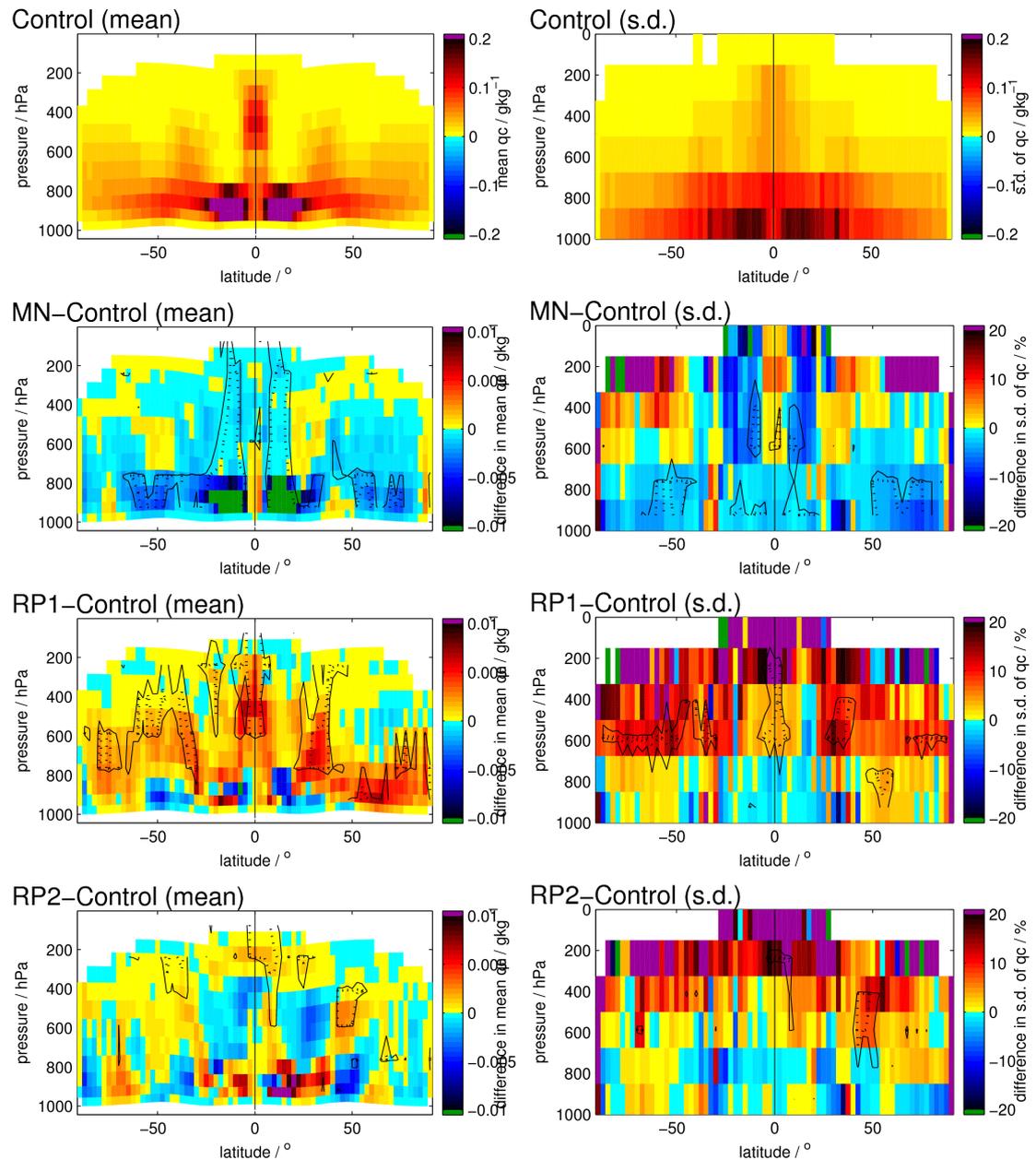


Figure 4.16: Convective Cloud water content q_c ; (left) time-averaged zonal means and (right) standard deviations; for the control run, and differences of the stochastic runs from the control, as labelled. Black contours denote statistical significance of the differences, at the (solid) 5% and (dotted) 1% levels.

budgets shown in figures 4.12, 4.13 and 4.14, the radiative changes appear consistent with the cloud changes. The significant increases in net downwards SW in the sub-tropics in the MN run are due to the reduction in boundary-layer cloud there (the major contribution coming from convective cloud, with some effect from layer cloud). The decreases in downward SW in the extra-tropics in the RP runs are due to increases in both layer cloud and convective cloud at a range of depths in the same regions. Polewards of $\pm 30^\circ$ latitude, the changes in cloud in the MN run appear to have no effect on the radiation budgets; presumably the effects of the increase in mid-tropospheric layer cloud and the decrease in low-level convective cloud roughly cancel.

The stochastic schemes also induce significant changes in the variance of the cloud variables. These changes are broadly consistent with the changes in the means (for a noisy variable with a frequently-realised zero base-value, the mean and standard deviation both indicate the typical magnitude of departures from zero), but there are some less trivial changes. In the MN run, there is a significant increase in the variance of convective cloud in the equatorial mid-troposphere, where there is a slight decrease in its mean. And in the RP runs, the convective cloud variance increases in the mid-troposphere in parts of the extra-tropics. These changes are consistent with the stochastic schemes increasing the variability of the height reached by convective cloud tops in those regions.

To further investigate the convective activity in the runs, figure 4.17 shows time-averaged zonal mean convective heating and moisture tendencies from the control run, and their differences between the stochastic runs and the control. The statistical test for the significance of the differences in mean-state could not be applied for the convective tendencies, as only the time-averaged zonal means of these fields were output.

The Multiplicative Noise run gives a 5% reduction in convective moistening of the lower troposphere in the sub-tropics, and some increase in convective heating / drying over the equator. The sub-tropical changes in convective activity are consistent with the reduction in shallow convective cloud there, which was found to be statistically significant (figure 4.16). Whilst analysis of the moisture profiles did not identify any statistically significant changes in mean humidity in the MN run, the reduction in the convective moisture flux out of the boundary layer in the sub-tropics likely explains the coincident reductions in layer cloud in that run (figure 4.15). It is not clear whether the changes over the equator are significant; they largely do not coincide with

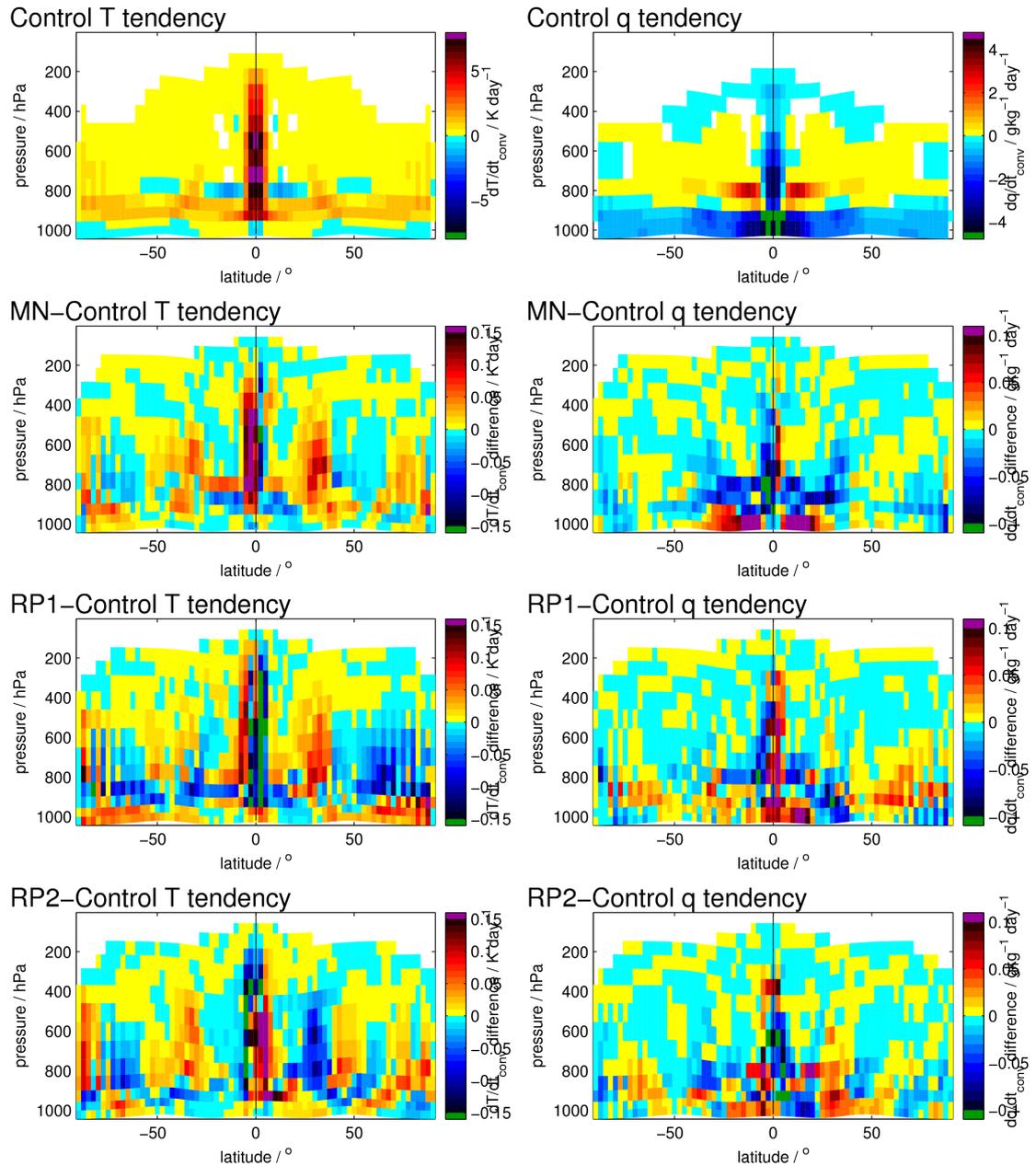


Figure 4.17: Time-averaged zonal mean convective tendencies in (left) temperature, and (right) specific humidity. Top two panels show the mean-fields for the 4-year control run, others show differences between stochastic runs and the control run as labelled.

significant changes in convective cloud.

The changes in convective tendencies in the RP integrations relative to the control are on the whole hemispherically asymmetric and/or not consistent between the two RP runs, so they are unlikely to be significant.

4.3.4 LARGE-SCALE CIRCULATION

As well as clouds and their effects on radiation budgets, the large scale circulations in the aqua-planet simulations were studied by analysing the wind, temperature and moisture fields (not shown). However, none of the stochastic runs were found to have any statistically significant effect on the mean-states of these variables. On the whole, the variances of these fields were also un-changed in the stochastic runs, suggesting that the increase in high-frequency variability introduced by the stochastic schemes had no systematic effects on the slower, larger-scale dynamical modes which dominate the overall variance of temperature, humidity and winds.

Or if there were any systematic effects, they were too small to be detected compared to the very slow modes whose residuals dominate the mean-field differences between each stochastic run and the control run. The statistical tests at least ruled out any systematic changes in mean-state greater than of the order 0.5K in temperature, 0.5% in specific humidity, and 5% in zonal, meridional and vertical winds.

The one exception found was a significant increase in the variance of vertical velocity w in the tropics and sub-tropics in the Multiplicative Noise run. This is shown in figure 4.18, along with the mean and variability of the control run's vertical velocity.

The standard deviation of w in the MN run is around 5% greater than in the control run in the ascending and descending branches of the Hadley circulation in the lower and mid-troposphere, and in the stratosphere above. The mean w in the MN run has no systematic difference from the control, so it is just the variability of w which has been affected by the MN scheme. The Random Parameters runs (not shown) didn't give any significant change in the mean or variance of w . To investigate the variability in w further, figure 4.19 shows power spectra for w at three locations where its variance increased significantly.

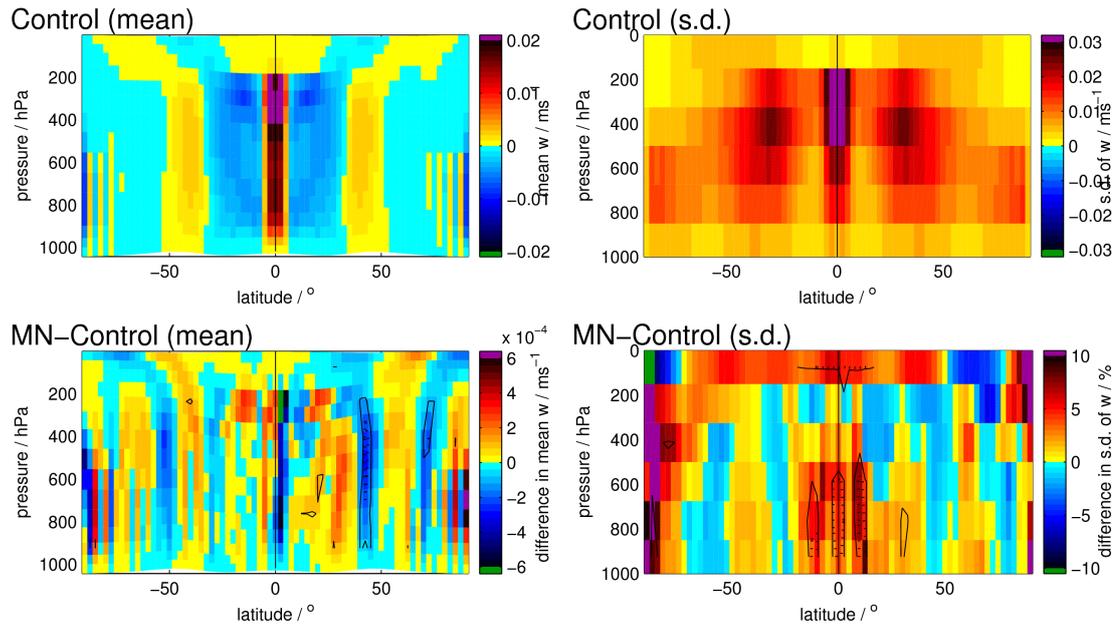


Figure 4.18: Vertical winds w ; (left) time-averaged zonal mean, and (right) total standard deviation, for (top) the control run, and (bottom) differences between the Multiplicative Noise run and the control run. Black contours denote statistical significance of the differences, at the (solid) 5% and (dotted) 1% levels.

In the lower troposphere over the equator, there is a peak in the spectrum at a period of 10 days, consistent with the Kelvin waves which dominate the rainfall variability shown in section 4.3.2. As noted in that section, the stochastic schemes do not significantly affect this mode, but the multiplicative noise scheme significantly boosts the variability at timescales faster than around 2 days. At all three locations this boost is sufficient to affect the overall variance of w , by simply adding the high-frequency multiplicative noise on top of the leading modes rather than actually altering them.

Of course, the model-field variances and power spectra studied have all been calculated from the output 6-hourly means of the data, so any changes which only affect variability on timescales faster than 6 hours have been neglected in this chapter. The aim was to look for changes in the slower, resolved modes rather than in the fast on-off convective noise known to be directly perturbed by the stochastic schemes. Nonetheless, the stochastic forcings applied by the MN and RP schemes have components of variability at timescales longer than 6 hours; the MN scheme applies white noise above this timescale and thus has a flat spectrum, whilst the RP scheme gives a heavily reddened time-series of parameter values. One might expect these to show up in the

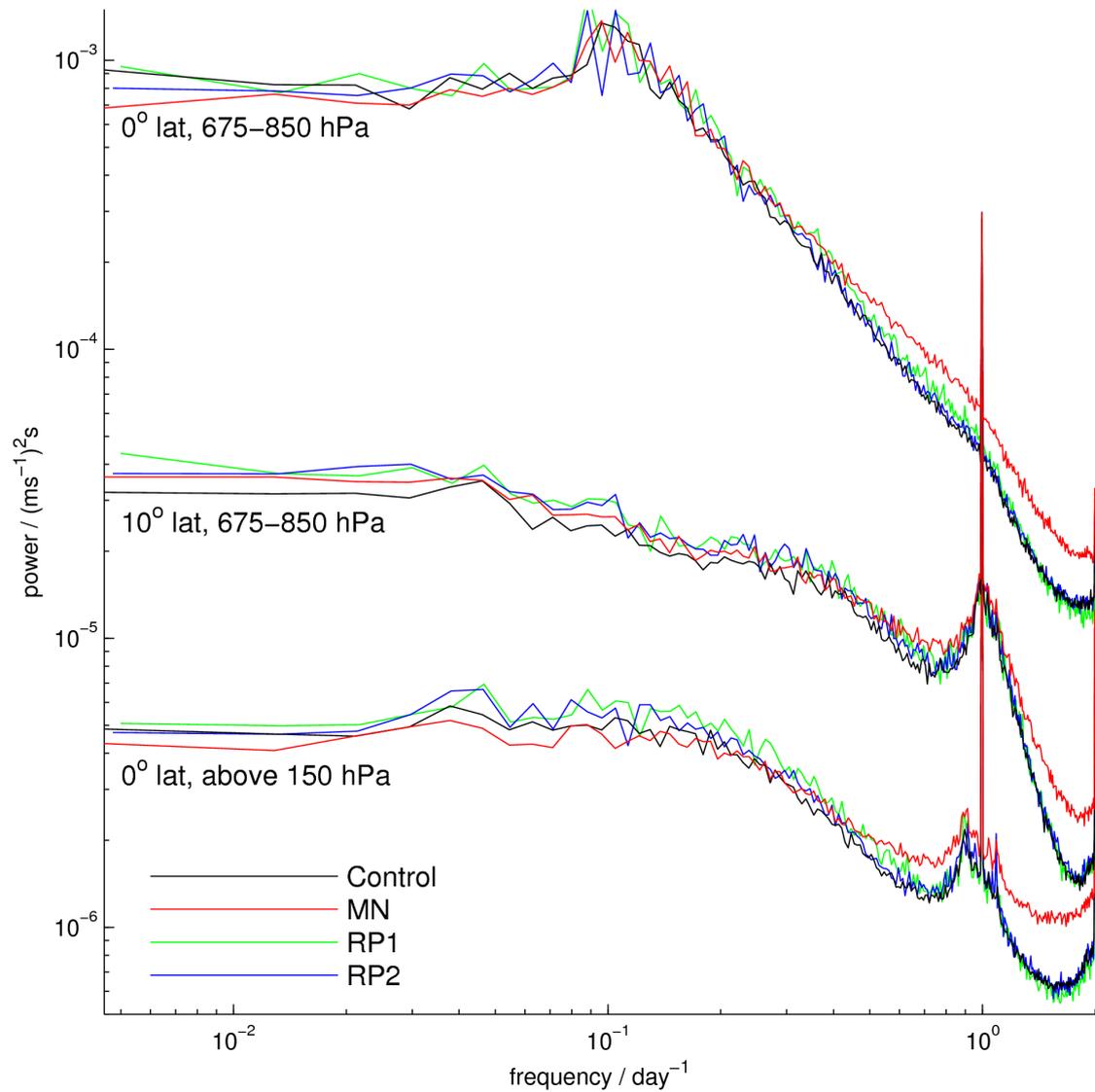


Figure 4.19: Mean power spectra for vertical velocity w at three latitudes / heights where a significant increase in variance was found in the MN run, for (black) the aqua-planet control simulation, (red) the Multiplicative Noise run, and (green, blue) the two Random Parameters runs. Spectral power has been normalised as in figure 4.10, but the axes scales are logarithmic.

model data, yet in most variables they don't.

It is interesting that vertical velocity gains high-frequency noise from the MN scheme although it is not directly perturbed, whereas the variables which are directly perturbed by the scheme (temperature, moisture and horizontal winds) do not show any significant change in variance. For temperature and moisture, this can be explained by a dynamical stabilisation response. The convective triggering events which dominate the high frequency variability act to warm and dry the model atmosphere. If the MN scheme acts to amplify (suppress) such an event, the dynamics respond with an increased (decreased) vertical velocity so-as to remove the temperature perturbation. This also creates vertical moisture advection in the right sense to oppose the moisture perturbation. In this way, the high-frequency variability added to the temperature and moisture fields is suppressed by the dynamics, producing vertical velocity variability instead. The dynamical stabilisation clearly acts on timescales faster than 6 hours, so that no increases in temperature and moisture variance in the stochastic runs are evident in the 6-hourly data.

It is tantalising to hypothesise that the increased high-frequency variability in the stratosphere above the tropics could be a result of vertically propagating gravity waves excited by the increased convective heating variability below. However, it was shown in chapter 3 in the Single-Column Model framework that the MN scheme produces considerable heating variability directly in the stratosphere as it perturbs the strong radiative tendencies there. A dynamical stabilisation response to these stochastic radiation perturbations is likely to be the primary cause of the increase in variance of w in the stratosphere.

4.4 DISCUSSION AND CONCLUSIONS

Idealised aqua-planet simulations of global atmospheric circulation were performed; two with different stochastic schemes designed to represent parameterisation uncertainty were compared against a deterministic control run, to investigate the sensitivity of the global atmospheric circulation to changes in the high frequency variability of parameterised processes.

The control run was found to simulate the basic features of the large-scale circulation, but with a more equatorially-confined Hadley circulation and convective showers over the polar regions,

both due to large departures of the prescribed SST pattern from reality in those regions.

The aqua-planet's tropical rainfall variability was found to be dominated by equatorially trapped Kelvin waves, which propagate Eastwards around the globe in around 30 days, and slower-moving Westward-propagating North-South oscillations in the latitude of the ITCZ. These modes were found to be insensitive to the stochastic parameterisations used in this study, as was the full spectrum of rainfall variability. This is in contrast to other studies such as Lin & Neelin (2000), who found that modelled tropical rainfall variability was highly sensitive to stochastic perturbation of the convection parameterisation.

However, the stochastic schemes were found to give a systematic increase in rainfall from the large-scale precipitation scheme over the tropics (although this was insignificant compared to the total rainfall there, which was mostly convective). For the RP scheme, this is in agreement with results from the Single-Column Model (SCM), presented in chapter 3. However, the MN scheme did not appear to give a significant increase in large-scale precipitation in the SCM (figure 3.15 in chapter 3). This could be because the SCM runs were not representative of the full tropical variability explored by the aqua-planet. There is also some ambiguity over the amount of large-scale precipitation in the MN runs, as the MN scheme does not form a closed moisture budget. Large-scale rainfall is produced from condensate, but condensate variables are not perturbed by the scheme, such that their tendencies become inconsistent with the specific humidity tendencies.

The increases in large-scale precipitation are expected given the systematic increases in layer cloud condensate in the stochastic aqua-planet runs. Similar increases in layer cloud also occurred for all the stochastic schemes in the SCM experiments (figure 3.20 in chapter 3). These were found to be a generic response to an increase in the range of humidities occurring in the model; positive humidity perturbations above saturation always increase condensation, but negative humidity perturbations can make no difference beyond removing all the condensate; a net increase in mean condensate inevitably results from this asymmetry.

Analysis of the 6-hourly temperature and moisture profiles in the aqua-planet found no significant differences between the stochastic runs and the control, in either the mean-state or the variance. It is at first confusing that significant increases in layer cloud still occurred, despite the fact that the key mechanism for the cloud-increases is thought to rely on an increase in humidity

variance. However, only the variance on timescales slower than 6 hours was studied in the aqua-planet, as 6-hourly mean data was used. The stochastic schemes presumably forced increased variability in temperature and humidity associated with the time-step-wise on-off variability of the convection scheme, as in the SCM. But this appears to have been suppressed on timescales of 6 hours or more by an atmospheric dynamical stabilisation response.

In the SCM, the RP and MN schemes gave similar increases in high-frequency temperature variability. For the MN scheme, the dynamical stabilisation response translated this in to vertical velocity variability in the aqua-planet, giving overall increases in the variance of w where convection and radiation give large parameterised tendencies. But the RP scheme did not give a significant increase in the variability of temperature or vertical velocity in the aqua-planet. A likely explanation for this is that the convective variability is too concentrated at high frequencies for significant stochastic effects to be seen in 6-hourly mean data, and the vertical velocity variance generated by the MN scheme is almost entirely forced by perturbing the more slowly-varying radiative tendencies.

The MN scheme also gave a significant reduction in convective cloud, especially in sub-tropical shallow convection. This was associated with a substantial decrease in the amount of moisture vented from the boundary layer in to the lower free troposphere by the convection parameterisation in the sub-tropics. Given that no significant change occurred in the humidity profiles, moisture budget or large-scale circulation, the sub-tropical convective flux decreases must be balanced by compensating increases in the moisture fluxes from the boundary layer parameterisation. A similar effect was seen for all the perturbed runs (including the MN scheme) in the SCM; increases in the boundary layer scheme moisture fluxes at 950hPa consistently occurred, and compensating reductions in convective moisture fluxes occurred at times. Whilst the SCM runs simulated the deep tropics, rather than sub-tropics, the boundary layer changes were especially pronounced during phases when deep convection was suppressed and shallow convection dominated, as in the sub-tropics in the aqua-planet.

For the RP run, the changes in the sub-tropical convective cloud and moisture fluxes seen in the SCM were not clearly apparent in the aqua-planet. The RP plots in figures 4.16 and 4.17 hint at some changes in low-level sub-tropical cloud and convective tendencies, but these are not consistent or statistically significant. Also, the RP scheme gives significant increases in convective

cloud in the upper troposphere over the tropics in the aqua-planet, but there was no clear sign of this in the SCM simulations. These results suggest the response of the convection and boundary layer parameterisations to the RP scheme is either highly modified by the dynamical response, or is highly dependent on details of the model or atmospheric state which differ between the SCM and aqua-planet. Perhaps there are two competing responses; the increased boundary layer moisture flux / decreased convective flux (seen in the SCM RP runs) could be competing with a separate response mechanism which gives a general increase in convection in the aqua-planet. Further work is needed to investigate this hypothesis, or alternative mechanisms.

The changes in cloud in both the MN and RP runs were found to significantly affect the radiation budgets in the model. The radiative changes were driven primarily by the convective cloud changes, with a secondary effect from the layer cloud increases, as convective cloud amounts were generally larger than layer cloud amounts in these runs. In the MN run, the reduction in low-level convective cloud gave a 3 Wm^{-2} increase in net downwelling Short-Wave radiation in the sub-tropics, resulting in the MN scheme giving a global mean net positive radiative forcing on the planet of nearly 1 Wm^{-2} relative to the control simulation. The RP scheme gave far less systematic change in the net radiation budget, with decreases in both net downwelling Short-Wave and net upwelling Long-Wave at the surface in the extra-tropics. This is consistent with the increases in cloud in the mid- and upper- troposphere in these regions, and yielded a net negative radiative forcing on the extra-tropics at top-of-atmosphere, a positive forcing on the surface, and some net loss of heat by the atmosphere. In a global-mean sense, these forcings were largely compensated for by changes in other regions which were not statistically significant; longer runs (or ensembles) are needed in order to investigate these robustly.

These results suggest that the simple stochastic schemes used here would likely have profound effects on coupled simulations with interactive ocean models. The surface radiative forcings associated with the cloud changes would alter the sea-surface temperatures and oceanic heat transports; this would in-turn alter the large-scale atmospheric circulation. Although the radiative forcing changes are fractionally small, coupled feedbacks in the ocean-atmosphere system could potentially yield substantial changes in the equilibrium climate mean-state.

However, in the absence of any coupled feedbacks from the ocean, the stochastic schemes had no detectable effect on the large-scale circulation or thermodynamic profiles in the aqua-

planet, despite significantly perturbing the cloud amounts and radiative fluxes. Analysis of the control run's radiation budgets suggested that variations in cloud primarily affect the amount of Short-Wave reflected or transmitted by the atmosphere, but have less effect on the net absorption / emission by the atmosphere. This is consistent with the results for the stochastic schemes; the changes in cloud had a significant effect on the top-of-atmosphere and surface radiation budgets, but not on the atmosphere's radiation budget. This, and the dynamical stabilisation response discussed earlier, likely explain the insensitivity of the atmosphere's mean-state to the stochastic parameterisations.

Further discussion and overall conclusions are given in the next and final chapter.

CHAPTER 5:

CONCLUSIONS

5.1 SUMMARY

The atmosphere's climatic behaviour is thought to exhibit key sensitivities to the nature of its high-frequency variability, but these sensitivities are not fully understood. This issue is of practical importance in weather forecasting and climate modelling, as parameterisations of sub-grid processes often exhibit deficient or otherwise unrealistic high-frequency variability, and the modelling community has paid increasing attention to developing stochastic parameterisations to improve the realism of high-frequency variability in models of the atmosphere.

Whilst a broad range of different stochastic methods have been developed, they share theoretical underpinnings in attempting to represent the uncertainty in / variability of the unresolved components of atmospheric variables. Deterministic formulations cannot wholly represent this as the sub-grid details of the atmosphere's state are by definition undefined in models. The continuous variability spectra observed in the atmosphere show that for any plausible model resolution, there will always be a component of variability which affects the resolved scale, but can neither be adequately resolved nor described entirely as a function of resolved variables using a statistical equilibrium assumption.

As well as leading to deficient variability at the smaller resolvable scales in models, this issue also leads to inherent uncertainties in tendencies produced by sub-grid parameterisations. In ensemble forecasting systems, it is important to account for all sources of uncertainty in the forecast, otherwise they give overconfidently small ensemble spreads. Various stochastic schemes have been developed to generically sample uncertainties in the formulation of parameterisations within ensemble forecasts, but it has long been recognised that some component of parameterisation uncertainty derives from the inherent sub-grid uncertainty in their tendencies.

In order to design appropriate stochastic parameterisations, both for improving the fidelity of model climates and for representing model uncertainties in ensemble forecast systems, a greater understanding of the atmosphere's sensitivity to the variability of sub-grid processes would be greatly beneficial. This thesis aims to address three key questions in particular, which were posed in section 1.7 of chapter 1; in summary:

1. Is the atmosphere's response to high-frequency variability highly dependent on the internal properties of that variability, the atmospheric state, or model-specific details? Or does it exhibit generic responses to any increase in "noise"? This is an important question regarding how sophisticated and realistic stochastic parameterisations need to be in order to yield realistic responses on the resolved scale.
2. How important is the theoretically underpinned notion of sub-grid uncertainty relative to generic model uncertainties, such as those associated with different choices of parameterisation formulation? If it is a major component of model uncertainty, ensemble forecasts should incorporate stochastic schemes consistent with known sub-grid uncertainties. If not, schemes which fully sample structural uncertainties in parameterisation formulation are more important.
3. What are the key mechanisms by which high-frequency variability can influence the global atmosphere's climatic behaviour? This question is of general interest in understanding atmospheric scale-interactions, and is relevant to guiding well-targeted efforts to develop stochastic (and other) parameterisations aimed at improving the realism of climate simulations.

To address these questions, the sensitivity of an atmospheric GCM (the UK Met Office Unified Model) to various alterations to its high-frequency variability was investigated. Sensitivity to model uncertainties was also explored by comparing multiple different deterministic convection parameterisations in the host model; various studies have shown that the parameterisation of moist convection is the leading source of high-frequency variability and modelling uncertainties in atmospheric GCMs, and the results in chapters 2 and 3 of this thesis are in agreement with this. The high-frequency variabilities of some of these deterministic configurations were then altered

using several stochastic or other methods drawn from the literature, to investigate the sensitivity of the atmosphere to the sources of variability / uncertainty they represent.

First, a broad comparison of stochastic and deterministic model configurations was made in a UM Single-Column Model simulation of transitions between active and suppressed phases of convection over the Tropical West Pacific. As well as providing useful results regarding the response of parameterised processes to different variabilities, these experiments guided the choice of a smaller set of configurations likely to yield interesting results in a 3D framework consistent with climate simulations. A comparison of this more limited set of configurations was then performed using the UM aqua-planet, to investigate the nature of the full, dynamically-coupled atmospheric response, and to what extent the dynamics modify the parameterised response.

Discussion and conclusions drawing from the results of these experiments is given in the following three sections, addressing each of the three questions posed above.

5.2 IS THE ATMOSPHERE’S RESPONSE TO HIGH-FREQUENCY VARIABILITY HIGHLY DEPENDENT ON IT’S INTERNAL PROPERTIES?

All of the different stochastic parameterisation methods investigated in this study were found to give qualitatively similar high-frequency responses, constituting a simple “scaling-up” of the host model’s existing convective variability. Whilst different deterministic convection parameterisations were found to have very different variabilities, a generic scaling-up of those variabilities was found for stochastic schemes applied to two different convection parameterisations. The resulting increase in the range of thermodynamic states available to the system induced some further generic responses from parameterised processes. The mechanisms for the changes in convective variability, parameterisation responses, and the implications, are discussed further in section 5.4).

In both the SCM and aqua-planet simulation frameworks, all of the stochastic schemes gave a systematic increase in layer cloud condensate due to the increased spread of humidities explored relative to the saturation threshold.

The SCM experiments also indicated a generic response from the boundary layer parameterisation, which gave consistently increased fluxes in the upper boundary layer in response to increased variability. However, there were also (possibly related) responses from the convection parameterisation. One stochastic scheme (a stochastic multiplicative perturbation method) gave similar moisture flux changes in the sub-tropics in the aqua-planet to those in the SCM, increasing the fluxes from the boundary layer scheme and reducing the activity of the convection scheme in the lower troposphere. But this resulted in a large decrease in low-level convective cloud in the aqua-planet not evident in the SCM. Meanwhile, another stochastic scheme (stochastic perturbation of model parameters) gave very different moisture flux and convective cloud changes in the SCM and aqua-planet, giving an increase in convective cloud in the latter. Clearly the convective cloud response to high-frequency variability maybe highly sensitive to particulars of the model, the atmospheric state, dynamical feedbacks and / or the type of stochastic parameterisation employed.

In the SCM, the play-off between the boundary layer scheme and convection responses gave subtle, highly state-dependent changes in the ensemble mean temperature and humidity profiles, with little in the way of systematic changes in a time-averaged sense. Correspondingly, the stochastic schemes gave no detectible mean-state response in temperature and humidity in the aqua-planet. The convective responses dominated the stochastic effects on the aqua-planet's radiation budget, due to associated changes in convective cloud, but differed enormously between the two different stochastic methods.

So, whilst the variability of the parameterised convection responded quite generically to various different stochastic methods, and the mean layer cloud condensate responded generically to the increased variability, the convective cloud / radiative response to the high-frequency variability was highly non-generic and there was little or no systematic mean response from the atmosphere in terms of thermodynamic profiles or the large-scale circulation.

These results suggest that in the design of stochastic parameterisations, much attention should be given to their interactions with the convection parameterisation. In particular, if the convection is tuned to give appropriate profiles of convective cloud, it will likely need to be retuned if stochastic parameterisations are subsequently introduced.

As is discussed in more detail in section 5.4 below, the responses of the convective cloud and boundary layer fluxes are likely to be symptomatic of non-linearities in the way the model treats various forms of buoyancy-driven fluxes, through arbitrary partition between separate boundary layer, shallow convection and deep convection parameterisation modules. In this study, the introduction of stochastic parameterisations which merely aim to symmetrically sample model uncertainties may have served to highlight unphysical non-linearities in this modelling strategy. Further work is needed to investigate the precise mechanisms involved (e.g. to what extent the convection and boundary scheme responses to stochastic forcing are related).

However, these results hint at the prevalence of purely numerical sources of uncertainty / variability where the different sub-grid parameterisation modules compete to handle some of the same physics, with the modules differing in their treatment of processes for which they overlap (e.g. there is no exact physical definition to partition which fluxes should be handled by the boundary layer and shallow convection schemes and in practice they overlap in the host model, but produce very different amounts of cloud depending on which one dominates). This points towards a need for a more unified understanding of sub-grid buoyancy-driven motions in the atmosphere, and a correspondingly unified framework to represent them in models. A proposal for such a framework is given in section 5.5.

5.3 HOW IMPORTANT ARE INHERENT SUB-GRID UNCERTAINTIES RELATIVE TO GENERIC PARAMETERISATION UNCERTAINTIES?

The SCM framework was used to directly compare ensemble spreads associated with estimates of sub-grid uncertainty and generic parameterisation uncertainty.

Sub-grid uncertainty was quantified by the Plant & Craig (2008) stochastic scheme, which simulates the fluctuations in grid-mean convective tendencies associated with likely sub-grid variations in the population of convective plumes within a grid-area. Since convection was shown to be the primary source of variability / uncertainty in the SCM, and is the most energetic unresolved atmospheric process, focussing purely on convection (i.e. neglecting the sub-grid fluctuations in

other parameterised processes) shouldn't be too detrimental to this comparison. The smaller the grid-size, the larger the fluctuations / sub-grid uncertainties in the convective tendencies produced by the stochastic convection scheme, as a smaller grid area will encompass a more limited sampling of the possible spectrum of different convective plumes, and so tend to venture further from statistical equilibrium.

Generic parameterisation uncertainty was quantified by the spread of two stochastic schemes from the literature aimed at accounting for just that, as well as an analogous constant perturbed parameter ensemble, and a combined deterministic ensemble directly sampling the structural uncertainties between three different convection parameterisation formulations. All of these measures of parameterisation uncertainty were found to be broadly in good agreement, although it would seem surprising if the uncertainties could be fully sampled by just three closely related deterministic model configurations.

The comparison suggested that sub-grid uncertainty as measured by the Plant & Craig (2008) scheme will dominate the modelling uncertainties sampled by the aforementioned methods at horizontal scales of 50 km or less, but become an increasingly unimportant contribution to model uncertainty at scales larger than 100 km. This would suggest that it is important to use stochastic schemes consistent with physical sources of sub-grid variability / uncertainty in state-of-the-art weather forecast models, which typically have grid-sizes smaller than 50 km, but that methods aimed at sampling generic parameterisation uncertainties should be a higher priority for climate simulations, which usually have grid-sizes of around 100 km or more.

This conclusion comes with the caveat that both sub-grid uncertainty and parameterisation uncertainty maybe underestimated in this study. In the case of parameterisation uncertainty, a larger number of current deterministic model formulations need to be compared to get a more representative measure of the structural uncertainties. In the case of sub-grid uncertainty, it is not clear that the effects of mesoscale organisation of convection are accounted for in the Plant & Craig (2008) stochastic scheme. At smaller grid-sizes, such systems become increasingly resolvable by the model dynamics, making this less of an issue. But at climate GCM grid-sizes, organised convective systems represent the most energetic mode of sub-grid variability.

5.4 WHAT ARE THE KEY MECHANISMS INVOLVED IN THE GLOBAL ATMOSPHERE'S SENSITIVITY?

In the SCM experiments, the primary response to stochastically forced increases in the variability of parameterised processes appeared to constitute a more-or-less linear scaling-up of the fast, “on-off” variability of the convection scheme. It was shown that this convective variability is the leading mode at short timescales, and also the main source (among sub-grid or non-dynamical processes) of non-linear sensitivity to perturbations. Whilst the changes in ensemble variability seen when stochastic schemes were introduced acted through the strong non-linearities in the “on-off” convective triggering, their dominant effect was more akin to a linear stochastic forcing response (the *Type 0* response illustrated in a toy-model in section 1.6 of chapter 1). The various stochastic parameterisations were found to directly force an increase in the amplitude of the convective variability, but did not yield major qualitative changes in its behaviour, or in the slower modes of variability. Only subtle, state-dependent changes in ensemble mean thermodynamic profiles occurred when stochastic parameterisations were introduced, and these were far smaller than the mean-state differences between different deterministic convection scheme formulations.

Results in the aqua-planet framework showed that the dynamical response to simple stochastic schemes designed to generically sample modelling uncertainties was similarly linear, with no significant changes in the atmosphere's mean-state or internal modes of variability. Further, the variability in the temperature and moisture profiles generated by convective variability appears to be suppressed by a dynamical stabilisation response, at least on timescales greater than 6 hours.

Whilst parameterised convection is known to have unrealistic variability (the “on-off noise” driving the stochastic response in this study is considered by many to be numerical artefact), its' strongly non-linear sensitivities are presumably in some ways representative of those in the real world. For example, convection-resolving weather forecasts diverge very rapidly relative to the observed locations and intensities of thunderstorms. Therefore, the increased variability of parameterised convection in response to stochastic perturbations should be in some way useful for quantifying the instantaneous uncertainty in convective tendencies.

These results support the use of generic stochastic forcings to represent uncertainties in pa-

parameterised processes in ensemble forecasts; if appropriately tuned, the forcings can yield appropriately increased ranges of parameterised tendencies in an ensemble, whilst having minimal effect on the finely tuned mean-state and modes of variability of the host model.

However, comparison of different deterministic convection parameterisations in the SCM showed that structural uncertainties in the convection scheme can yield large uncertainties in the shape of the profile of heating variability. If this is important in the fidelity of atmospheric simulations (and studies investigating the importance of gravity wave excitation by convective heating suggest it might be), then the profile of heating variability needs to be tuned in climate simulations, or its' internal uncertainties well-sampled in ensembles. However, the stochastic parameterisations investigated in this study on the whole didn't seem to change the shape of the heating profile; they simply increased its amplitude, so they may not be effective at tuning or sampling its shape. To more fully represent modelling uncertainties, it may be necessary to develop stochastic parameterisations which sample uncertainties in the convective plume model formulation, rather than being restricted to perturbing certain control parameters or the overall amplitude of the convective response.

It was also notable in this study that the stochastic parameterisations which attempted to directly force prescribed timescales on the "on-off" convective variability (the Random Parameters and Multiplicative Noise schemes) failed in this endeavour; they simply acted to boost the existing convective noise without affecting its internal timescales. It therefore seems likely that efforts to improve the realism of the timescales of variability of penetrative mass-flux convection schemes, through direct stochastic forcing of their outputs or closure parameters, will have limited success.

This study found that, consistent with studies using prognostic closures on the convection parameterisation, altering the convective closure to time-smooth its response is far more effective at imposing longer timescales on the convective variability. However, the time-smoothing scheme studied in the SCM in this study also caused a large drift in the atmospheric mean-state. This was not a response specifically to the increased timescales of convection; time-smoothing the response to convective instability partially delays the response, resulting in a more unstable atmospheric state and more vigorous convection. In the SCM, further changes then resulted from radiative feedbacks to the associated drying of the moisture profile.

Whilst the mean-states of thermodynamic profiles and winds were not found to have major sensitivity to an increase in convective variability in this study, the cloud variables were found to be quite sensitive to the introduction of stochastic parameterisations, or other schemes which increase the range of thermodynamic states explored by the parameterised processes.

Firstly, as discussed earlier, in both the SCM and the aqua-planet frameworks, all of the stochastic schemes investigated yielded a general increase in layer cloud condensate, independent of any changes in mean humidity. The mechanism for this is straightforward; a symmetric increase in the range of humidities explored by the system usually gives an asymmetric (positive) response in mean cloud condensate, because greater positive humidity excursions yield equivalently greater condensation, whereas greater negative excursions can make no difference beyond evaporating all of the condensate. In the aqua-planet, and in some of the SCM runs, the increased mean layer cloud condensate also resulted in an increase in the amount of precipitation produced by the layer-cloud microphysics scheme.

In the design of model cloud schemes, much attention has been given to accounting for the effects of sub-grid inhomogeneity, which yield increases in cloud relative to a homogeneous moisture field, essentially via the same mechanism as that described above. Sub-grid fluctuations can cause cloud to form in some parts of a model grid area even when the grid-mean humidity is well below saturation. Failure to account for this sub-grid variability results in unrealistically low cloud amounts and inflated humidities. The results of this study show that the same will be true to some extent in models which have unrealistically low humidity variability at and above the grid-scale. In practice, the cloud schemes in models are often tuned so as to yield realistic cloud amounts and humidities overall, but the results of this study suggest that this will yield inappropriately large sub-grid cloud contributions if the grid-scale humidity variability in the host model is too small. Where stochastic parameterisations are implemented in forecast or climate models, the cloud scheme may need to be retuned, just as it would under a change in model resolution.

The increase in the range of thermodynamic states associated with increased variability of parameterised processes was also found to give a systematic response in the activity of the boundary layer and convection schemes in the lower troposphere. In the SCM, all the stochastic schemes gave systematic increases in the moisture fluxes from the boundary layer parameterisation, driven by increases in its activity near the boundary layer top. This effect was largely off-set in a time-

averaged sense by a highly state-dependent response from the convection scheme, which gave decreased moisture fluxes in the boundary layer at times.

Whilst the sum of these two effects gave very little net change in the time-averaged moisture budget of the lower troposphere, the changes in the activity of the convection scheme yielded large changes in convective cloud amounts. In the UM, the radiative effects of convective cloud are parameterised independently of the grid-scale cloud variables, as cloud within convective updrafts and spreading anvil formations largely evaporates or is rained-out once the convective moisture tendencies are fed-back to the grid-scale variables seen by the radiation scheme. The shallow convection scheme tends to parameterise substantially larger cloud amounts than are generated by the boundary layer / large-scale cloud schemes (this was evident in the aqua-planet simulations). Therefore, changes in the relative activity of these schemes in the lower troposphere gave substantial changes in cloud.

It is unclear whether the boundary layer and convection scheme responses have any physical mechanism underlying them. It might be purely a model issue related to the somewhat arbitrary partitioning of buoyancy-driven fluxes in to “boundary layer” overturning, “shallow convection” and “deep convection”. Further work is needed to investigate the mechanisms for these flux changes.

Whilst the mechanism for the convective cloud response to stochastic forcings is unclear, these results strongly suggest that the changes in cloud will themselves form a mechanism for significant sensitivities of the mean-states of coupled climate simulations to high-frequency variability, as they yield substantial radiative forcings on the surface.

5.5 FUTURE WORK

Whilst the results presented in this thesis have allowed some conclusions to be drawn, there are also some areas where further analysis or improvements to the methodology are needed in order to make more definitive statements regarding the questions posed. Suggested avenues for further work within the present methodological framework (comparisons of stochastic and deterministic model configurations in SCM and aqua-planet simulations) are given in subsection 5.5.1. Whilst

it is unlikely that the author will have the opportunity to carry out these improvements to the methodology, they are described here for the benefit of anyone conducting related experiments in the future. There are also some areas which could be better addressed using alternative methods. These are proposed in subsection 5.5.2. Finally, the results of this study have led the author to suggest some potential avenues for improving GCM parameterisation schemes, described in subsection 5.5.3.

5.5.1 EXTENSION OF THE PRESENT METHODOLOGY

Firstly, the Single-Column Model experiments revealed some parameterised responses to changes in high-frequency variability for which the underlying mechanisms were not apparent. Further work to identify these mechanisms would be useful.

In particular, the increase in moisture fluxes at 950 hPa from the boundary layer parameterisation was identified as a consistent response to increasing the range of available states, and as a key driver of changes in the lower troposphere moisture profiles and surface fluxes, but its cause is unclear. The SCM ensemble simulations could be repeated with further diagnostics output from the boundary layer parameterisation to investigate this further. The mechanism behind the highly state-dependent changes in mean convective activity were also unclear, and again further diagnostics from the convection scheme could be output to investigate this. The relationship between these two responses is also worth investigating; it might be that the boundary layer scheme changes are driving the convective changes. This could be investigated by performing SCM ensemble runs in which the convection scheme is disabled and replaced by prescribed tendencies from the default UM run, and others in which the same is done to the boundary layer scheme instead, and investigating whether the boundary layer and convection responses to stochastic schemes still occur in each of these configurations. Repeating the deterministic and stochastic ensemble runs with alternative boundary layer scheme formulations would also be useful, to identify whether the responses are specific to the UM boundary layer scheme.

Another unexplained stochastic response in the SCM experiments was the increased rainfall and more stable thermodynamic profile when stochastic convective fluctuations were introduced in the Plant & Craig (2008) scheme. This could be explored by repeating the Plant & Craig

scheme runs with more diagnostics relating to the convection scheme. For example, the effects of reducing the prescribed grid-size (which controls the magnitude of fluctuations in the scheme) on the population of convective plumes simulated, the spectral CAPE closure and the convective triggering might give some clues.

It would also be good to clear up the moisture budget ambiguities in the implementation of the stochastic multiplicative perturbation scheme (based on Buizza *et al.* 1999) used in this study; simply perturbing the condensate, precipitation and surface moisture fluxes consistently with the specific humidity during the run would address this. More reliable difference plots for precipitation and surface fluxes relative to the Default UM could then be studied.

Whilst the crash of the aqua-planet simulation with the Random Parameters (RP) stochastic scheme based on Bowler *et al.* (2008) was frustrating, it was also potentially an interesting result in itself, if it was actually forced by the stochastic scheme rather than being a chance occurrence in those runs. The crash was caused by the appearance of an unrealistically intense polar cyclone which fed off the unphysically warm prescribed sea-surface temperatures over the poles, but more data is needed to identify why exactly this cyclone appeared in the RP runs and not the other runs. No significant changes in the mean-state of the RP simulation were apparent from the available data, but very little data was output from upper levels. The available data suggested a cooling over the pole at upper levels just as the cyclone became established. Runs could be repeated with more data from the stratosphere, to investigate the hypothesis that the cyclone was able to intensify due to upper-level cooling, induced by a response of stratospheric dynamics to the RP scheme.

A single, long integration was performed for each aqua-planet configuration, as other studies suggested that slow modes of tropical rainfall variability might be sensitive to stochastic forcing. Whilst some interesting low-frequency modes were present, they displayed no detectable sensitivity to the stochastic schemes investigated. Instead, responses from parameterised processes were found. These did not require such long runs to spin-up, and the long time-series of 6-hourly mean data, with few physical diagnostics, were not ideal for investigating them. And the differences between the parameterisations and model-states between the SCM and aqua-planet frameworks prevented more direct comparisons of the parameterisation responses in the two. A shorter length of run in an ensemble framework would better suit such comparisons (in particular, using ensembles of runs would allow more statistically robust comparisons of different model configurations).

But it would nonetheless be beneficial to sample the full range of model states explored by the aqua-planet's long timescale modes.

An appropriate method might be to perform a long control integration, say 20 years in length, and then use 6-monthly model dumps from that run to initialise 40-member ensembles for the stochastic schemes, each 1 month in length. The 40 1-month-long segments of the control run following each dump could be analysed as an ensemble for comparison. The 6-month spacing of the dumps would make the ensemble members reasonably independent of one-another (adjacent segments would be more correlated, undermining the use of simple statistical methods). In each ensemble member, detailed diagnostics could be output at a few representative grid-points (tropical, sub-tropical, mid-latitude and polar locations). These could then be analysed using the same methods as in the Single-Column Model. Further, the dynamical tendencies from the aqua-planet at those points could be output and used to force additional Single-Column Model runs, which could be compared directly to the aqua-planet ensemble data to better quantify the dynamical and parameterised responses to stochastic forcings.

Much attention has been given by operational forecasting centres to the development of stochastic backscatter schemes, which directly perturb the model winds and aim to excite high-frequency dynamical modes, consistent with upscale transfer of energy from unresolved modes. Studies have found that such schemes can influence model climatologies and improve the probabilistic skill of ensemble forecasts. Backscatter schemes could not meaningfully be investigated in the Single-Column Model, which prescribes the dynamics, but it would be interesting to investigate the mechanisms for the noise-induced effects of a backscatter scheme in the aqua-planet framework.

Results from the Single-Column Model using the stochastic convection scheme of Plant & Craig (2008) suggested that the sub-grid convective fluctuations it simulates could play a key role in high frequency variability at resolutions consistent with weather forecast models. The aqua-planet used in this study has a much larger grid-size, at which that scheme would likely have little effect. It would be interesting to investigate the role of sub-grid convective fluctuations in the climate of a higher resolution aqua-planet, with a horizontal resolution of say N240, consistent with a weather forecast model.

5.5.2 OTHER SUGGESTED EXPERIMENTS

In the Single-Column Model (SCM), different deterministic convective formulations were found to give very differently-shaped profiles of variability in convective heating, implying large parameterisation uncertainties, which were not adequately sampled by the stochastic parameterisations tested. However, it seems likely that constraints on the profile of convective heating variability could be derived from Cloud-Resolving Models (CRMs). The observation-derived forcings and initial condition perturbations used to drive the SCM ensembles (see chapter 2) could be applied to an ensemble of CRM integrations covering a domain consistent with a climate-model grid-box, say 100km square. The domain-mean heating rates could then be computed for each ensemble member to obtain an ensemble of “grid-mean” sub-grid tendencies, from which the profiles of ensemble spread and decorrelation timescale could be computed equivalently to the SCM data. By comparing the SCM ensemble profiles of heating variability for various different parameterisation formulations to the profiles from the CRM, it should be possible to determine which convection parameterisations give the most realistic profiles, and which are clearly unrealistic. Whilst comparison of parameterisations to coarse-grained data from CRMs is already a widely used methodology, most studies focus on the ensemble means of thermodynamic profiles or sub-grid tendencies. The author believes such methods would be well-suited to investigating high-frequency variability.

Also, the SCM comparison of sub-grid uncertainties (as quantified by the Plant & Craig 2008 stochastic convection scheme) to model uncertainties could be validated by using the ensemble variability of the domain-mean CRM data as another, perhaps more robust, measure of sub-grid uncertainties. In that comparison, model uncertainties also need to be better quantified. This could be achieved by expanding the combined deterministic SCM ensemble described in chapter 3 to include a larger, more representative range of different parameterisation formulations. The spread of such an ensemble could be compared to the spread of the ensemble of CRM domain-mean data, to make more definitive statements about the relative magnitudes of sub-grid uncertainties and generic parameterisation uncertainties.

5.5.3 SUGGESTED MODEL DEVELOPMENTS

It was noted in section 5.2 that unphysical non-linear sensitivities may arise in models from the arbitrary partitioning of buoyancy-driven fluxes between the boundary layer, shallow convection and deep convection parameterisations, and that this issue could be addressed through a more unified treatment of buoyant plumes. One way to achieve this might be through an extension of the stochastic spectral convection parameterisation paradigm on which the Plant & Craig (2008) scheme is based. That scheme already replaces the separate shallow and deep convection schemes with a spectrum of plume sizes which bridges the shallow and deep modes consistently. Why not also replace the non-local buoyancy-driven component of the boundary layer parameterisation with a unified scheme simulating the full spectrum of sub-grid buoyancy-driven mass-fluxes, from boundary-layer thermals to thunderstorms? The Plant & Craig (2008) scheme aborts plumes which terminate without producing cloud; this check could be turned off to include the smaller plumes internal to the boundary layer.

Whilst the Plant & Craig (2008) scheme automatically computes appropriate entrainment rates for the different plume sizes, it applies a single spectral dilute CAPE-based closure to the overall mass-flux, which may not be as appropriate for the smallest plume sizes; shallow convection schemes are often closed on a measure of surface buoyancy flux rather than CAPE. To bridge these two types of closure consistently, a buoyancy-flux term could be added to the plume work function for all the plumes used in the spectral closure to calculate total mass-flux; the plume work function would still dominate where deep plumes release significant CAPE, but where this does not occur the surface buoyancy flux term would dominate, yielding shallow convective behaviour. Crucially, the intermediate spectrum of “cumulus congestus” type plumes would then be handled consistently.

The distribution of plume mass-fluxes used in the Plant & Craig (2008) scheme is based on a radiative-convective equilibrium in which a full spectrum of plume sizes has become established. This equilibrium assumption (not to be confused with the statistical equilibrium assumption!) could be relaxed by modifying the distribution of plume mass-fluxes in the scheme to account for non-equilibrium behaviour. A simple modification would be to impose an evolving maximum bound on the allowed size of individual plumes, consistent with a horizontal scale equal to the

depth of the boundary layer, or the maximum height reached by other recent plumes if this is greater. This would incorporate the observed tendency of convective plumes not to have a much greater horizontal scale than the depth of over-turning (with the exception of organised convective systems, but these only occur under specific conditions). Crucially, this would only allow the parameterised convection to access the small entrainment rates consistent with well-established deep cumulonimbus once it has managed to penetrate to such depths via smaller congestus-type plumes with higher entrainment rates, as observed. This should improve the realism of the diurnal cycle of parameterised convection, which is known to trigger deep plumes too early in the cycle.

A further point of potential unphysical non-linearity highlighted in this work is the very different cloud amounts seen by the radiation scheme, depending on whether the cloud is actually handled by the UM cloud scheme or represented diagnostically as an output from the convection scheme. Scinocca & McFarlane (2004) showed that alterations to the convective closure can drastically alter the partitioning of cloud and rainfall-production between the equivalent two schemes in a different model, and in this study stochastic perturbations applied to an aqua-planet yielded large changes in sub-tropical boundary-layer cloud due to a change in the partitioning of fluxes between the convection and boundary-layer schemes. This suggests a more unified framework for representing clouds in models would be beneficial.

At present, the condensate within convective plumes is accounted for diagnostically, and then all either rained-out or evaporated on detrainment. If the detrained moisture gives sufficient moistening of the grid-mean state, it is then handled by the large-scale cloud scheme, potentially resulting in “double counting” of its’ radiative effects. An alternative might be to handle only the rained-out convective cloud condensate diagnostically, whilst all the remaining moisture from convective plumes is handled by a unified cloud scheme, at the level where it is detrained. The scheme would need to account for the increased sub-grid inhomogeneity in the moisture field associated with buoyancy-driven plumes, or it would likely underestimate the cloud amounts. This could be done by tuning the cloud scheme interactively to produce appropriate cloud amounts where buoyancy-driven moisture fluxes occur, but relaxing it back towards lower sub-grid inhomogeneities where they do not. Crucially, such a scheme could give a more consistent treatment of boundary-layer top cloud (in which very little convective rain-out occurs) as all the cloud would be handled within the same framework and experience the same microphysics.

Appendices

APPENDIX A:

**DETAIL OF THE PARAMETERISATION
SCHEMES USED IN THE MET OFFICE
UNIFIED MODEL**

This appendix describes in some detail the parameterisations of sub-grid and non-dynamical processes applied in the UM, with appropriate references given. This material is included here, particularly for reference regarding the Random Parameters stochastic model uncertainty scheme (Arribas 2004, Bowler *et al.* 2008) tested in the SCM and the Aqua-Planet in this study, which perturbs parameters in the UM parameterisations. The UM includes parameterisations for boundary-layer turbulent mixing and surface fluxes, sub-grid convection, cloud (accounting for sub-grid moisture inhomogeneity), cloud microphysics / precipitation, and radiative transfer. Each of these schemes is described in the sections below. The UM also includes a parameterisation for orographically generated gravity waves, but this is not applied in the experiments conducted in this thesis (neither the SCM or the Aqua-Planet contain any orography), so it is not described here.

A.1 BOUNDARY LAYER

A Richardson number based scheme is applied for local mixing (Smith 1990). The vertical turbulent flux of each model variable is assumed to be proportional to its vertical gradient, the magnitude of the wind-shear, an empirical stability function dependent purely on the Richardson number, and an estimate of the neutral mixing length at each model level.

There is also a non-local component which applies for unstable boundary layer conditions, to parameterise the effects of buoyant motions driven by heating at the surface or cooling at

the boundary layer cloud-top. This is described by Lock *et al.* 2000. It applies turbulent mixing profiles covering multiple model-levels, consistent with the expected influence of surface or cloud-top buoyancy sources.

A.2 CONVECTION

The convection parameterisation used is based on the penetrative mass-flux bulk plume model of Gregory & Rowntree (1990), but with a different and more complex closure and interface to the host model (briefly described in Martin *et al.* 2006).

The scheme first employs a test parcel ascent calculation and an analysis of the stability of the boundary layer temperature and moisture profiles to estimate whether convection capable of penetrating the boundary layer top is possible. If so, it then assesses whether the convection will likely be “deep” or “shallow” (based on the height reached by the test parcel and a check for large-scale subsidence at the boundary layer top) and calls either a deep or shallow version of the Gregory & Rowntree (1990) plume model accordingly. Each version then applies a trigger function, only actually producing any convection if the profile contains a parcel which, if given a small buoyancy excess (which is calculated as a function of the environment buoyancy gradient), remains buoyant if lifted by one model level. If this condition is met, a bulk entraining-detraining plume calculation is performed to estimate convective cloud condensate, precipitation and the effect of plumes on their environment.

The deep convection scheme represents the cumulonimbus clouds which generate most of the convective heating and rainfall. It employs the CAPE closure method of Fritsch & Chappell (1980). The convective mass-flux is scaled such that the convection will remove the dilute CAPE calculated over the plume ascent at a specified rate, which is set by a timescale τ_{CAPE} .

The shallow convection scheme represents the effects of the small but widespread “fair weather” or “trade wind” cumulus clouds, which produce little rain or heating but play a vital role in moistening the lower free troposphere. It uses a closure based on a boundary layer turbulent kinetic energy budget, relating cloud-base mass flux to a turbulence vertical velocity scale which is a function of the surface buoyancy flux (Grant 2001). It also employs different entrainment and

detrainment calculations (those of Grant & Brown 1999).

Additionally, there is a third “mid-level” version of the scheme, which is designed to represent convection that originates outside the boundary layer. It therefore calls at all points independently of the boundary layer tests applied for the deep and shallow schemes. It is subject to a more stringent trigger function, with a smaller parcel buoyancy excess constrained to be less than 0.2K. It uses the same CAPE closure method as the deep scheme.

All three schemes incorporate the convective downdraft scheme of Gregory & Allen (1991). The convective momentum transport scheme of Gregory *et al.* (1997) is also included, although this will have little effect in the SCM as winds are strongly relaxed towards observed profiles.

A.3 LARGE-SCALE CLOUD

The cloud scheme of Smith (1990) is used. Cloud fraction and condensed water content are calculated as diagnostic functions of the total water content, temperature and pressure. The scheme accounts for sub-grid inhomogeneity by assuming a symmetric triangular distribution for the difference between specific total water content and the saturation specific water vapour content within each grid-box. This has the effect of starting to form clouds when the grid-point value of relative humidity (RH) passes some critical value RH_{crit} somewhat below 100%. Cloud fraction and condensed water content then increase smoothly with relative humidity above this threshold. The scheme can be tuned via the free parameter RH_{crit} .

Note that the scheme of Cusack *et al.* 1999b used in HadGEM1, which calculates RH_{crit} from the variability between neighbouring grid-points based on a scaling relationship, is not used in this study. Clearly it cannot be formulated in an SCM. Instead, unless otherwise stated RH_{crit} is prescribed at a value of 80% at all levels except within the boundary layer, where it increases towards the surface to reach 95% at the lowest model level.

A.4 MICROPHYSICS

The scheme used is essentially the mixed-phase cloud and precipitation scheme of Wilson & Ballard (1999). There are four water species; vapour, cloud-water, ice and rain. Conversions between these phases are simulated by physical process equations, which depend upon parameterisations for the size distributions and fall-speeds of condensate species.

Rain is assumed to fall to the surface during the time-step in which it forms, whilst the other three species are prognostic variables in the model. The single variable for ice represents the sum of ice cloud particles, snow, hail etc. with an assumed continuous size distribution which has a longer tail towards large particle sizes as the total ice content increases. Ice is always assumed to fall at a parameterised speed, which is a function of the particle size. Combining this with the assumed distribution of ice particle sizes yields a bulk ice fall-rate which increases with specific ice content. Liquid water can only fall by conversion to rain. In the absence of ice formation, this can only occur if the liquid water content exceeds an auto-conversion threshold, beyond which a “warm rain” parameterisation takes effect.

A.5 RADIATION

The fully interactive 2-stream radiative transfer scheme of Edwards & Slingo (1996) is used, with gaseous absorption parameterised according to Cusack *et al.* (1999a). The radiative effect of clouds is represented by the condensate variables and cloud fraction produced by the large-scale cloud and microphysics schemes, along with assumptions about condensate particle sizes and overlap of partial cloud at different levels.

However, cloud condensate in the convection parameterisation is not passed to the model’s condensate variables but is all either rained out or evaporated. To represent the radiative effect of convective cloud, the convection scheme produces diagnostics for the total cloud condensate in the bulk plume, the convective cloud fraction, and height of the cloud base and top, which are used to construct profiles of convective cloud properties for the radiation scheme. The spread-out “anvil clouds” seen at the tops of deep convective plumes are represented according to Gregory

(1999) by applying a larger cloud fraction on model-levels near the cloud top.

APPENDIX B:

STATISTICAL METHODS USED TO

TEST THE SIGNIFICANCE OF

DIFFERENCES BETWEEN

AQUA-PLANET SIMULATIONS

This appendix describes the statistical methods used to ascertain which differences between the aqua-planet simulations described in chapter 3 were likely to have resulted from systematic effects of the stochastic parameterisations, and which were not significant enough for such a claim to be justifiable. Section B.1 describes the statistical test used to assess differences in the mean-states of the simulations, whilst section B.2 describes the test used to assess differences in variance. Section B.3 details the method used to estimate effective sample sizes for use in these statistical tests, to adequately account for autocorrelation in the data.

B.1 TEST FOR DIFFERENCES IN MEAN-STATES

The Welch T-test (Welch 1947) is used to test whether each stochastic aqua-planet simulation's mean-state differs significantly from that of the control run. In this method, the test statistic for two samples x_1 and x_2 is given by:

$$T = \frac{\bar{x}_2 - \bar{x}_1}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \tag{B.1}$$

where s_1^2 and s_2^2 are the unbiased estimators of the sample variances. In the case of the null

hypothesis where the samples (model runs) are drawn from populations (hypothetical infinitely long model runs) with equal means, T follows a Student's T distribution with degrees of freedom given by:

$$d.o.f. = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left(\frac{s_1^2}{n_1}\right)^2/(n_1 - 1) + \left(\frac{s_2^2}{n_2}\right)^2/(n_2 - 1)} \quad (\text{B.2})$$

If T exceeds the cumulative T-distribution score (with the above degrees of freedom) for a specified probability threshold, the test may reject the null hypothesis and conclude a significant difference between the mean-states of the model runs. To account for autocorrelation in the data, reduced effective sample sizes are substituted for n_1 and n_2 in equations B.1 and B.2 when performing the T-test.

B.2 TEST FOR DIFFERENCES IN VARIANCE

Levene's test (Levene 1960) is used to test whether the variances of the stochastic simulations differ significantly from those in the control run. This test uses the absolute deviates of variables from their sample means to test for differences in variance across any number of samples being compared. For the case of a comparison between two samples, the test statistic is:

$$W = (n_1 + n_2 - 2) \frac{n_1(\bar{Z}_1 - \bar{Z})^2 + n_2(\bar{Z}_2 - \bar{Z})^2}{\sum_{j=1}^{n_1} (Z_{1,j} - \bar{Z}_1)^2 + \sum_{j=1}^{n_2} (Z_{2,j} - \bar{Z}_2)^2} \quad (\text{B.3})$$

where n_1 and n_2 are the samples sizes, Z_1 and Z_2 are the absolute deviates $|x - \bar{x}|$ of the sample variables x from their means, and \bar{Z} is the combined mean of Z_1 and Z_2 . In the case of the null hypothesis where the samples are drawn from populations with equal variance, W follows an f-distribution with degrees of freedom 1 and $n_1 + n_2 - 2$. Thus if W exceeds the cumulative f-distribution score for a specified probability threshold, the test may reject the null hypothesis and conclude a significant difference in variance between the two model runs. To incorporate reduced effective sample sizes as described in chapter 4, equation B.3 can be re-written:

$$W = (n_1 + n_2 - 2) \frac{\frac{1}{n_2}(\bar{Z}_1 - \bar{Z})^2 + \frac{1}{n_1}(\bar{Z}_2 - \bar{Z})^2}{\frac{1}{n_2}(\bar{Z}_1 - \bar{Z}_1)^2 + \frac{1}{n_1}(\bar{Z}_2 - \bar{Z}_2)^2} \quad (\text{B.4})$$

To properly account for autocorrelation when establishing the significance of a difference in variance, the effective sample sizes need to be substituted for n_1 and n_2 in equation B.4 *and* in the degrees of freedom used in the f-distribution test score.

B.3 ESTIMATION OF EFFECTIVE SAMPLE SIZE

It was shown in chapter 4 that a commonly used method to estimate the effective sample size of autocorrelated data, using only the lag-1 autocorrelation coefficient, would yield misleading results in this study. This is because the aqua-planet data exhibits complex autocorrelation on a range of scales and so cannot be modelled as a first-order autoregressive process, as that method assumes. Instead, effective sample sizes are estimated based on the sampling distributions of the statistics concerned (mean and variance), as quantified by repeatedly sub-sampling the data. The method used to estimate the effective sample size for comparing variances is described in subsection B.3.1, and that for comparing means in B.3.2

B.3.1 EFFECTIVE SAMPLE SIZE FOR THE SAMPLE VARIANCE

For data which is approximately normal, mean-squared-differences of the data from their mean follow a χ^2 distribution. This can be written in terms of the variance of a sample as

$$n_s \frac{\sigma_s'^2}{\sigma_p^2} \sim \chi_{n_s-1}^2 \quad (\text{B.5})$$

where n_s is the effective size of sub-samples, σ_p^2 is the population variance, and $\sigma_s'^2$ is the sample variance. Here and elsewhere in this section, the superscript ' denotes the sample variance normalised by n_s (not $n_s - 1$ as in the un-biased estimator of variance). An un-biased estimate of σ_p^2 is related to the variance of the whole available data-set, $\sigma_d'^2$, by $s_d^2 = \sigma_d'^2 n_{d\sigma} / (n_{d\sigma} - 1)$, where $n_{d\sigma}$ is the effective sample size of the whole data-set (with respect to variance), which is not

known *a priori*. If sub-samples are selected by evenly dividing up the entire data-set, we expect the overall sample-size to be the number of sub-samples m_s times the size of each sub-sample ($n_{d\sigma} = m_s n_s$). Substituting the above in (B.5), we obtain

$$\frac{(m_s n_s - 1) \sigma_s^2}{m_s \sigma_d^2} \sim \chi_{n_s - 1}^2 \quad (\text{B.6})$$

The χ^2 distribution with $n_s - 1$ degrees of freedom has mean $n_s - 1$ and variance $2(n_s - 1)$. Thus expressions for the mean $\mu_{\sigma_s^2}$ and variance $\sigma_{\sigma_s^2}^2$ of the sampling distribution of sub-sample variances can be obtained:

$$\mu_{\sigma_s^2} = (n_s - 1) \frac{m_s}{m_s n_s - 1} \sigma_d^2 \quad (\text{B.7})$$

$$\frac{m_s}{m_s - 1} \sigma_{\sigma_s^2}^2 = 2(n_s - 1) \left(\frac{m_s}{m_s n_s - 1} \right)^2 \sigma_d^4 \quad (\text{B.8})$$

Note that on the left-hand side of B.8, the theoretical variance of sub-sample variances has been replaced by an un-biased estimator based on the variance of the finite sample of m_s sub-samples. Taking (B.8), dividing by (B.7) and rearranging, a simple expression for the effective sample size of the data-set, $n_{d\sigma} = m_s n_s$, can be obtained:

$$n_{d\sigma} = 1 + 2(m_s - 1) \frac{\sigma_d^2 \mu_{\sigma_s^2}}{\sigma_{\sigma_s^2}^2} \quad (\text{B.9})$$

This equation estimates $n_{d\sigma}$ based on the width of the sampling distribution of sub-sample variances relative to the population variance and the mean of the sub-sample variances. The less spread the sub-sample variances are, the more comprehensive the sampling within each sub-sample, and the greater the implied degrees of freedom.

However, even when taking sub-samples of a substantial size, some autocorrelation between the variances of adjacent samples may occur, and will need to be accounted for if present (be it far smaller than the autocorrelation between adjacent points in the original data). Otherwise, assuming the sub-samples to be independent will lead to an over-estimate of the effective sample

size. If the sub-samples each cover an appropriately long period of time, many of the dynamical modes which lead to complex non-first-order autocorrelations in the data will be averaged out and accounted for by the sub-sampling method, and the remaining inter-sub-sample variability should be closer to a first-order autoregression model than the original data. Therefore, for the statistical tests used in this study, it was deemed appropriate to use the Dawdy & Matalas (1964) method to estimate the effective size of the sample of sub-samples, to account for timescales of variability in the data which are longer than the sub-sample size. To do this, m_s in equation B.9 is calculated as:

$$m_s = \frac{1 - r_{\sigma_s^2}}{1 + r_{\sigma_s^2}} m_{sub} \quad (\text{B.10})$$

where $r_{\sigma_s^2}$ is the autocorrelation of adjacent sub-sample variances, and m_{sub} is the actual number of sub-samples.

B.3.2 EFFECTIVE SAMPLE SIZE FOR THE SAMPLE MEAN

According to the Central Limit Theorem, the mean μ_s of a sample s drawn from a population with mean μ_p and variance σ_p^2 follows a sampling distribution which rapidly tends towards normal with increasing sample size n_s , with mean μ_p and variance $\sigma_{\mu_s}^2$ given by

$$\sigma_{\mu_s}^2 = \sigma_p^2 / n_s \quad (\text{B.11})$$

Thus, if sub-samples are drawn from the available data and means computed for each, the variance $\sigma_{\mu_s}^2$ of the means can be used to estimate the effective size n_s of each sub-sample. Note that n_s here is the effective sub-sample size with respect to the mean, and so may differ from that in earlier equations which were with respect to the variance. For unbiased results, unbiased estimators should be used for both $\sigma_{\mu_s}^2$ and σ_p^2 . These are $s_{\mu_s}^2 = \sigma_{\mu_s}^2 m_s / (m_s - 1)$ and $s_d^2 = \sigma_d^2 n_{d\sigma} / (n_{d\sigma} - 1)$ respectively, where m_s is the number of sub-samples, $n_{d\sigma}$ is the effective sample size of the data-set with respect to variance, calculated using equation B.9, $\sigma_{\mu_s}^2$ is the variance of the finite sample of sample-means, and σ_d^2 is the variance of the whole available data-set. As before, the ' denotes

variances normalised by sample size rather than degrees of freedom. Substituting the above in equation B.11, we obtain:

$$\sigma_{\mu_s}^2 \frac{m_s}{m_s - 1} = \sigma_d^2 / n_s \frac{n_{d\sigma}}{n_{d\sigma} - 1} \quad (\text{B.12})$$

which can be trivially rearranged to yield an expression for the effective sample size of the whole data-set with respect to sample means, $n_{d\mu} = m_s n_s$:

$$n_{d\mu} = (m_s - 1) \frac{\sigma_d^2}{\sigma_{\mu_s}^2} \frac{n_{d\sigma}}{n_{d\sigma} - 1} \quad (\text{B.13})$$

This equation estimates $n_{d\mu}$ based on the width of the sampling distribution of sub-sample means; the less spread the sub-sample means are, the greater the implied degrees of freedom within each sub-sample.

As with the effective sample size with respect to variance, autocorrelation between the means of adjacent sub-samples can then be accounted for using the Dawdy & Matalas (1964) method to estimate the effective number of samples m_s to use in (B.13):

$$m_s = \frac{1 - r_{\mu_s}^2}{1 + r_{\mu_s}^2} m_{sub} \quad (\text{B.14})$$

where $r_{\mu_s}^2$ is the autocorrelation of adjacent sub-sample means, and m_{sub} is the actual number of sub-samples.

References

- Amodei, M., Pawson, S., Scaife, A. A., Langematz, U., Lahoz, W., Li, D. M. & Simon, P. 2001: The SAO and Kelvin waves in the EuroGRIPS GCMS and the UK Met. Office analyses. *Ann. Geophys.* **19**, 99–114.
- Arakawa, A. & Schubert, W. H. 1974: Interaction of a cumulus cloud ensemble with the large-scale environment. Part I. *J. Atmos. Sci.* **31**, 674–701.
- Arribas, A. 2004: Results of an initial stochastic physics scheme for the Met Office Unified Model. *Forecasting research technical report, Met Office, UK*. No. 452.
- Ball, M. A. & Plant, R. S. 2008: Comparison of stochastic parameterisation approaches in a single-column model. *Phil. Trans. R. Soc. A* **366**, 2605–2623.
- Bell, S. 2008: The Met Office NWP system: a) Description of current system, b) Recent and planned changes. *Met Office Scientific Advisory Committee (MOSAC) report, Met Office, UK*. Paper 13.1a.
- Berner, J., Doblas-Reyes, F. J., Palmer, T. N., Shutts, G. & Weisheimer, A. 2008: Impact of a quasi-stochastic cellular automaton backscatter scheme on the systematic error and seasonal prediction skill of a global climate model. *Phil. trans. R. Soc. A*. **366**, 2561–2579.
- Berner, J., Shutts, G. J., Leutbecher, M., & Palmer, T. N. 2009: A Spectral Stochastic Kinetic Energy Backscatter Scheme and Its Impact on Flow-Dependent Predictability in the ECMWF Ensemble Prediction System. *J. Atmos. Sci.*, **66**, 603–626.
- Bowler, N. E., Arribas, A., Mylne, K. R., Robertson, K. B., & Beare, S. E. 2008: The mogreps short-range ensemble prediction system. *Q. J. R. Meteor. Soc.* , **134**, 703–722.

- Bowler, N. E., Arribas, A., Beare, S. E., Mylne, K. R., & Shutts, G. J. 2009: The Local ETKF and SKEB: Upgrades to the MOGREPS Short-range Ensemble Prediction System. *Q. J. R. Meteorol. Soc.*, **135**, 767–776.
- Bretherton, C. S., Peters, M. E. & Back, L. E. 2004: Relationships between water vapour path and precipitation over the tropical oceans. *J. Climate* **17**, 1517–1528.
- Bright, D.R. & Mullen, S.L. 2002: Short-range ensemble forecasts of precipitation during the southwest monsoon. *Wea. Forecasting*, **17**, 1080–1100.
- Buizza, R., Miller, M. & Palmer, T. N. 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Q. J. R. Meteorol. Soc.* **125**, 2887–2908.
- Charney, J. G. 1963: A Note on Large-Scale Motions in the Tropics. *J. Atmos. Sci.* **20**, 607–609.
- Charron, M., Pellerin, G., Spacek, L., Houtekamer, P. L., Gagnon, N., Mitchell, H. L., & Michelin, L. 2010: Toward Random Sampling of Model Error in the Canadian Ensemble Prediction System. *Mon. Wea. Rev.*, **138**, 1877–1901.
- Ciesielski, P. E., Johnson, R. H., Haertel, P. T., Wang, J. 2003: Corrected TOGA COARE sounding humidity data: impact on diagnosed properties of convection and climate over the warm pool. *J. Climate* **16**, 2370–2384.
- Cohen, B. G. & Craig, G. C. 2006: Fluctuations in an equilibrium convective ensemble. Part II: Numerical experiments. *J. Atmos. Sci.* **63**, 2005–2015.
- Craig, G. C. & Cohen, B. G. 2006: Fluctuations in an equilibrium convective ensemble. Part I: Theoretical formulation. *J. Atmos. Sci.* **63**, 1996–2004.
- Cullen, M. J. P. 1993: The unified forecast/climate model. *Meteorol. Mag.* **122**, 89–94.
- Cusack, S., Edwards, J. M., & Crowther, J. M. 1999a: Investigating k-distribution methods for parameterising gaseous absorption in the Hadley Centre climate model. *J. Geophys. Res.* **104**, 2051–2057.
- Cusack, S., Edwards, J. M. & Kershaw, R. 1999b: Estimating the subgrid variance of saturation, and its parameterisation for use in a GCM cloud scheme. *Q. J. R. Meteorol. Soc.* **125**, 3057–3076.

- Dai, A. 2006: Precipitation characteristics in eighteen coupled climate models. *J. Climate* **19**, 4605–4630.
- Davies, T., Cullen, M. J. P., Malcolm, A. J., Mawson, M. H., Staniforth, A., White, A. A. & Wood, N. 2005: A new dynamical core for the Met Office’s global and regional modelling of the atmosphere. *Q. J. R. Meteorol. Soc.* **131**, 1759–1782.
- Dawdy, D. R., and Matalas, N. C., 1964: Statistical and probability analysis of hydrologic data, part III: Analysis of variance, covariance and time series. In *Ven Te Chow, ed., Handbook of applied hydrology, a compendium of water-resources technology: New York, McGraw-Hill Book Company*, p. 8.68–8.90.
- Done, J. M. 2002: Predictability and representation of convection in a mesoscale model. *PhD thesis, University of Reading, UK*, p.180.
- Emanuel, K. A. & Zivkovic-Rothman, M. 1999: Development and evaluation of a convection scheme for use in climate models. *J. Atmos. Sci.* **56**, 1766–1782.
- Frederiksen, J. S. & Davies, A. G. 1997: Eddy viscosity and stochastic backscatter parameterisations on the sphere for atmospheric circulation models. *J. Atmos. Sci.* **54**, 2475–2492.
- Fritsch, J. M. & Chappell, C. F. 1980: Numerical prediction of convectively driven mesoscale pressure systems. Part I: Convective parameterisation. *J. Atmos. Sci.* **37**, 1722–1733.
- Gage, K. S. & Nastrom, G. D. 1986: Theoretical interpretation of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft during GASP. *J. Atmos. Sci.* **43**, 729–740.
- Ghan, S., Randall, D., Xu, K.-M., Cederwall, R., Cripe, D., Hack, J., Iacobellis, S., Klein, S., Krueger, S., Lohmann, U., Pedretti, J., Robock, A., Rotstayn, L., Somerville, R., Stenchikov, G., Sud, Y., Walker, G., Xie, S., Yio, J. & Zhang, M. 2000: A comparison of single column model simulations of summertime midlatitude continental convection. *J. Geophys. Res.* **105**, 2091–2124.
- Grant, A. L. M. & Brown, A. R. 1999: A similarity hypothesis for shallow-cumulus transports. *Q. J. R. Meteorol. Soc.* **125**, 1913–1936.

- Grant, A. L. M. 2001: Cloud-base fluxes in the cumulus-capped boundary layer. *Q. J. R. Meteorol. Soc.* **127**, 407–421.
- Gray, M. E. B. & Shutts G. J. 2002: A stochastic scheme for representing convectively generated vorticity sources in general circulation models. *APR Turbulence and Diffusion Note* no. **285**, Met Office, UK.
- Gregory, D. & Rowntree, P. R. 1990: A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure. *Mon. Wea. Rev.* **118**, 1483–1506.
- Gregory, D. & Allen, S. 1991: The effect of convective scale downdrafts upon NWP and climate simulations. *Ninth Conf. on Numerical Weather Prediction, Denver, CO, Amer. Met. Soc.* 122–123.
- Gregory, D., Kershaw, R. & Inness, P. M. 1997: Parameterisation of momentum transport by convection. II: Tests in single-column and general circulation models. *Q. J. R. Meteorol. Soc.* **123**, 1153–1183.
- Gregory, J. 1999: Representation of the radiative effects of convective anvils. *Hadley Centre Technical Note 7*, pp. 20.
- Grell, G. A., Kuo, Y.-H. & Pasch, R. 1991: Semiprognostic tests of cumulus parameterisation schemes in the middle latitudes. *Mon. Wea. Rev.* **119**, 5–31.
- Hack, J. J., Schubert, W. H. & Dias, P. L. S. 1984: A spectral cumulus parameterisation for use in numerical models of the tropical atmosphere. *Mon. Wea. Rev.* **112**, 704–716.
- Holloway, C. E. & Neelin, J. D. 2009: Moisture vertical structure, column water vapour, and tropical deep convection. *J. Atmos. Sci.* **66**, 1665–1683.
- Horinouchi, T., Pawson, S., Shibata, K., Langematz, U., Manzini, E., Giorgetta, M. A., Sassi, F., Wilson, R. J., Hamilton, K., de Grandpre, J. & Scaife, A. A. 2003: Tropical cumulus convection and upward propagating waves in middle-atmospheric GCMs. *J. Atmos. Sci.* **60**, 2765–2782.
- Houtekamer, P. L., Lefaiivre, L., Derome, J., Ritchie, H. & Mitchell, H. L., 1996: A System Simulation Approach to Ensemble Prediction. *Mon. Wea. Rev.*, **124**, 1225–1242.

- Hume, T. & Jakob, C. 2005: Ensemble single column modeling (ESCM) in the tropical western Pacific: forcing data sets and uncertainty analysis. *J Geophys. Res.* **110**, D13109.
- Inness, P. M., Slingo, J. M., Woolnough, S. J., Neale, R. B. & Pope, V. D. 2001: Organisation of tropical convection in a GCM with varying vertical resolution; implications for the simulation of the Madden-Julian oscillation. *Climate Dyn.* **17**, 777–793.
- Jung, T., Palmer, T. N. & Shutts, G. J. 2005: Influence of a stochastic parameterisation on the frequency of occurrence of North Pacific weather regimes in the ECMWF model. *Geophys. Res. Lett.* **32**, L23811.
- Kain, J. S. & Fritsch, J. M. 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterisation. *J. Atmos. Sci.* **47**, 2784–2802.
- Kain, J. S. 2004: The Kain-Fritsch convective parameterisation: an update. *J. Appl. Meteorol.* **43**, 170–181.
- Kalnay, E. 2003: Atmospheric modelling, data assimilation and predictability. *Cambridge University Press*, ISBN 0-521-79179-0, ISBN 0-521-79629-6 (pbk.)
- Keane, R. J. & Plant, R. S. 2012: Large-scale length and time-scales for use with stochastic convective parametrization. *Q. J. R. Meteorol. Soc.* **138**, 1150–1164.
- Krueger, S. K. & Lazarus, S. M. 1998: Intercomparison of multi-day simulations of convection during TOGA COARE with several Cloud-Resolving and Single-Column Models. In *23rd Conf. on Hurricanes and Tropical Meteorology, Dallas, TX. Amer. Met. Soc.*, 643–647.
- Levene, H. 1960: Robust tests for equality of variances. In *Contributions to Probability and Statistics (I. Olkin, ed.)*, Stanford Univ. Press, Palo Alto, CA., 278–292.
- Li, X., Charron, M., Spacek, L., & Candille, G. 2008: A Regional Ensemble Prediction System Based on Moist Targeted Singular Vectors and Stochastic Parameter Perturbations. *Mon. Wea. Rev.*, **136**, 443–462.
- Lin, J. W.-B., & Neelin, J.D. 2000: Influence of a stochastic moist convective parameterization on tropical climate variability. *Geophys. Res. Lett.*, **27**, 3691–3694.

- Lin, J. W.-B., & Neelin, J. D. 2002: Considerations for stochastic convective parameterisation. *J. Atmos. Sci.* **59**, 959–975.
- Lin, J.-L., Kiladis, G. N., Mapes, B. E., Weickmann, K. M., Sperber, K. R., Lin, W., Wheeler, M. C., Schubert, S. D., del Genio, A., Donner, L. J., Emora, S., Gueremy, J.-F., Hourdin, F., Rasch, P. J., Roeckner, E., & Scinocca, J. F. 2006: Tropical intraseasonal variability in 14 IPCC AR4 climate models. Part I: convective signals. *J. Climate*. **19**, 2665–2690.
- Lock, A. P., Brown, A. R., Bush, M. R., Martin, G. M. & Smith, R. N. B. 2000: A new boundary layer mixing scheme. Part I: Scheme description and SCM tests. *Mon. Wea. Rev.* **128**, 3187–3199.
- Lord, S. J. 1982: Interaction of a cumulus cloud ensemble with the large-scale environment. Part III: semi-prognostic test of the Arakawa-Schubert cumulus parameterization. *J. Atmos. Sci.* **39**, 88–103.
- Lorenz, E. N. 1963: Deterministic Nonperiodic Flow. *J. Atmos. Sci.* **20**, 130–141.
- Lorenz, E. N. 1969: Atmospheric Predictability by Naturally Occuring Analogues. *J. Atmos. Sci.* **26**, 636–646.
- Lorenz, E. N. 1975: Climate predictability. The physical basis of climate modelling, *W.M.O. GARP Pub.Ser.*, **16**, 132–136.
- Luo, D., Huang, F. & Diao, Y. 2001: Interaction between antecedent planetary-scale envelope soliton blocking anticyclone and synoptic scale eddies: Observations and theory. *J. Geophys. Res.* **106**, 31795–31815.
- Martin, G. M., Ringer, M. A., Pope, V. D., Jones, A., Dearden, C. & Hinton, T. J. 2006: The physical properties of the atmosphere in the new Hadley Centre Global Environmental Model (HadGEM1). Part I: model description and global climatology. *J. Climate* **19**, 1274–1301.
- Mason, P. J. & Thomson, D. J. 1992: Stochastic backscatter in large-eddy simulations of boundary layers. *J. Fluid Mech.*, **242**, 61–78.

- Molteni, F. & Tibaldi, S. 1990: Regimes in the wintertime circulation over the northern extratropics. II: Consequences for dynamical predictability. *Q. J. R. Meteorol. Soc.* **116**, 1263–1288.
- Moncrieff, M. W., Krueger, S. K., Gregory, D., Redelsperger, J.-L. & Tao, W.-K. 1997: GEWEX Cloud System Study (GCSS) Working Group 4: precipitating Convective Cloud Systems. *Bull. Amer. Met. Soc.* **78**, 831–845.
- Murphy, J., Sexton, D. M. H., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M. & Stainforth, D. A. 2004: Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature* **430**, 768–772.
- Mylne, K.R., Evans, R.E. & Clark, R.T. 2001: Multi-model multi-analysis ensembles in quasi-operational medium-range forecasting. *Q. J. R. Meteorol. Soc.*, **128**, 361–384.
- Neale, R. B. & Hoskins, B. J. 2001a: A standard test for AGCMs including their physical parameterisations: I: The proposal. *Atmos. Sci. Lett.* **1** (doi:10.1006/asle.2000.0019).
- Neale, R. B. & Hoskins, B. J. 2001b: A standard test for AGCMs including their physical parameterisations: II: Results for the Met Office model. *Atmos. Sci. Lett.* **1** (doi:10.1006/asle.2000.0020).
- Neale, R. B. & Slingo, J. M. 2003: The Maritime Continent and Its Role in the Global Climate: A GCM Study. *J. Climate* **16**, 834–848.
- Palmer, T. N. 2001: A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parameterization in weather and climate prediction. *Q. J. R. Meteorol. Soc.*, **127**, 279–304.
- Palmer, T. N., Buizza, R., Doblas-Reyes, F., Jung, T., Leutbecher, M., Shutts, G. J., Steinheimer, M., & Weisheimer, A. 2009: Stochastic Parameterisation and Model Uncertainty. *ECMWF Research Department*, Technical Memorandum No. 598.
- Panofsky, H. A. & Brier, G. W. 1968: Some Applications of Statistics to Meteorology. *Pennsylvania State University Press*, pp 224.

- Petch, J. C., Willett, M., Wong, R. Y. & Woolnough, S. J. 2007: Modelling suppressed and active convection. Comparing a numerical weather prediction, cloud-resolving and single-column model. *Q. J. R. Meteorol. Soc.* **133**, 1087–1100.
- Piani, C., Frame, D. J., Stainforth, D. A. & Allen, M. R. 2005: Constraints on climate change from a multi-thousand member ensemble of simulations. *Geophys. Res. Lett.* **32**, L23825.
- Plant, R. S., & Craig, G. C. 2008: A stochastic parameterization for deep convection based on equilibrium statistics. *J. Atmos. Sci.*, **64**, 87–105.
- Rajendran, K., Kitoh, A., Mizuta, R., Sajani, S. & Nakazawa, T. 2008: High-resolution simulation of mean convection and its intraseasonal variability over the tropics in the MRI/JMA 20-km mesh AGCM. *J. Climate* **21**, 3722–3739.
- Randall, D. A., Xu, K.-M., Somerville, R. J. C. & Iacobellis, S. 1996: Single-column models and cloud ensemble models as links between observations and climate models. *J. Climate* **9**, 1683–1697.
- Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R. J., Sumi, A. & Taylor, K. E. 2007: Climate Models and Their Evaluation. In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., & Miller, H. L. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Ricciardulli, L. & Garcia, R. R. 2000: The excitation of equatorial waves by deep convection in the NCAR Community Climate Model (CCM3). *J. Atmos. Sci.* **57**, 3461–3487.
- Scinocca, J. F. & McFarlane, N. A. 2004: The variability of modelled tropical precipitation. *J. Atmos. Sci.* **61**, 1993–2015.
- Shutts, G. J. 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Q. J. R. Meteorol. Soc.*, **131**, 3079–3102.

- Shutts, G. J. & Palmer, T. N. 2007: Convective forcing fluctuations in a cloud-resolving model: Relevance to the stochastic parameterisation problem. *J. Climate* **20**, 187–202.
- Smith, R. N. B. 1990: A scheme for predicting layer clouds and their water content in a general circulation model. *Q. J. R. Meteorol. Soc.* **116**, 435–460.
- Slingo, J. M., Inness, P. M., Neale, R. B., Woolnough, S. J. & Yang, G.-Y. 2003: Scale interactions on diurnal to seasonal timescales and their relevance to model systematic errors. *Ann. Geophys.* **46**, 139–155.
- Stainforth, D. A., Aina, T., Christensen, C., Collins, M., Faull, N., Frame, D. J., Kettleborough, J. A., Knight, S., Martin, A., Murphy, J. M., Piani, C., Sexton, D. M. H., Smith, L. A., Spicer, R. A., Thorpe, A. J. & Allen, M. R. 2005: Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature* **433**, 403–406.
- Stiller, O. 2009: Efficient moist physics schemes for data assimilation. II: Deep convection. *Q. J. R. Meteorol. Soc.* **135**, 721–738.
- Teixeira, João, & Reynolds, C. A. 2008: Stochastic Nature of Physical Parameterizations in Ensemble Prediction: A Stochastic Convection Approach. *Mon. Wea. Rev.*, **136**, 483–496.
- Tennant, W. J., Shutts, G. J., Arribas, A., & Thompson, S. A. 2011: Using a Stochastic Kinetic Energy Backscatter Scheme to Improve MOGREPS Probabilistic Forecast Skill. *Mon. Wea. Rev.*, **139**, 1190–1206.
- Van Den Dool, H. M. & Chervin, R. M. 1986: Comparison of Month-to-Month Persistence of Anomalies in a General Circulation Model and in the Earth’s Atmosphere. *J. Atmos. Sci* **43**, 1454–1466.
- Wang, X. & Bishop, C. H. 2003: A comparison of breeding and ensemble transform kalman filter ensemble forecast schemes. *J. Atmos. Sci.* **60**, 1140–1158.
- Webster, P. J. & Lukas, R. 1992: TOGA COARE: the coupled ocean-atmosphere response experiment. *Bull. Amer. Met. Soc.* **73**, 1377–1416.
- Welch, B. L. 1947: The generalization of “Student’s” problem when several different population variances are involved. *Biometrika* **34** (1-2): 28–35, doi:10.1093/biomet/34.1-2.28, MR19277

- Willett, M. R. & Milton, S. F. 2006: The tropical behaviour of the convective parameterisation in aqua-planet simulations and the sensitivity to timestep. *Forecasting research technical report, Met Office, UK*. No. 482.
- Williams, P. D. 2005: Modelling climate change: the role of unresolved processes. *Phil. Trans. R. Soc. A* **363**, 2931–2946.
- Wilson, D. R., & Ballard, S. P. 1999: A microphysically based precipitation scheme for the Met Office Unified Model. *Q. J. R. Meteorol. Soc.* **125**, 1607–1636.
- Wu, X., Moncrieff, M. W. & Emanuel, K. A. 2000: Evaluation of large-scale forcing during TOGA COARE for Cloud Resolving Models and Single-Column Models. *J. Atmos. Sci.* **57**, 2977–2985.
- Xie, S., Xu, K.-M., Cederwall, R. T., Bechtold, P., Genio, A. D. D., Klein, S. A., Cripe, D. G., Ghan, S. J., Gregory, D., Iacobellis, S. F., Krueger, S. K., Lohmann, U., Petch, J. C., Randall, D. A., Rotstayn, L. D., Somerville, R. C. J., Sud, Y. C., Salzen, K. V., Walker, G. K., Wolf, A., Yio, J. J., Zhang, G. J. & Zhang, M. (2002): Intercomparison and evaluation of cumulus parametrizations under summertime midlatitude continental conditions. *Q. J. R. Meteorol. Soc.* **128**, 1095–1135.
- Yang, G.-Y. & Slingo, J. M. 2001: The diurnal cycle in the tropics. *Mon. Weath. Rev.* **129**, 784–801.
- Yang, Z. & Arritt, R.W. 2002: Tests of a perturbed physics ensemble approach for regional climate modelling. *J. Clim.*, **15**, 2881–2896.