Comparison of stochastic parameterisation approaches in a single-column model

Michael A. Ball and Robert S. Plant
Department of Meteorology, University of Reading
PO Box 243, Reading, Berkshire RG6 2BB, UK.

July 23, 2008

Abstract

We discuss and test the potential usefulness of single-column models (SCMs) for the testing of stochastic physics schemes that have been proposed for use in general circulation models (GCMs). We argue that although single column tests cannot be definitive in exposing the full behaviour of a stochastic method in the full GCM, and although there are differences between SCM testing of deterministic and stochastic methods, SCM testing remains a useful tool. It is necessary to consider an ensemble of SCM runs produced by the stochastic method. These can be usefully compared to deterministic ensembles describing initial condition uncertainty and also to combinations of these (with structural model changes) into poor man's ensembles. The proposed methodology is demonstrated using an SCM experiment recently developed by the GCSS (GEWEX Cloud System Study) community, simulating transitions between active and suppressed periods of tropical convection.

Keywords:
stochastic parameterisation, single-column model, ensemble, tropical convection

1 Introduction

In recent years, increasing attention has been given to the potential usefulness (Palmer 2001, Wilks 2005) of introducing some stochastic component(s) to the physical parameterisations used in GCMs. For example, many GCMs are known to have have insufficient high-frequency, small-scale variability of convective heating rates and precipitation in the tropics, which may damage their ability to represent low-frequency, large-scale climate variability (Ricciardulli and Garcia 2000; Horinouchi *et al.* 2003). A wide variety of plausible stochastic methods continue to be suggested and actively investigated, including perturbing the inputs to a parameterisation (e.g. Tompkins and Berner submitted), perturbing the parameters used within it (e.g. Byun and Hong 2007), perturbing its outputs (e.g. Teixeira and Reynolds 2008), and even constructing new parameterisations designed explicitly to be stochastic from the outset (e.g. Plant & Craig 2008). There is a growing acceptance that the use of stochastic elements in GCMs may be desirable for both theoretical and practical reasons (e.g. Penland 2003, Williams 2005). Thus, the time may soon be approaching when the key question changes from *why a stochastic method?* to *which stochastic method?* Here, we explore whether single column modelling might be able to provide some insights that could inform such decision-making.

The aim of a stochastic scheme is to introduce variability into the numerical representation of the climate system. In order to determine the variability of some climate phenomenon in the GCM, either multiple or long integrations are likely to be required. Further lengthy explorations would also be required if one wished to assess the impact of that variability on other aspects of the model climate. How then, in practice, should one choose the stochastic method(s) to be used in a GCM? The difficulty is not simply the range of possible schemes available in the literature, but also (at least) two other important considerations.

First, we do not know how well various methods might combine. The motivations behind various schemes, and the uncertainties they attempt to address, may often appear to be very different. At first sight then it may be attractive to use several methods. However, there may be a danger in this of some "double counting", particularly if attempting to combine some of the more generic methods to address parameterisation uncertainty. For instance, taking a single parameterisation and perturbing its inputs, parameters and outputs simultaneously might not be totally unreasonable, but it would be extremely naive to expect good performance by implementing three such methods directly "off-the-shelf".

Second, one's difficulties are compounded by the fact that many (if not all) of the stochastic schemes in the literature themselves contain free parameters and structural uncertainties. We can use the random parameters approach to offer a simple example. Suppose that one wished to choose random values for the entrainment rate and the CAPE closure timescale in the parameterisation of deep convection. Should those choices be correlated, and if so, then how?

It would surely be impractical to conduct full GCM testing of all plausible stochastic physics schemes and all possible variations on their basic themes. However, it should be possible to do better than testing some best guesses. As a first step, we describe in this paper essentially a test-of-concept for the idea that single-column model (SCM) experiments might have some useful value for comparing stochastic schemes. We are not at this stage attempting an assessment of the relative performance of various stochastic schemes. Rather, our objective is to demonstrate that simple methods to improve one's understanding of the behaviour of various stochastic schemes are both possible and worth pursuing. It does not seem unreasonable to hope that current best guesses could evolve into educated guesses.

The paper is organised as follows. In Section 2 we introduce some issues in single-column modelling and their implications for testing stochastic methods. Section 3 describes the modelling framework used in this study, including the stochastic methods (3.1) and the ensemble approach (3.2). Results (section 4) are shown for the sensitivity to initial condition (IC) perturbations (4.1), and for the mean states (4.2) and variabilities (4.3 and 4.4) of various stochastic methods. Conclusions are drawn in Section 5.

2 An SCM Approach for Stochastic Schemes

Single-column modelling has a long history as a useful guide towards understanding and testing the behaviour of deterministic parameterisations within a GCM. In the full GCM, a parameterisation interacts with model dynamics and with the other parameterisations.

Essentially the SCM is a means to understand the latter, which may or may not dominate in the full GCM. Usually the dynamical forcing of the SCM is determined beforehand, perhaps based on an observational campaign. The forcing is independent of the current model state, which constrains the possible responses. Thus the SCM may behave differently from the corresponding full GCM if dynamical feedback is an important aspect of the situation modelled. One consequence is that parameterisation errors in a GCM which adversely affect model dynamics may not be apparent in the SCM, which is kept on track by the prescribed dynamics.

It is not immediately apparent how an SCM might be used to make meaningful comparisons of stochastic methods. In some cases, the use of a single column may simply not be viable because the stochastic terms cannot easily be applied (e.g. Shutts 2005). Indeed, it has been suggested that an ideal stochastic method would probably be non-local (Palmer 2001; Craig *et al.* 2005; Ghil *et al.* 2005). For the present though, the majority of stochastic methods can be formulated for a single column. Nonetheless, an obvious objection to SCM comparisons remains: feedback from the introduced variability to the dynamics may be a key feature of the behaviour in a full GCM (e.g. Lin and Neelin 2002). This will be missing in a traditional SCM experiment with specified dynamics. SCMs can include appropriate dynamical feedbacks by using a parameterised dynamics formulation, such as a weak-temperature gradient approximation for the tropical atmosphere (e.g. Sobel *et al.* 2007), or by coupling vertical advection to the parameterised diabatic heating via a gravity-wave model (Bergman & Sardeshmukh 2004). Ultimately, we believe that these and similar frameworks would be particularly well-suited to studying stochastic physics schemes, but do not pursue them further here.

It may nonetheless be possible to gain some useful insights into the behaviour of stochastic methods through an SCM comparison. The results for each method must be considered in the form of an ensemble of SCM runs, each run having a different set of random numbers. Our proposal is to compare such ensembles to the SCM results obtained from multiple deterministic parameterisations, the suite of deterministic configurations being treated as a so-called poor-man's ensemble (e.g. Mylne *et al.* 2002).

3 Experimental setup

Experiments have been carried out using the single-column form of the UK Met Office Unified Model (UM, Cullen 1993). The model runs are based on GCSS PCCS case 5, the design of which is described by Petch *et al.* (2007). Specifically, we study here the consecutive time periods B and C. Model intercomparison cases have been a major part of the Global Energy and Water cycle EXperiment (GEWEX) Cloud System Study (GCSS), which aims to support the development of physically based parameterisations for cloud processes. An overview of the Precipitating Convective Cloud Systems (PCCS) working group can be found in Moncrieff *et al.* (1997).

Case 5 simulates a column of the atmosphere in the tropical West Pacific warm pool region, at 2°S 156°E, and the model runs presented here span the period 9–28 January 1993. The forcing data-set is derived from observations taken in the TOGA-COARE campaign (Webster and Lukas 1992). It contains temperature and moisture increments due to turbulent fluxes from the ocean surface and to large-scale vertical and horizontal advection[1]. Also prescribed are timeseries of observed winds, towards which the SCM is

[1]These data-sets are available for the whole of the TOGA-COARE observing period, along with

strongly relaxed, with a timescale of 1 hour. Any deviations of the winds in these runs from those observed are therefore limited. The forcings and initial conditions are derived directly from surface and radiosonde measurements averaged over the TOGA-COARE IFA (Intensive Flux Array, see figure 14 of Webster & Lukas 1992).

The focus of case 5 is the transition of tropical convection from suppressed to active phases, and two such transitions occur during these SCM runs. See figure 1 in which the periods are defined as in Petch *et al.* (2007). Here we label ActB, a very active period with heavy rain, SupC, in which convection is suppressed by the large scale forcing, and ActC, the subsequent active phase. Rain rates are similar to those found in other SCMs (Woolnough *et al.* submitted).

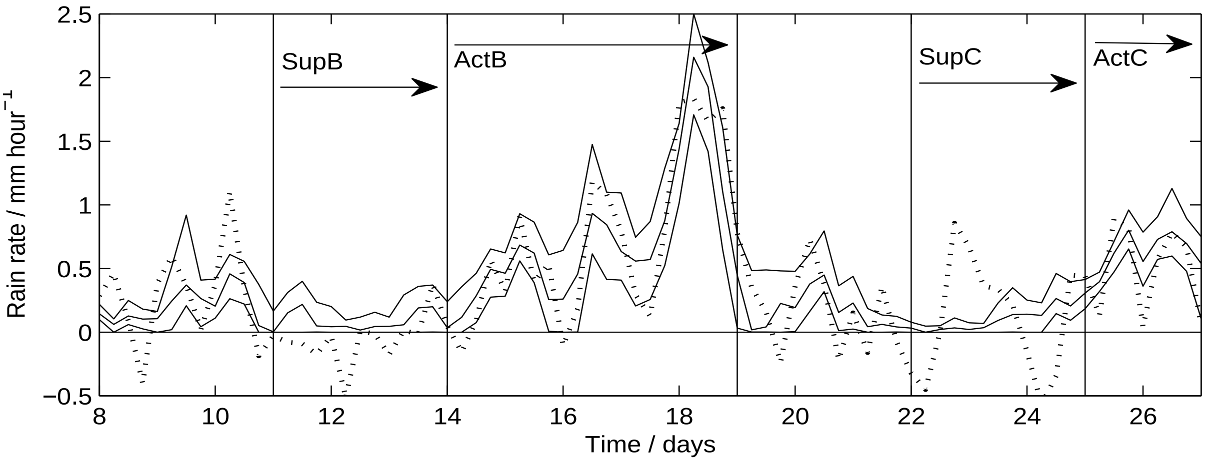

Figure 1: 6-hourly means of the 5th, 50th and 95th percentiles of rainfall rate from an ensemble of SCM runs using the default UM configuration (solid lines), and a budget-derived estimate for the TOGA-COARE IFA (dotted line). This estimate has some negative values, since the observations were insufficient to derive an accurate moisture budget (Petch *et al.* 2007). The time axis is labelled in whole days since the start of the month of January 1993, as in later figures. Annotations are explained in the text.

The SCM runs use a timestep of 30 minutes and there are 38 levels in the vertical. The performance of the default UM SCM for this case in comparison to other models is discussed by Petch *et al.* (2007) and Woolnough *et al.* (submitted). It is more consistent with CRM simulations than some of the SCMs, which were somewhat dry.

3.1 Model Variants

Taking the UM SCM as a basis, several model configurations have been tested, differing through either the convection parameterisation or the stochastic method used. These are described below and summarised in table 1. Most of the stochastic methods have been implemented by introducing a stochastic element to the pre-existing UM parameterisations. We quantify the variability associated with a stochastic method using the spread of an ensemble, with a different set of random numbers drawn for the stochastic component of each ensemble member. Small initial condition perturbations are also included in the ensembles; these are discussed in section 3.2.

information about their derivation, at http://tornado.atmos.colostate.edu/togadata/data/ifa_data.html

Scheme name	Description	Type
Default UM	Includes Gregory & Rowntree mass flux convection scheme	Deterministic
Multiplicative Noise	The default UM with multiplicative noise introduced to total physics tendencies	Stochastic
Random Parameters (Varying)	The default UM with randomly time-varying model parameters	Stochastic
Random Parameters (constant)	The default UM with randomly selected but constant model parameters	Deterministic
Kain-Fritsch	The UM with the Kain-Fritsch mass flux convection scheme instead of the default	Deterministic
Plant & Craig (stochastic)	The UM with the Plant & Craig stochastic convection scheme	Stochastic
Plant & Craig (deterministic)	The UM with the Plant & Craig convection scheme in deterministic mode	Deterministic

Table 1: Summary of the SCM configurations used.

- *Default UM*
  The default UM configuration contains parameterisations for layer-cloud microphysics, radiation, boundary layer processes and convection. Martin *et al.* (2006) provide an overview of the current set of schemes. Convection is represented by a deterministic bulk mass flux scheme based on Gregory & Rowntree (1990), but which has since been modified (Martin *et al.* 2006). There are prognostic moisture variables for specific humidity, cloud liquid water content and cloud ice water content.

- *Kain-Fritsch convection scheme*
  An alternative deterministic mass flux scheme for convection is that of Kain & Fritsch (KF, 1990). The version described by Kain (2004) has been implemented here. For a discussion of the differences between the schemes of Gregory & Rowntree (1990) and KF in the UM in a forecasting context, see Done (2002).

- *Multiplicative Noise scheme*
  This scheme follows the method of Buizza *et al.* (1999) and is designed to represent parameterisation uncertainty. At each timestep, the total parameterised tendencies for each model variable are multiplied by a random number ϵ_1 chosen from a uniform distribution between $1 - k$ and $1 + k$, where the constant k sets the amplitude of the stochastic perturbations. The random number is the same for each model variable and at each vertical level. Temporal correlation is enforced by keeping the same random number for multiple timesteps. Buizza *et al.* (1999) found that the greatest improvements to the performance of the ECMWF ensemble prediction system occurred for $k = 0.5$ and a new random number every 6 hours. The same choices are made here. Total tendencies from the default UM are multiplied by ϵ_1 at the end of each timestep, with a check to restore moisture to zero if the stochastic perturbation implies a negative value.

- *Random Parameters scheme*
  GCM parameterisations include parameters for which the appropriate value is not

well determined. The Random Parameters scheme attempts to account for parameterisation uncertainty by allowing such parameters to vary within a plausible range. This scheme follows the system used (Arribas 2004) in the Met Office Global and Regional Ensemble Prediction System (MOGREPS, Mylne *et al.* 2005). The relevant parameters and ranges can be found in Arribas (2004), but include the entrainment rate and CAPE closure timescale from the UM convection scheme. Temporal correlations are described by a first-order auto-regression model,

$$P_{n+1} = \mu_P + r(P_n - \mu_P) + k_P \epsilon_2 \tag{1}$$

in which the parameter is labelled P and the update number n. μ_P is the default value of P, r is an auto-correlation coefficient and $k_P \epsilon_2$ is a stochastic shock term (see section 3.2 of Mylne *et al.* 2005) in which ϵ_2 is a random number uniformly distributed between -1 and 1, and k_P is a parameter-dependent normalisation. Each parameter is subject to maximum and minimum acceptable bounds, and the same random number ϵ_2 is used for all parameters at each update, every 3 hours.

- *Random but constant parameters*
  Another approach to sampling parameter uncertainty is to run an ensemble in which each run has a fixed, but different parameter set. This approach has been used to make probabilistic predictions of future climate (e.g. Murphy *et al.* 2004). For this study, we simply adapt the random parameters scheme above by choosing initial parameter values randomly within the acceptable range and holding these fixed. Our method does not explore parameter space in an un-biased way, as it is constrained by the correlations between parameters assumed in the random parameters scheme above. Nonetheless, it allows for an interesting test of the temporal correlations in that scheme.

- *Plant & Craig stochastic convection scheme*
  In the Plant & Craig (2008) parameterisation, a finite number of distinct plumes are present in a grid-box area at any instant, resulting in a random sampling of the full spectrum for an ensemble of cumulus clouds. The spectrum used is based on an equilibrium exponential distribution of cloud-base mass flux (Craig & Cohen 2006) and plumes are produced at random, with the properties of each based on an adaptation of the KF plume model. The smaller the grid-box size, the more limited the sampling and the larger the fluctuations from statistical equilibrium. A sounding averaged over nearby grid-points and recent timesteps provides a smoothed input for the CAPE closure calculation. Of course, spatial averaging is not possible in an SCM. Preliminary tests showed that the scheme behaved sensibly in the SCM when averaging over 20 timesteps. This choice represents a compromise between providing a smooth input profile and the need to capture variations in the dynamical forcings.

- *Deterministic limit of the Plant & Craig scheme*
  The Plant & Craig (2008) scheme can operate as a spectral convective parameterisation by running the plume model for every category of cloud and weighting the tendencies according to the probability of that cloud occuring. This corresponds to the deterministic limit of a very large grid-box in which the cumulus ensemble is well sampled.

3.2 Initial Condition Ensembles and Ensemble Size

It should be noted that a stochastic scheme is not required in order for a parameterisation embedded in an SCM to exhibit variability; even if the prescribed forcings remain constant, a purely deterministic SCM will vary from one timestep to the next. This is particularly associated with switches in parameterisations, the most important of which is the trigger function in the convection scheme. This can often exhibit exaggerated on-off behaviour (e.g. Willett and Milton 2006), and the exact set of timesteps on which the convection scheme triggers can be very sensitive to small changes in the model state (this was found to be true in deterministic SCM ensembles used in this study, not shown). Such unsteady behaviour inherent to convective and other parameterisations provides a source of variability in deterministic and stochastic SCMs alike.

Part of the variability in a stochastic physics SCM ensemble[2] may arise simply because the stochastic (ST) perturbations force each ensemble member to follow a different realisation, with the convection being triggered on a different set of timesteps. Such realisations can also be explored in a deterministic model by running an ensemble with initial condition (IC) perturbations. We suggest that such IC ensembles should be run in order to make meaningful comparisons of stochastic schemes with their deterministic counterparts. Hack and Pedretti (2000) suggest that an ensemble approach is appropriate for SCM studies as an SCM can be sensitive to small differences in the initial conditions.

The perturbations for an IC SCM ensemble should be small enough not to introduce significant bias to any of the ensemble members, but large enough to force the ensemble members to diverge into an unbiased sample of probable realisations early in the model run. It should be emphasised that the IC perturbations are only used in this study to provide a sample of realisations and are not necessarily intended to represent realistic IC uncertainty. Thus, there is no requirement for the perturbations to match instrumental and sampling errors in the observations that provide the ICs. Results of such ensemble tests are discussed in section 4.1.

Another aspect to consider is the ensemble size required. We use 39 member ensembles, which appears to be sufficient to produce useable results. The robustness of results derived from ensembles of this size can be estimated from a brief statistical consideration:

- *Ensemble means.* Assuming some model variable to be approximately normally distributed, its ensemble mean has a standard error of $\sigma/\sqrt{K}$, where σ is the standard deviation and K is the number of ensemble members. For example, in the IC ensemble for the default UM, the ensemble standard deviation for temperature is of the order 0.5 K. This gives the ensemble mean a standard error of approximately 0.08 K, leading to a 95% confidence interval of ± 0.16 K, which is less than 10% of the amplitude of typical temperature variations during the model runs (see figure 2).

- *Ensemble spread.* This can be quantified by the root-mean-square deviation of ensemble members from their mean (the standard deviation σ). It can be shown that the sampling distribution of a standard deviation is approximately normal with standard error $\sigma/\sqrt{2K}$, for ensemble sizes $K \gtrsim 25$. In our case, with $K = 39$, this leads to a 95% confidence interval of $\pm 22\%$ of σ. An interval this broad suggests

[2]That is, an ensemble in which each member has the same stochastic parameterisation scheme but draws a different set of random numbers for it.

that ensemble spreads calculated for single variables should be interpreted with care. More accurate ensemble spreads occur for an error norm which sums the ensemble spread over the model column. We present below results for the Total Column root-mean-square Ensemble Spread, given by

$$\mathrm{TCES} = \left[\frac{\int_{p_{\mathrm{top}}}^{p_{\mathrm{surf}}} \sum_{k=1}^{K} \left(F_k - \overline{F} \right)^2 dp}{(p_{\mathrm{surf}} - p_{\mathrm{top}})(K-1)} \right]^{1/2} \tag{2}$$

where F is a model variable, $\overline{F}$ its ensemble mean, p is the pressure coordinate, and k labels ensemble members. Assuming hydrostatic balance, the TCES is the square root of the mass-weighted vertical integral of estimated model-field variance. Since there are vertical correlations between model levels in the SCM, it is not obvious how much more accurate this norm will be than the ensemble standard deviation at a single level. We have therefore performed some simple tests to estimate a sampling error for the TCES and found a 95% confidence interval of around ±10%.

4 Results

4.1 Ensemble Sensitivity to Initial Condition Perturbations

We have constructed IC ensembles from two sets of IC perturbations. In set 1, random temperature perturbations are applied to the lowest model level, chosen from a uniform distribution between ±0.25 K. Temperature perturbations for set 2 are larger and cover a greater vertical extent. A uniform distribution is again used with amplitude 0.5 K at the surface and decreasing exponentially above with a height scale of 1 km. For these larger perturbations, it is desirable to ensure that no spurious super-saturation occurs, and so corresponding perturbations are applied to the specific humidity field in order to maintain the relative humidity.

Figure 2 shows spaghetti plots of temperature for a single model level in the lower troposphere. Set 1 IC perturbations have been added to the default UM. Figure 3 is equivalent, but for the Kain-Fritsch convection scheme. For the default UM, the perturbations appear to produce a good spread of realisations, but using the KF scheme the ensemble members fail to diverge. Even at the end of the 19 day runs they remain clustered in six distinct groups, the members of each group triggering convection on the same set of timesteps (not shown). Figure 4 shows the temperature plume for the Kain-Fritsch scheme using the set 2 IC perturbations. It is clear that the ensemble members are less tightly clustered than with the set 1 perturbations, producing a more representative sample of realisations. But the spread is still smaller than in the default UM using set 1. It is perhaps slightly surprising that the SCM responds so differently to IC perturbations when different convection schemes are used. This is in contrast to Hume and Jakob (2005).

An important point to note from the simulations of Hack and Pedretti (2000) is their observation of bifurcations in SCM solutions (e.g. their figure 4), with ensemble members dividing into two or more preferred modes. Clearly, this raises issues with the representativeness of statistics such as the ensemble mean, since the mean state may lie between modes and never actually occur. Little evidence for multi-modal behaviour was found in the present study. Clearly separated modes do occasionally occur, as seen for

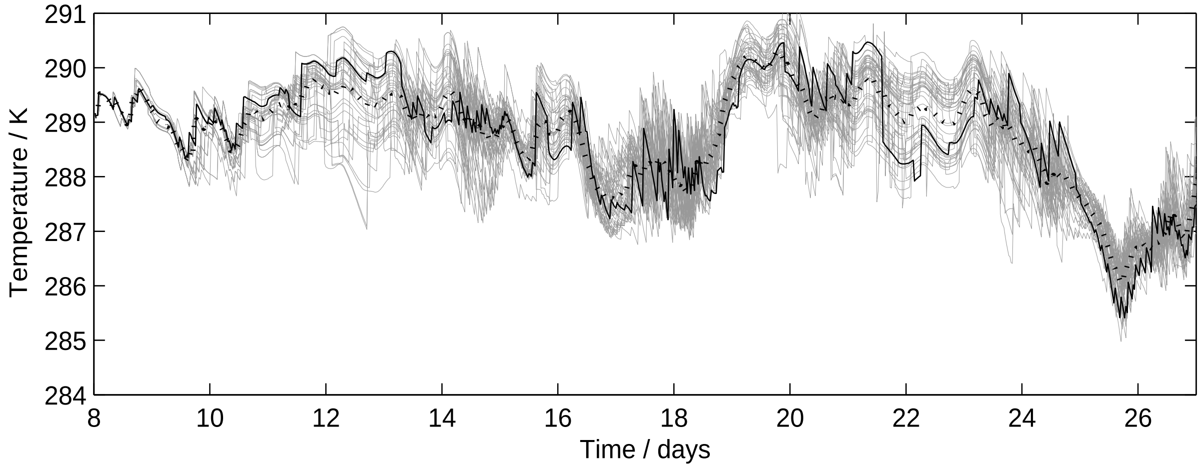

Figure 2: Ensemble plume plot of temperature on model level 10 (800 hPa) for the default UM with the set 1 IC perturbations. The solid line denotes the control run (with no IC perturbations) and the dotted line denotes the ensemble mean.

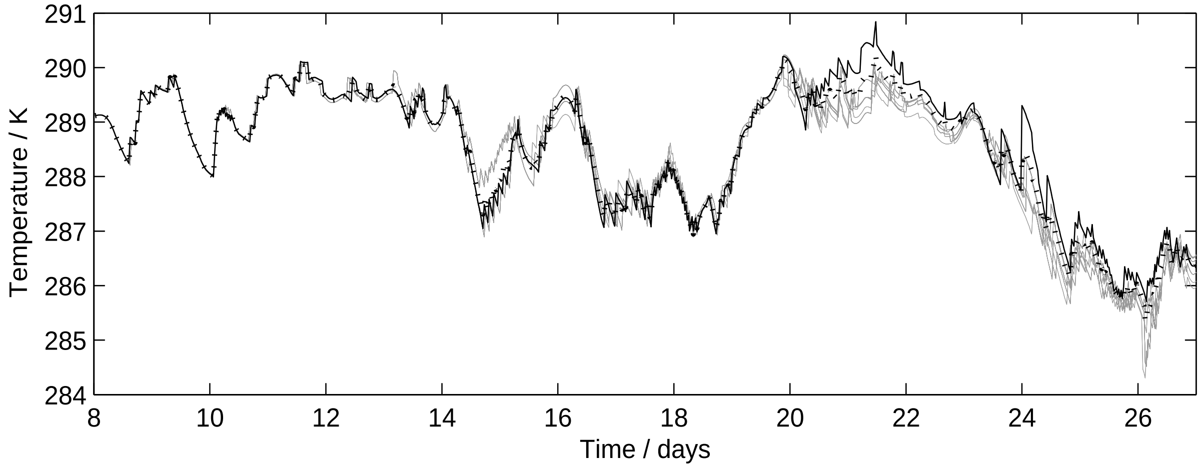

Figure 3: As figure 2 but with the Kain-Fritsch convection scheme.

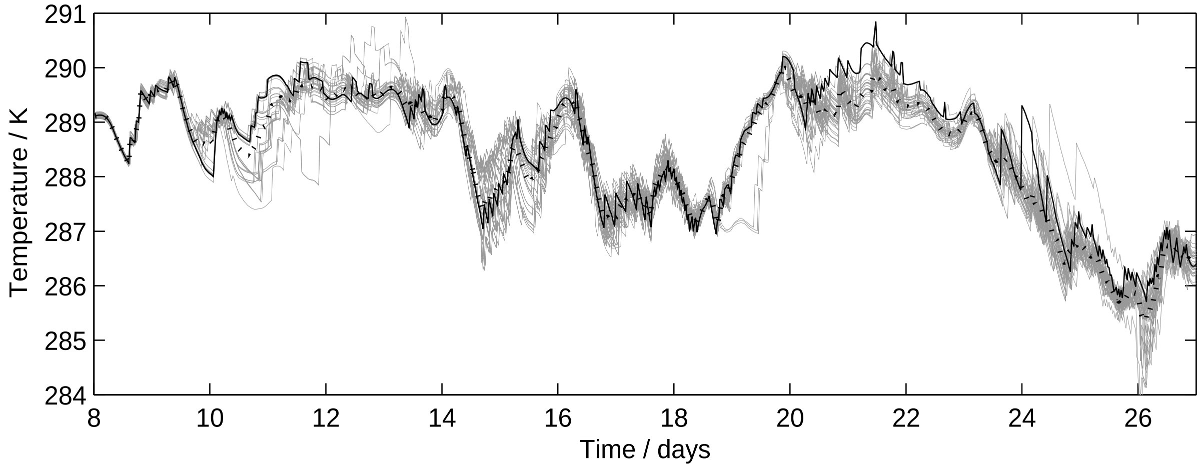

Figure 4: As figure 2 but with the Kain-Fritsch convection scheme and the set 2 IC perturbations.

example, using the Kain-Fritsch convection scheme around the 19th (figure 4). However, these persist for no longer than a day or so. The presence or absence of bifurcations is presumably related to the character of either (or both) the SCM or the large-scale forcing. We do not speculate further here, but rather note that the ensemble mean and standard deviation appear to be genuinely useful diagnostics for the present study.

Figure 5 shows timeseries of the TCES of temperature, for the Default UM ensemble with set 1 and set 2 IC perturbations, and also for an ensemble which includes the multiplicative noise scheme described in section 3.1, both with and without set 2 IC perturbations added. The corresponding plots for relative humidity are shown in Figure 6.

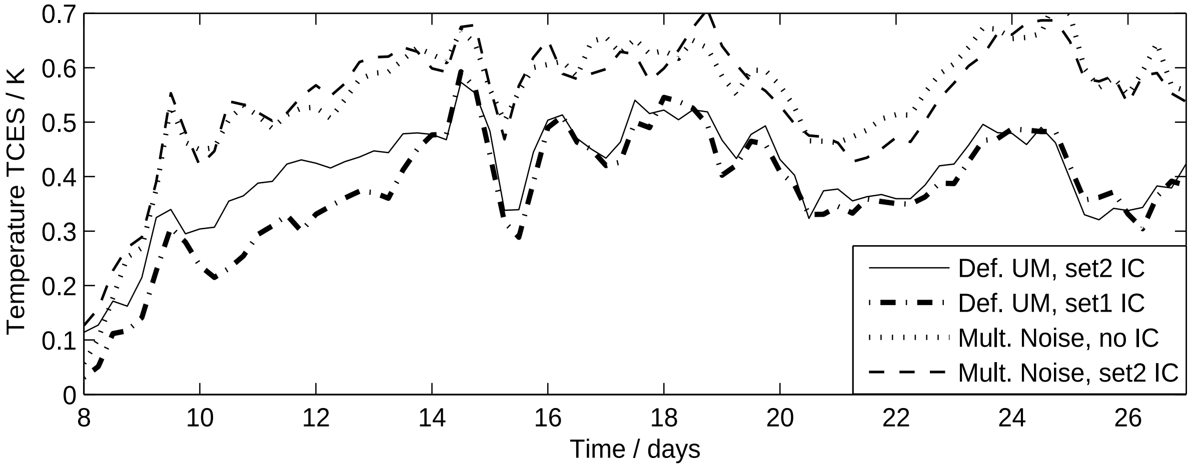

Figure 5: $TCES$ of temperature, in the default UM with set 1 (dash-dotted line) and set 2 (solid line) IC perturbations, with multiplicative-noise perturbations but no IC perturbations (dotted line), and with both set 2 IC perturbations and multiplicative noise (dashed line).

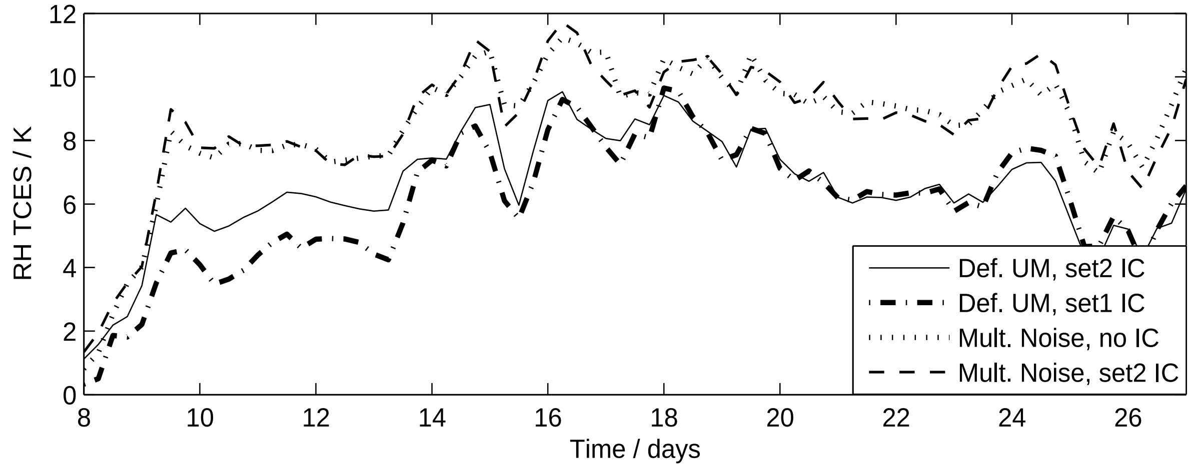

Figure 6: As figure 5, but for relative humidity rather than temperature.

Looking first at the two Default UM TCES ensembles, there is more spread over the first 6 days using the larger set 2 IC perturbations. However, the set 1 and 2 ensembles look very similar beyond 6 days, suggesting that the ensemble spread has saturated in both. This is reassuring as it suggests 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. For the Kain-Fritsch scheme, the larger IC perturbations produce larger ensemble spreads throughout (figures 3 and 4), but this is because the spread did not saturate when the set 1 IC perturbations were used.

In a stochastic physics (ST) SCM ensemble, the stochastic method provides some physically-motivated source of variability. One might anticipate that the physics perturbations would allow the ST ensemble to explore at least those realisations accessible to its deterministic analogue[3]. If this is true, then IC perturbations should have little effect on ensemble spread when implemented in an ST ensemble. It is clear from figures 5 and 6 that beyond the first 36 hours or so, the ensembles including multiplicative noise ST 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 probable realisations which 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 ST 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 ST-only ensembles, and that ensembles with IC and ST perturbations produced greater spread still[4]. Teixeira and Reynolds (2008) 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[5]. Also those studies used much larger IC perturbations designed to sample IC uncertainty.

It is interesting to note in Figures 5 and 6 the time-variability of the ensemble spreads. The spread clearly has some dependence on the large scale forcing, with a peak followed by a sudden drop in spread occurring at the start of each convectively active phase. To study this in more detail, we show in figure 7 a time–height plot of the ensemble spread in temperature in the set 2 Default UM ensemble. The spread appears to follow different characteristic regimes during suppressed and active phases, while behaving in a more unsteady manner during transition periods between the two. Note that in figure 7 the active and suppressed phases labelled in figure 1 have been redefined in order to separate out the transition periods. This will allow the ensemble variability characteristic of suppressed and active phases to be analysed in section 4.3.

Most notably during transitions from suppressed to active phases, the pattern of ensemble spread appears to be related to the convective cloud top height. For example, peaks in TCES on the 15th (figures 5 and 6) correspond to large spreads in the mid-troposphere where the ensemble produces a broad range of convective cloud tops. During the following day this range suddenly narrows and the ensemble spread drops throughout

[3]By 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.

[4]This is shown for forecast days 3, 5 and 7 in their table 1a.

[5]See Teixeira and Reynolds (2008), figure 7a for example.

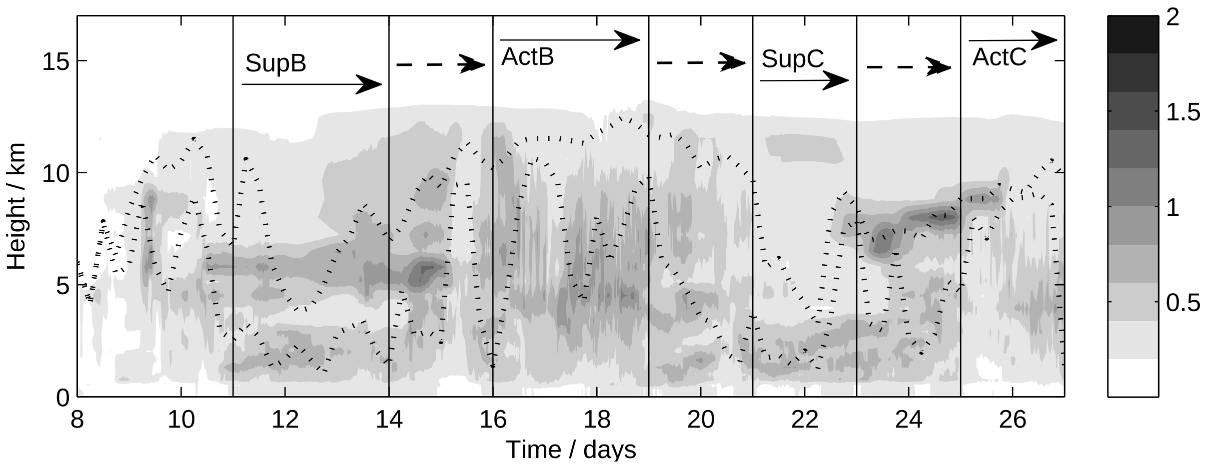

Figure 7: Ensemble r.m.s. spread in temperature (K) for the Default UM ensemble using set 2 IC perturbations (shaded). Overlaid are the 25th and 75th percentiles of the convective cloud top height (dotted lines). These values are determined at each timestep from the subset of ensemble members in which the convection scheme triggers and are smoothed for presentation here using 6-hourly averaging. Annotations are explained in the text.

the troposphere. Another interesting feature is the sloping layer of high ensemble spread around the 24th and 25th, which increases in height from roughly 7 km to 9 km during the transition from SupC to ActC. This layer closely follows the 75th percentile of convective cloud top height, indicating an ascending lid on the convection. 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 (not shown).

4.2 Intercomparison of Ensemble Mean States

An ensemble was produced for each of the SCM configurations described in section 3.1. In the case of the stochastic Plant & Craig scheme, two separate ensembles were produced for columns with horizontal scales Δx of 50 and 100 km (as explained in section 3.1, 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, we included the set 2 IC perturbations in all ensembles, although beyond the first day they make very little difference to the stochastic ones.

We show results here for the precipitable water content (PWC), the mass-weighted integral of specific humidity through the column. Figure 8 shows timeseries of observation-derived PWC and ensemble means for three deterministic SCM configurations. The SCMs exhibit large systematic biases from the observed PWC; this is probably due to discrepancies in the large scale forcings used to drive the SCMs, as found in other SCM and CRM studies which use advective forcings derived from observations (e.g. Krueger and Lazarus 1999). The Kain-Fritsch and Plant & Craig schemes (which are based on the same convective plume model) produce a drier state than the default UM, although well within the range of values seen when comparing various SCMs (Woolnough *et al.* submitted). The drying is associated with tropospheric cooling in these schemes relative to the default UM (not shown).

Several of the SCM configurations in this study include ST perturbations but are

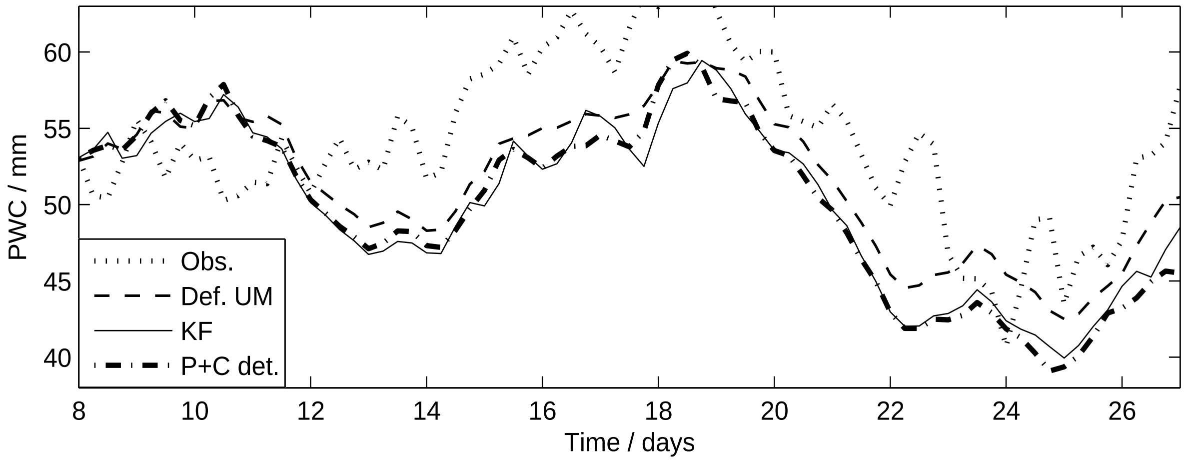

Figure 8: Six-hourly mean PWC derived from TOGA-COARE observations (dotted line), and ensemble means for the Default UM (dashed line), Kain-Fritsch (solid line), and Plant & Craig deterministic mode (dash-dotted line) configurations.

based on the UM convection scheme. In figure 9, 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 10 shows similar plots for the Plant & Craig scheme. In terms of ensemble mean PWC, the difference between the two convection parameterisations (default UM and Kain-Fritsch) is several times the difference between any of the stochastic schemes and its deterministic analogue. Similar remarks apply to other variables and suggest that the ensemble mean fields are more sensitive to structural differences in the convection scheme than they are to the introduction of stochastic schemes.

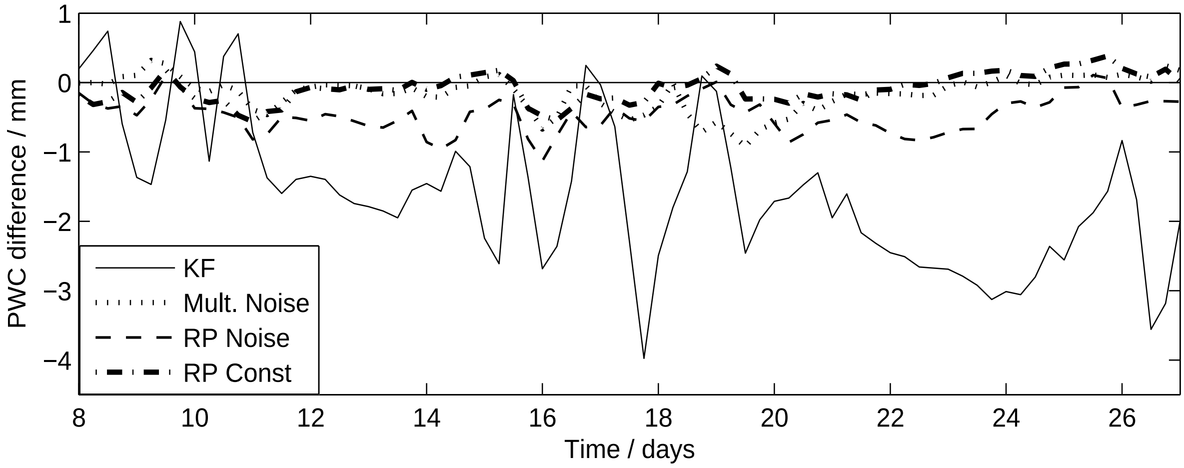

Figure 9: Six-hourly mean difference in ensemble mean PWC relative to the default UM, for the multiplicative noise (dotted line), time-varying random parameters (dashed line), constant random parameters (dash-dotted line), and Kain-Fritsch (solid line) configurations.

However, 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 interesting. Statistical tests indicate that the

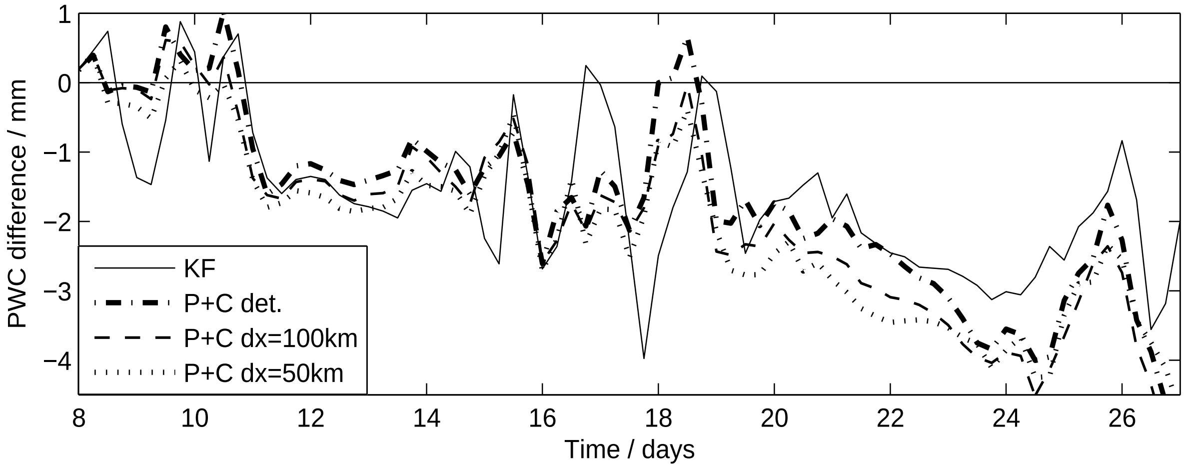

Figure 10: As figure 9, but for the Plant & Craig scheme with gridlengths of 50 km (dotted line) and 100 km (dashed line) and in deterministic mode (dash-dotted line), and for the Kain-Fritsch scheme (solid line).

ensemble mean state of the time-varying random parameters ensemble is significantly cooler and dryer than the default UM for much of the model run, especially during periods of suppressed convection. However the constant random parameters scheme did not produce this deviation despite sampling the same range of values for model parameters (see figure 9). This suggests noise induced drift; the random noise introduced by the time-variation of the model parameters causes the SCM to explore a region of phase-space which is asymmetric about the mean state of the deterministic analogue. Note in figure 10 that the stochastic Plant & Craig scheme produces a similar drift relative to its deterministic analogue (most clearly seen around the 22nd), which is also found to be statistically significant during the suppressed phases. This drift is smaller when the larger column size is used.

4.3 Intercomparison of Ensemble Variability

Figure 11 shows time-mean vertical profiles of ensemble spread in temperature for the active and suppressed periods ActB, SupC and ActC, as labelled in figure 7. There are marked differences between active and suppressed phases. This is most apparent in the mid-troposphere where the spread tends to be higher during active phases, whereas in the lower troposphere most of the SCM configurations exhibit greater spread during the suppressed phase (the Plant & Craig scheme is an exception during ActC). These observations are consistent with the notion that convective variability is a key ingredient in producing spread.

There are distinct differences between the profiles in panels (a), (b) and (c), which are for configurations using the default UM's convection scheme, and those in panels (d), (e) and (f), which are for configurations based on the Kain-Fritsch convective plume model. The latter grouping exhibits large peaks in ensemble spread 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. The vertical structure of the ensemble spread profile appears to be primarily dependent on the convective plume model used, with the ST perturbations primarily

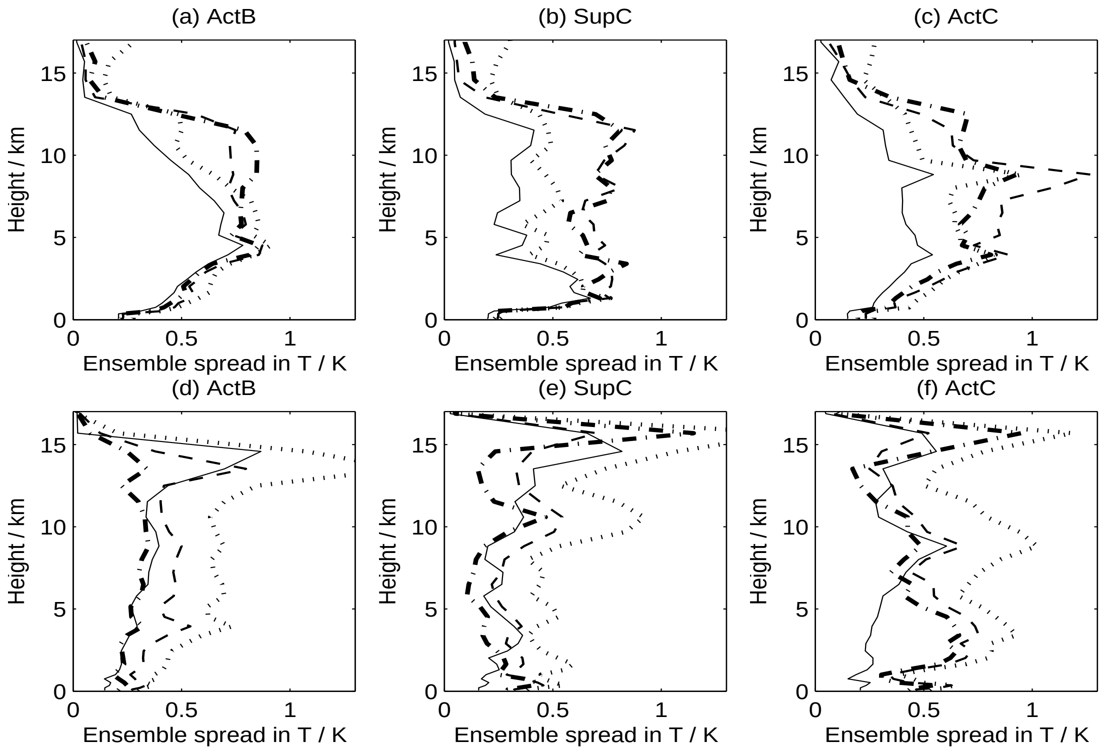

Figure 11: Vertical profiles of ensemble r.m.s. spread in temperature, for eight different SCM ensembles, averaged over the periods ActB (panels a and d), SupC (panels b and e), and ActC (panels c and f). The periods are as defined in figure 7. The profiles in panels a, b and c are for the Default UM configuration (solid line), the Multiplicative Noise scheme (dotted line), the time-varying Random Parameters scheme (dashed line), and the constant Random Parameters scheme (dot-dashed line). Those in panels d, e and f are for the Kain-Fritsch configuration (solid line), and the Plant & Craig scheme with a gridlength of 50 km (dotted line), 100 km (dashed line), and in deterministic mode (dot-dashed line).

affecting its amplitude.

In the troposphere, the default UM ensemble produces more spread than the Kain-Fritsch ensemble. These profiles confirm that the convection parameterisation is an important source of variability and also that different deterministic convection parameterisations produce rather different variabilities 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.

The schemes that 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 in the mid and upper troposphere, but have relatively little effect on the lower troposphere. The multiplicative noise scheme also affects the stratosphere, as it directly perturbs the radiative tendencies which dominate there. The stochastic Plant & Craig scheme also tends to scale-up the profile of spread produced by its deterministic mode, but differs from the other methods in that during ActB and SupC it creates substantial increases in spread in the lower troposphere.

The deterministic Plant & Craig scheme generally produces small ensemble spreads, often smaller than those in the Kain-Fritsch ensemble. This is consistent with its design, since it uses time-averaged profiles to reduce timestep-to-timestep variability in its closure calculations. The stochastic form of this scheme is not designed to represent generic model uncertainty, but specifically the variability arising from sub-sampling the cumulus ensemble within a finite area. For an area of side $\Delta x = 100$ km, the scheme is certainly more spread than in deterministic mode, but still comparable with the deterministic Kain-Fritsch ensemble and in the troposphere is much less spread than any of the model-uncertainty schemes. However, with $\Delta x = 50$ km the ensemble spread has tropospheric values comparable to those produced by model-uncertainty schemes. These results suggest that local fluctuations about convective equilibrium become as important as generic model uncertainty at resolutions of around 50 km, and a key mechanism for variability at smaller grid-lengths.

4.4 Comparison of Stochastic Physics SCM Ensemble Spread 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. To examine this point, it is useful to compare the ensembles that are designed to represent model uncertainty with the range of model states produced by different deterministic structural configurations. A poor-man's ensemble is produced by combining the 39-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 117-member ensemble can be used as a simple measure of the spread of model states associated with model uncertainty. With only three different model configurations, its representativeness is questionable, but we suggest that a stochastic scheme that aims to represent model uncertainty should produce at least comparable levels of spread. Figure 12 shows timeseries of several ensemble percentiles of PWC for each of the stochastic schemes and for the constant random parameters scheme, compared with the same percentiles of the combined deterministic ensemble. (The ensemble mean PWC for each deterministic scheme was shown in figure 8).

It is encouraging to find that the three schemes designed to represent generic model uncertainty do indeed produce spread comparable to the combined deterministic ensemble (see panels a, b and c). However, these schemes tend simply to broaden the ensemble about the ensemble mean state 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. As discussed in section 4.2, it produces some noise-induced drift, and this appears to be favourable, in the sense that the distribution of PWC is nudged towards that of the combined deterministic ensemble. It only fails to encompass the full range of model uncertainty during the last day or two of the runs.

Panels d and e show results for the stochastic Plant & Craig scheme. These confirm the points made in section 4.3, that fluctuations about convective equilibrium result in ensemble spread similar to key aspects of model uncertainty on scales of around 50 km.

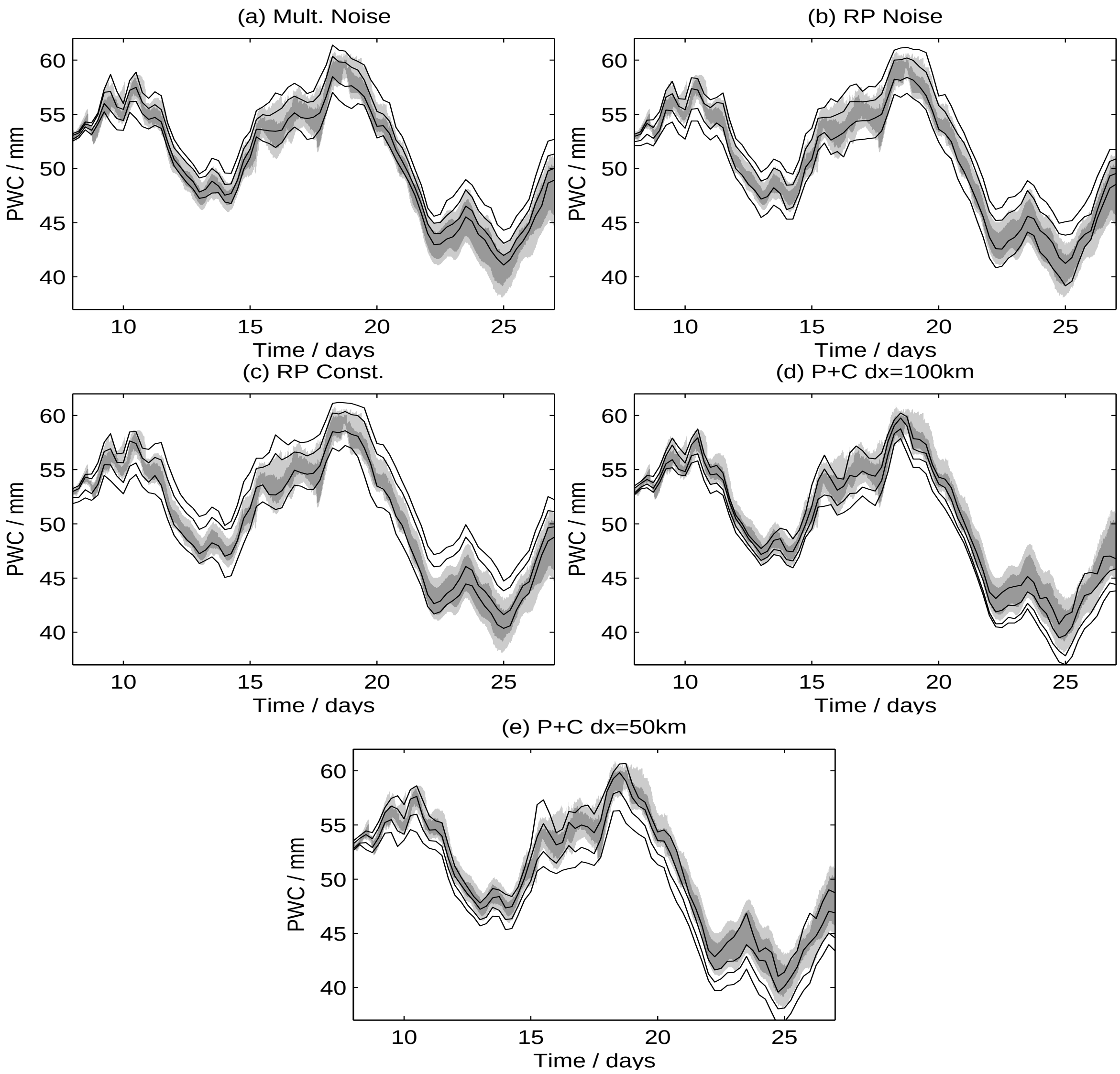

Figure 12: The 5th, 25th, 75th and 95th percentiles of PWC (solid lines) for ensembles using (a) the multiplicative noise scheme, (b) the time-varying Random Parameters scheme, (c) the constant Random Parameters scheme, and (d,e) the Plant & Craig scheme with a gridlength of (d) 100 km and (e) 50 km. Each panel also shows (shaded) the same ensemble percentiles for the combined deterministic ensemble.

5 Conclusions

Single column tests isolate the parameterisation schemes of a GCM and allow one to study their interactions under prescribed forcings. The strengths of the 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. We have performed single-column tests of tropical convection, comparing various stochastic physics methods. The interactions of stochastic perturbations with model dynamics are likely to be an important aspect of their behaviour in a full GCM, but we wished to consider whether SCM tests may nonetheless have value.

It is necessary to study stochastic methods with ensembles of SCM runs. Here we have used an ensemble size of 39 which is certainly practical, and sufficient to make some inferences about the methods, but additional members might have allowed more definitive statements to be made in some cases. The stochastic ensembles can be usefully compared with deterministic ensembles produced by initial condition uncertainty and also with combinations of these into poor man's ensembles. Such comparisons allow one to judge not merely that the ensemble mean from some stochastic method is sensible, but also to assess the variability that the method produces through its interactions with the GCM parameterisation set. For example, if the spread of a stochastic ensemble designed to represent model uncertainty were much larger (smaller) than the poor man's ensemble, then the SCM would imply that for a good performance of the method in a full GCM the interactions of the stochastic perturbations with GCM dynamics should be such as to strongly dampen (amplify) the variability introduced.

In agreement with Hack and Pedretti (2000), deterministic SCM runs were found to be sensitive to small initial condition perturbations. The variability of initial condition ensembles is strongly dependent on the convection parameterisation used, according to the timings and frequency of triggering. The perturbations chosen allowed the runs to diverge into a set of independent realisations within a few days. The same initial condition perturbations had very little effect on the ensembles produced by stochastic methods, beyond the first day or two.

Three methods designed to represent model uncertainty appeared to perform well, the ensemble spreads being broadly similar to that of the deterministic poor man's ensemble. The ensemble mean states were close to the ensemble means of the deterministic analogue (differences in the convective parameterisation produced substantially larger changes to the mean). For the random parameters scheme, however, there was a statistically significant noise-induced drift of ensemble mean PWC. The Plant & Craig scheme produced levels of spread similar to the model-uncertainty approaches for $\Delta x = 50$ km, suggesting that fluctuations about convective equilibrium form an important component of variability at and below this scale.

Although there are some changes in methodology and philosophy for SCM testing of stochastic methods, we are inclined to view our results as encouraging and to speculate that SCM testing may have a useful role to play in studying stochastic parameterisations. This may be particularly the case for studying possible variations of a method, and determining whether further, more comprehensive testing is desirable. For example, it seems clear from our results that a GCM integration with a horizontal resolution of 2.5° would not be a good test of the impact of the convective fluctuations parameterised in the Plant & Craig scheme. Given that comprehensive testing of all details of all plausible stochastic methods will remain impractical, we contend that the indications of potential impact that may be gleaned from SCM tests are far preferable to no indications at all.

Acknowledgements

We are grateful to A. Arribas, N. Bowler and K. Mylne for their discussions of stochastic methods, to S. Woolnough for discussions of GCSS case 5 and for providing the observational data, to R. Wong for providing us the default UM SCM configuration for that case, and to NCAS Computer Modelling Support. M. A. Ball is funded by the NERC award NER/S/A/2006/14189 with CASE support from the Met. Office.

References

- Arribas, A. 2004 Results of an initial stochastic physics scheme for the Met Office Unified Model. Forecasting Research Technical Report no. 452, Met Office, UK.

- Bergman, J. W. & Sardeshmukh, P. D. 2004 Dynamic stabilisation of atmospheric single column models, *J. Clim.*, **17**, 1004-1021.
(doi:10.1175/1520-0442(2004)017<1004:DSOASC>2.0.CO;2)

- 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.
(doi:10.1256/smsqj.56005)

- Byun, Y.-H. & Hong, S.-Y. 2007 Improvements in the subgrid-scale representation of moist convection in a cumulus parameterisation scheme: The single-column test and its impact on seasonal prediction. *Mon. Weather. Rev.* **135**, 2135–2154.
(doi:10.1175/MWR3397.1)

- Craig, G. C. & Cohen, B. G. 2006 Fluctuations in an equilibrium convective ensemble. Part I: theoretical formulation. *J. Atmos. Sci.* **63**, 1996–2004.
(doi:10.1175/JAS3709.1)

- Craig, G., Fraedrich, K., Lin, J., Moncrieff, M., Plant, R., Thomson, D. & and Tompkins, A. 2005 Working group 1 report: issues in convection. In: *ECMWF Workshop on Representation of Sub-grid Processes using Stochastic–Dynamic Models, Reading, UK, 6-8 June 2005*, pp. ix–xii.

- Cullen, M. J. P. 1993 The unified forecast/climate model. *Meteorol. Mag.* **122**, 89–94.

- Done, J. M. 2002 Predictability and representation of convection in a mesoscale model, p, 180. PhD thesis, University of Reading, UK.

- Ghil, M., Berloff, P., Buizza, R., Hermanson, L., Miller, M., Palmer, T., Roulstone, I., Sardeshmukh, P. & Williams, P. 2005 Working group 3 report: the parameterisation problem in weather and climate models In: *ECMWF Workshop on Representation of Sub-grid Processes using Stochastic–Dynamic Models, Reading, UK, 6-8 June 2005*, pp. xvii–xxii.

- Gregory, D. & Rowntree, P. R. R. 1990 A mass flux convection scheme with representation of cloud ensemble characteristics and stability dependent closure. *Mon. Weather. Rev.* **118**, 1483–1506.
(doi:10.1175/1520-0493(1990)118<1483:AMFCSW>2.0.CO;2)

- Hack, J. J. & Pedretti, J. A. 2000 Assessment of solution uncertainties in single-column modeling frameworks. *J. Clim.* **13**, 352–365.
(doi:10.1175/1520-0442(2000)013<0352:AOSUIS>2.0.CO;2)

- Horinouchi, T., Pawson, S., Shibata, K., Langematz, U., Manzini, E., Giorgetta, M. A., Sassi, F., Wilson, A. J., Hamilton, K., de Grandpré, J. & Scaife, A. A. 2003

Tropical cumulus convection and upward-propagating waves in middle-atmospheric GCMs. *J. Atmos. Sci.* **60**, 2765–2782.
(doi:10.1175/1520-0469(2003)060<2765:TCCAUW>2.0.CO;2)

- 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. (doi:10.1029/2004JD005704)

- Kain, J. S. 2004 The Kain-Fritsch convective parameterization: an update. *J. Appl. Meteorol.* **43**, 170–181.
(doi:10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2)

- Kain, J. S. & Fritsch, J. M. 1990 A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.* **47**, 2784–2802.
(doi:10.1175/1520-0469(1990)047<2784:AODEPM>2.0.CO;2)

- Krueger, S. K., and Lazarus, S. M. 1999 Intercomparison of multi-day simulations of convection during TOGA COARE with several cloud-resolving and single-column models. In *Twenty-third Conf. Hurricanes and Tropical Meteorology, Dallas, TX.* Preprint, pp. 643-647. Boston, MA: American Meteorological Society.

- Lin, J. W.-B. & Neelin, J. D. 2002 Considerations for stochastic convective parameterization. *J. Atmos. Sci.* **59**, 959–975.
(doi:10.1175/1520-0469(2002)059<0959:CFSCP>2.0.CO;2)

- 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. Clim.* **19**, 1274–1301.
(doi:10.1175/JCLI3636.1)

- 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. Am. Meteorol. Soc.* **78**, 831–845.
(doi:10.1175/1520-0477(1997)078<0831:GCSSGW>2.0.CO;2)

- Murphy, J., Sexton, D. M. H., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M. & Staniforth, D. A. 2004 Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature* **430**, 768–772.
(doi:10.1038/nature02771)

- Mylne, K. R., Evans, R. E. & Clark, R. T. 2002 Multi-model multi-analysis ensembles in quasi-operational medium-range forecasting. *Q. J. R. Meteorol. Soc.* **128**, 361–384.
(doi:10.1256/00359000260498923)

- Mylne, K., Bowler, N., Arribas, A. & John, S. 2005: MOGREPS (Met Office Global and Regional Ensemble Prediction System). Met Office Scientific Advisory Committee, 10 to 11 November 2005, Paper No. 10.5.

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

- Penland, C. 2003 A stochastic approach to nonlinear dynamics: a review *Bull. Am. Meteorol. Soc.* **84**, ES43–ES52.
(doi:10.1175/BAMS-84-7-Penland)

- 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.
(doi:10.1002/qj.109)

- Plant, R. S. & Craig, G. C. 2008 A stochastic parameterization for deep convection based on equilibrium statistics. *J. Atmos. Sci.* **65**, 87–105.
(doi:10.1175/2007JAS2263.1)

- 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.
(doi:10.1175/1520-0469(2000)057<3461:TEOEWB>2.0.CO;2)

- Shutts, G. 2005 A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Q. J. R. Meteorol. Soc.* **131**, 3079–3102.
(doi:10.1256/qj.04.106)

- Sobel, A. H., Bellon, G. & Bacmeister, J. 2007 Multiple equilibria in a single-column model of the tropical atmosphere. *Geophys. Res. Lett.* texbf34, L22804.
(doi:10.1029/2007GL031320)

- Teixeira, J. & Reynolds, C. A. 2008 The stochastic nature of physical parameterisations in ensemble prediction: a stochastic convection approach *Mon. Weather. Rev.* **136**, 483-496.
(doi:10.1175/2007MWR1870.1)

- Tompkins, A. M. & Berner, J. Submitted. Stochastic convective input based on subgrid humidity distributions *J. Geophys. Res.*

- Webster, P. J. & Lukas, R. 1992: TOGA COARE: the coupled ocean-atmosphere response experiment. *Bull. Am. Meteorol. Soc.* **73**, 1377–1416.
(doi:10.1175/1520-0477(1992)073<1377:TCTCOR>2.0.CO;2)

- Wilks, D. S. 2005 Effects of stochastic parameterisations in the Lorenz '96 system. *Q. J. R. Meteorol. Soc.* **131**, 389–407.
(doi:10.1256/qj.04.03)

- Willett, M. R. & Milton, S. F. 2006 The tropical behaviour of the convective parameterisation in aquaplanet simulations and the sensitivity to timestep. Forecasting Research Technical Report no. 482, Met Office, UK.

- Williams, P. D. 2005 Modelling climate change: the role of unresolved processes. *Phil. Trans. R. Soc. A* **363**, 2931–2946. (doi:10.1098/rsta.2005.1676)

- Woolnough, S. J., Bechtold, P., Blossey, P., Chaboureau, J.-P., Hosomi, T., Iacobellis, S., Luo, Y., Petch, J. C., Wong, R. & Xie, S. Submitted. Convective processes during the suppressed phase of the Madden-Julian oscillation *Q. J. R. Meteorol. Soc.*