The impact of stochastic physics on tropical rainfall variability in global climate models on daily to weekly time scales

Peter A. G. Watson, Judith Berner, Susanna Corti, Paolo Davini, Jost von Hardenberg, Claudio Sanchez, Antje Weisheimer, Tim N. Palmer

AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2016JD026386

The impact of stochastic physics on tropical rainfall variability in global climate models on daily to weekly time scales

Peter A. G. Watson[1], Judith Berner[2], Susanna Corti[3], Paolo Davini[4], Jost von Hardenberg[5], Claudio Sanchez[6], Antje Weisheimer[1,7,8], and Tim N. Palmer[1]

[1]Atmospheric, Oceanic and Planetary Physics, University of Oxford, Oxford, UK, [2]National Center for Atmospheric Research, Boulder, Colorado, USA, [3]Institute of Atmospheric Sciences and Climate (ISAC-CNR), Bologna, Italy, [4]Laboratoire de Météorologie Dynamique/IPSL, Ecole Normale Supérieure, Paris, France, [5]Institute of Atmospheric Sciences and Climate (ISAC-CNR), Torino, Italy, [6]Met Office, Exeter, UK, [7]European Centre for Medium-Range Weather Forecasts, Reading, UK, [8]National Centre for Atmospheric Sciences, University of Oxford, Oxford, UK

Abstract Many global atmospheric models have too little precipitation variability in the tropics on daily to weekly time scales and also a poor representation of tropical precipitation extremes associated with intense convection. Stochastic parameterizations have the potential to mitigate this problem by representing unpredictable subgrid variability that is left out of deterministic models. We evaluate the impact on the statistics of tropical rainfall of two stochastic schemes: the stochastically perturbed parameterization tendency scheme (SPPT) and stochastic kinetic energy backscatter scheme (SKEBS), in three climate models: EC-Earth, the Met Office Unified Model, and the Community Atmosphere Model, version 4. The schemes generally improve the statistics of simulated tropical rainfall variability, particularly by increasing the frequency of heavy rainfall events, reducing its persistence and increasing the high-frequency component of its variability. There is a large range in the size of the impact between models, with EC-Earth showing the largest improvements. The improvements are greater than those obtained by increasing horizontal resolution to ~20 km. Stochastic physics also strongly affects projections of future changes in the frequency of extreme tropical rainfall in EC-Earth. This indicates that small-scale variability that is unresolved and unpredictable in these models has an important role in determining tropical climate variability statistics. Using these schemes, and improved schemes currently under development, is therefore likely to be important for producing good simulations of tropical variability and extremes in the present day and future.

Plain Language Summary Simulations from climate models have been found to lack day-to-day variability in tropical rainfall, with there being too many rainy days and not enough days with very heavy rainfall. A possible contributor to this problem is that the schemes the models use to predict rainfall try to predict the average rainfall that would be expected for given large-scale conditions. In reality, unpredictable small-scale features like eddies and gravity waves may contribute to the formation of severe storms or prevent them from developing. We test whether using stochastic methods to represent the effectively random impact of these small-scale features improves the variability of tropical rainfall simulated by three climate models. We find evidence that it does, and this indicates that treating the prediction of tropical rainfall probabilistically rather than deterministically will give improvements in climate simulations.

Key Points:
- Stochastic physics schemes improve simulated tropical precipitation variability on daily to weekly time scales in climate models
- Large improvements are found in simulating extreme event frequency, persistence, and power spectra of precipitation
- Small-scale variability has an important role in determining tropical climate variability statistics on scales larger than ~100 km

Correspondence to:
P. A. G. Watson,
peter.watson@physics.ox.ac.uk

Citation:
Watson, P. A. G., J. Berner, S. Corti, P. Davini, J. von Hardenberg, C. Sanchez, A. Weisheimer, and T. N. Palmer (2017), The impact of stochastic physics on tropical rainfall variability in global climate models on daily to weekly time scales, *J. Geophys. Res. Atmos.*, *122*, 5738–5762, doi:10.1002/2016JD026386.

Received 16 DEC 2016
Accepted 13 MAY 2017
Accepted article online 24 MAY 2017
Published online 8 JUN 2017

1. Introduction

Variability of tropical rainfall on daily to weekly time scales is a highly important aspect of climate. The occurrence of heavy rain events in the tropics can reduce crop yields, cause flooding, and increase the incidence of disease, and changes in tropical variability are expected to contribute to the damage of climate change [*Intergovernmental Panel on Climate Change*, 2014]. Simulating tropical rainfall well is crucial for activities like estimating flood risks [*Ward et al.*, 2015]. It is therefore important that climate models have a realistic representation of the statistics of this variability on these time scales on relevant length scales, such as those of river drainage basins.

Numerous studies have found that many climate models do not simulate tropical variability well on daily to weekly time scales, as discussed in the review by *Westra et al.* [2014]. For example, *Stephens et al.* [2010] showed that tropical locations had too many rainy days in a selection of models, and the rainfall was therefore typically too light. *Lin et al.* [2006] presented space-time spectra of tropical rainfall in Coupled Model Intercomparison Project Phase 3 (CMIP3) models and showed that generally they have too little spectral power associated with rainfall events lasting a few days or less, and *Hung et al.* [2013] found that there was not much improvement in CMIP5 models. *Crétat et al.* [2014] showed that most CMIP5 models have much less frequent intense daily mean rainfall over Africa than in satellite observations. Some models do produce intense rain in individual time steps, however [*Klingaman et al.*, 2017].

These model errors can affect other aspects of the simulations. Tropical convective systems can affect weather and climate at higher latitudes by acting as sources of Rossby waves [*Branstator*, 2014]. Additionally, inaccuracy in simulating tropical rainfall variability causes errors in simulating the amount of moisture that leaves the land surface as runoff and as evaporation [*Qian et al.*, 2006]. This affects coupling of the land surface and atmosphere, with *Saeed et al.* [2013] finding that the simulated tropical variability can affect the sign of projections of changes in mean rainfall due to climate change.

In the tropics, parameterizations of unresolved processes like deep convection have a very important role in climate models. One reason for the problems in simulating tropical rainfall variability in climate models may be that in the parameterizations, it is assumed that the impacts of small-scale phenomena like tropical convective systems on resolved scales can be represented by deterministic functions of variables averaged over a model grid box. In reality, a given grid box-average state may be associated with many possible realizations of the small-scale phenomena [e.g., *Davies et al.*, 2013; *Peters et al.*, 2013]. For example, two grid box-sized regions in a similar convective-equilibrium state could have different numbers of convective clouds [*Cohen and Craig*, 2006]. Therefore, more realistic statistics of the effects of subgrid phenomena can be obtained by modeling them as random samples from probability distributions conditioned on the grid box-scale state, rather than just the mean effect across all possible realizations, so that the effects of these different individual realizations are included in the simulations [*Palmer*, 2001, 2012]. This would be expected to increase the variability in the simulation of parameterized processes, potentially giving more realistic intermittency and extremes of quantities like rainfall on resolved scales. Also, due to nonlinearities in the dynamics, improving the representation of subgrid variability may also improve a model's mean state through feedbacks [e.g., *Sardeshmukh et al.*, 2003; *Williams*, 2012; *Weisheimer et al.*, 2014]. This may indirectly lead to improvements in the simulated variability, which can depend on the mean state.

In recent years, much effort has been put into developing stochastic parameterizations. Such parameterizations have been used in several numerical weather prediction models, including some run operationally, and have been found to improve forecast skill by both increasing the spread of ensemble forecasts and reducing the size of errors of the ensemble mean forecasts [e.g., *Buizza et al.*, 1999; *Palmer et al.*, 2009; *Reynolds et al.*, 2011; *Yonehara and M. Ujiie*, 2011; *Bouttier et al.*, 2012; *Sušelj et al.*, 2014; *Berner et al.*, 2015; *Sanchez et al.*, 2016]. There are also numerous other promising schemes that have been or are being developed [e.g., *Plant and Craig*, 2008; *Khouider et al.*, 2010; *Bengtsson et al.*, 2013; *Rochetin et al.*, 2014; *Kober et al.*, 2015; *Shutts*, 2015; *Dorrestijn et al.*, 2016; *Sakradzija et al.*, 2016; *Ollinaho et al.*, 2017; *Peters et al.*, 2017].

Several recent studies have found that adding stochasticity to atmospheric parameterization schemes improves aspects of the climate simulated in models, indicating that the stochasticity is indeed improving the representation of unresolved processes, as reviewed by *Berner et al.* [2017]. This includes improvements to simulated tropical climate. For example, *Berner et al.* [2008] and *Weisheimer et al.* [2014] showed that stochastic parameterizations applied in the European Centre for Medium-Range Weather Forecasts (ECMWF) coupled atmosphere-ocean model reduce biases in tropical mean rainfall, and *Sanchez et al.* [2016] found a similar result in the Met Office atmospheric model. *Lin and Neelin* [2000, 2002] showed that including stochasticity in the convection parameterization of intermediate complexity general circulation models (GCMs) improves aspects of the tropical variability, and *Lin and Neelin* [2003] also showed this in the National Center for Atmospheric Research (NCAR) Community Climate Model. *Davini et al.* [2017] showed that stochastic parameterizations in EC-Earth improve the simulation of tropical rainfall rate distributions and the Madden-Julian Oscillation, and *Wang et al.* [2016] found that the scheme of *Plant and Craig* [2008] improves the simulated tropical rainfall rate distribution in the NCAR Community Atmosphere Model as well. *Dorrestijn et al.* [2016], *Goswami et al.* [2016], and *Peters et al.* [2017] showed that variants of the stochastic multicloud model of

Khouider et al. [2010] improved aspects of tropical variability simulated in different GCMs, and *Frenkel et al.* [2012] showed similar results in a single-column model context. *Christensen et al.* [2017] also found that stochastic physics greatly reduced excessive El Niño variability in the NCAR Community Atmosphere Model, version 4 (CAM4).

The stochastic physics schemes evaluated here are variants of the stochastically perturbed parameterization tendencies scheme (SPPT) [*Buizza et al.*, 1999; *Palmer et al.*, 2009] and the stochastic kinetic energy backscatter scheme (SKEBS) [*Shutts*, 2005; *Berner et al.*, 2009]. We focus on these schemes because they have been tested extensively at ECMWF and the Met Office and have been shown to improve weather forecast skill [*Palmer et al.*, 2009; *Sanchez et al.*, 2016], and therefore, they are good candidates for being implemented into climate models at this time.

SPPT treats the total parameterized tendencies in prognostic model variables, such as temperature, humidity, and wind that are produced by the subgrid parameterizations as uncertain quantities. The tendencies are multiplied by a random number that scales them up or down. The random number is typically correlated in space and time, to account for the fact that parameterization errors are spatially and temporally correlated. Rainfall is not directly perturbed, but perturbations to atmospheric tendencies will affect rainfall at subsequent time steps. For example, if convective heating is increased, this will result in greater stabilization of the atmospheric column, possibly reducing precipitation at the next time step.

SKEBS provides a representation of errors resulting from energy dissipation by subgrid-scale processes affecting larger scales—a process that is not parameterized in deterministic atmospheric models. The atmospheric stream function tendency is perturbed at each grid point and time step by a random pattern with a specified amplitude, spatial power spectrum, and temporal autocorrelation. The temperature may also be perturbed, depending on the implementation. The amplitude of the perturbations may depend on dissipative processes acting at each location associated with the model numerics and subgrid-scale processes such as convection. These perturbations to the model state will influence the behavior of subgrid parameterization schemes and hence affect precipitation.

Note that SPPT and SKEBS represent only part of the uncertainty in parameterized processes—for example, uncertainty in whether convection is present in a grid box is not represented.

One simple conceptual model for how SPPT and SKEBS perturbations affect tropical precipitation variability, hereafter referred to as the "No Feedbacks" model, is that it is purely a result of the direct effect of the perturbations on a GCM's prognostic variables, from which precipitation is diagnosed. More precisely, let $\mathbf{X}$ be the prognostic variables in the deterministic configuration of the GCM, such as the grid-scale temperatures and winds, and let $P = P(\mathbf{X})$ be the precipitation diagnosed by the parameterization schemes. Let $\mathbf{X}'$ be the prognostic variables in a stochastic configuration, with stochastic schemes adding perturbations ϵ to this in a single model time step before precipitation $P' = P(\mathbf{X}' + \epsilon)$ is diagnosed. Assume that the stochastic perturbations do not affect the climatological statistical properties of the prognostic variables, so that the properties of $\mathbf{X}'$ equal those of $\mathbf{X}$. Then the differences between the statistics of P' and P are only due directly to the statistics of ϵ. *Subramanian et al.* [2017] demonstrated that this would typically cause the frequency of the heaviest and lightest rain rates to increase and that of rates in between to decrease. It would also be expected to reduce the temporal autocorrelation of precipitation, if the perturbations have a smaller autocorrelation than $\mathbf{X}$, and increase the precipitation's standard deviation. Deviations from these predictions indicate that feedbacks from changes in the statistics of resolved variables are important. Consistency with the predictions does not imply that feedbacks from resolved variables are necessarily unimportant, however. Note that the mean precipitation could potentially change even if the mean of ϵ is zero without there being feedbacks, since P is a nonlinear function of $\mathbf{X}$.

In this paper we extend upon previous work investigating the impact of SPPT and SKEBS on models' climate statistics, focusing on tropical rainfall variability on daily to weekly time scales in three different atmospheric models. We evaluate whether stochastic physics causes any improvements, which helps to indicate how realistically SPPT and SKEBS model the statistics of parameterization errors. Our analysis shows how important it is to refine the way stochasticity is included in parameterizations for simulating tropical variability well, on top of improving the deterministic aspects of schemes. Our results also indicate which diagnostics are sensitive to using stochastic physics and are therefore good targets to use when evaluating new stochastic schemes—statistics of tropical rainfall that have previously been used for this purpose have been shown not

to be sensitive to the way stochastic physics is implemented [*Watson et al.*, 2015]. We also compare the impact of increasing model resolution with using stochastic physics in two models, to test if stochastic physics can give similar or greater benefits, as investigated previously in the context of midlatitude dynamics [*Dawson and Palmer*, 2015]. We evaluate to what extent the results are consistent with the No Feedbacks conceptual model as well.

While previous studies on the impact of stochastic physics on model climate have examined the impact in individual models, the use of multiple models in our work allows evaluation of the sensitivity of the impact to the deterministic model being used and the details of the stochastic schemes. This may help to inform model developers about what impacts SPPT and SKEBS implemented in a new model can be expected to have and which impacts are sensitive to the details of the deterministic model and stochastic scheme.

2. Data Sets and Methods

We use precipitation data from two observational and three atmospheric model data sets, taking approximately 10 years of data from each. We use data between 1998 and 2007 for the observational data sets, which is the closest available decade to the periods covered in our model data sets. For the models, we use approximately 10 years of data that are the closest to this period, as specified in section 2.2. These periods are not the same for the different models. We have verified that our results are similar in separate subsamples of these periods, indicating that this has not had a large effect on the results. This included testing sensitivity to using just the 1993–1996 period in the Met Office Unified Model (UM) and CAM4 data to exclude any impact of the El Niño events that occurred shortly before and after this period.

The precipitation is conservatively interpolated to a 2.5° × 2.5° grid. We use daily mean precipitation for diagnostics of precipitation variability. We have verified that our results do not change qualitatively if data interpolated to a 1° × 1° grid is used.

2.1. Observational Data

In order to estimate biases in annual mean precipitation in the atmospheric models in section 3.1, we use monthly mean observational estimates from the Version-2.2 Global Precipitation Climatology Project Monthly Precipitation Analysis product [*Huffman et al.*, 1997; *Adler et al.*, 2003, hereafter "GPCP-Mon"]. This combines satellite-based estimates of rainfall from microwave imagers, cloud top temperatures, and rain gauges. For each model, we estimate the biases using the same 10 year period of these observations as we used for the given model (see section 2.2).

For diagnostics of precipitation variability, we use estimates of daily mean rainfall from the Global Precipitation Climatology Project One-Degree Daily (1DD) product [*Huffman et al.*, 2001, hereafter "GPCP"] and the Tropical Rainfall Measuring Mission Multisatellite Precipitation Analysis 3B42 version 7 [*Huffman et al.*, 2007, hereafter "TRMM"]. Both products are derived by combining satellite-based estimates of rainfall from microwave imagers and cloud top temperatures and scaled so their monthly means match those of GPCP-Mon.

We have found that while the daily precipitation values in the GPCP and TRMM data sets are very well correlated, the estimated magnitude of the precipitation can differ significantly, with TRMM generally indicating a larger rainfall amount in heavy rainfall events, particularly over oceanic regions. This is discussed in more detail in Appendix A. We do not attempt to determine which data set is more realistic here and treat the difference between the data sets as an estimate of observational uncertainty. We discuss alongside the results (section 3) where this observational uncertainty is too large for it to be clear what the model biases are and whether stochastic physics is making an improvement.

2.2. Model Data

We use data from experiments using three atmosphere-only general circulation models. All of the models include parameterizations of small-scale atmospheric processes such as convection, cloud processes, and gravity wave drag and a representation of land surface dynamics. Observational estimates of sea surface temperatures (SSTs) and sea ice concentrations (SICs) are used as boundary conditions. The models use historical atmospheric composition, except that after 2005 in the EC-Earth experiments the composition is taken from the CMIP5 Representative Concentration Pathway 8.5 scenario [*Moss et al.*, 2010].

The model data sets and their stochastic physics schemes are described in the following sections and summarized in Table 1.

Table 1. Summary of Model Configurations Evaluated and Their Abbreviated Names

Base Model	Configuration	Resolution (at Equator)	Short Name
EC-Earth	Deterministic low resolution	T255 (~80 km)	EC Det Low
	Stochastic low resolution	T255 (~80 km)	EC Stoch Low
	Deterministic high resolution	T1279 (~16 km)	EC Det High
Unified model	Deterministic low resolution	1.25° × 1.875° (~ 140 km × 210 km)	UM Det Low
	Stochastic low resolution	1.25° × 1.875° (~140 km × 210 km)	UM Stoch Low
	Deterministic high resolution	0.23° × 0.35° (~25 km × 40 km)	UM Det High
CAM4	Deterministic	0.9° × 1.25° (~100 km × 140 km)	CAM4 Det
	SPPT	0.9° × 1.25° (~100 km × 140 km)	CAM4 SPPT
	SKEBS	0.9° × 1.25° (~100 km × 140 km)	CAM4 SKEBS

2.2.1. EC-Earth (EC)

We used atmosphere-only integrations of EC-Earth version 3.1, as performed and documented by *Davini et al.* [2017]. The atmospheric component is based on the ECMWF Integrated Forecast System (IFS) cycle 36r4 in the "System 4" seasonal forecast model configuration [*Molteni et al.*, 2011]. It uses a bulk mass flux convection scheme based on the work of *Tiedtke* [1989], described in detail in the IFS Cy37r2 documentation at http://www.ecmwf.int. The model has been adapted and tuned for performing climate simulations. The "low-resolution" configurations used here have spectral atmospheric resolution T255, with model physics calculated on a reduced Gaussian grid with a grid spacing of approximately 80 km globally. We consider deterministic and stochastic configurations at this resolution ("EC Det Low" and "EC Stoch Low"). We also compare these with a "high-resolution" deterministic configuration ("EC Det High"), with spectral resolution T1279, corresponding to an approximate grid spacing of 16 km. Each configuration has 91 atmospheric levels. Radiative fluxes were tuned in the low-resolution deterministic configuration, and the model parameters were kept the same in each configuration except for small changes in the convection and nonorographic gravity wave drag schemes made in the high-resolution configuration. We use data between 1998 and 2007, which matches the period of the observational data (section 2.1). Ten ensemble members were used for each low-resolution configuration and one member for the high-resolution configuration. SSTs and SICs are specified according to the HadISST (Hadley Centre Sea Ice and Sea Surface Temperature) data set, version 2.1.1 [*Titchner and Rayner*, 2014].

In the stochastic model configuration, schemes very similar to the ECMWF SPPT scheme and SKEBS are activated [*Berner et al.*, 2009; *Palmer et al.*, 2009; IFS cycle 37r2 documentation Part V, available at http://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model/ifs-documentation]. In the SPPT scheme, the total physics tendencies of temperature, humidity, and winds are perturbed. The random number pattern that is multiplied against the tendencies is a sum of three independent random patterns with horizontal correlation scales of 500 km, 1000 km, and 2000 km, correlation time scales 6 h, 3 days, and 30 days and standard deviations 0.52, 0.18, and 0.06, respectively. The random number pattern is the same between the boundary layer and the stratosphere and is tapered in the vertical at the top and bottom of this region. The same random number pattern is used for each variable.

SKEBS perturbs vorticity tendencies by adding a random pattern that varies in all three spatial dimensions with a 7 h autocorrelation time scale. The spatial power spectrum of the perturbations follows a power law, with larger perturbations made to larger scales. At each point the perturbation is drawn from a distribution with a standard deviation that increases with increasing local total dissipation by the model numerics, orographic drag, and deep convection.

The only difference between the EC-Earth and ECMWF stochastic schemes is that in EC-Earth, *Davini et al.* [2017] implemented a fix in SPPT so that it conserves the global mean tendencies, which they found to be particularly important for conserving water and energy in long simulations.

Weisheimer et al. [2014] found that in the ECMWF seasonal forecasting model, which has a very similar atmospheric component to the EC-Earth model used here, SKEBS had little effect on the mean state compared to SPPT. We have confirmed, using the same data, that the impact of SPPT and SKEBS on the statistics of tropical rainfall variability that we present below is quite similar to that in EC-Earth (not shown). SKEBS also has a much

smaller impact than SPPT. Therefore, it seems reasonable to assume that SKEBS is also having little effect on these statistics in EC-Earth, and the effects we present are mostly due to SPPT.

Subramanian et al. [2017] tested the impact of amplifying perturbations associated with the 500 km-6 h scale in SPPT and of removing the perturbations associated with each of the other scales in the ECMWF seasonal forecasting model. Using the same data, we found that for the diagnostics of precipitation variability that we use, amplifying the 500 km-6 h scale perturbations increases the impact of SPPT, removing the 1000 km-3 day scale slightly reduces the impact and removing the 2000 km-30 day scale has little effect (not shown). Therefore, most of the impact of SPPT seems to be associated with the 500 km-6 h scale.

We also used coupled atmosphere ocean configurations of EC-Earth [*Davini et al.*, 2017] to examine the effect of stochastic physics on the simulated impact of increasing greenhouse gases on tropical variability (section 3.6). These use the same atmospheric component as the atmosphere-only configurations, coupled to the NEMO 3.3.1 ocean model [*Madec*, 2008] and the Louvain-la-Neuve sea ice model version 3 [*Vancoppenolle et al.*, 2012]. The atmospheric component has "low" spectral resolution T255 and the ocean model has 1° resolution in middle latitudes and higher resolution near the equator. We used periods 1980–2009 and 2070–2099 to represent "present-day" and "future" conditions, respectively.

2.2.2. Met Office Unified Model (UM)

Atmosphere-only integrations of the Met Office Unified Model Global Atmosphere 6.0 configuration [*Walters et al.*, 2017] were used, performed by *Sanchez et al.* [2016]. The model uses a bulk mass flux deep convection scheme based on the work of *Gregory and Rowntree* [1990]. We use "low-resolution" 1.25° × 1.875° (~140 km × 210 km at the equator) deterministic and stochastic configurations ("UM Det Low" and "UM Stoch Low") and a high-resolution 0.23° × 0.35° (~25 km × 40 km at the equator) deterministic configuration ("UM Det High"). Each configuration has 85 atmospheric levels and uses the same parameter settings except for a small change in one gravity wave drag parameter. We mostly use the final 10 years of the low-resolution runs (1992–2001) and the full 9 year and 7 month duration of the high-resolution run (December 1981 to June 1991). When comparing the annual mean precipitation between UM Det High and UM Det Low and GPCP-Mon in section 3.1, we use the December 1981 to June 1991 period for each. A single model run was used for each configuration. SSTs and SICs are specified according to the method of *Reynolds et al.* [2007].

Similarly to the EC-Earth and ECMWF models, the stochastic configuration has versions of the SPPT scheme and SKEBS activated [*Sanchez et al.*, 2016]. There are several notable differences between the SPPT schemes in the UM and EC-Earth. In the UM, the random pattern follows a Gaussian power law, so that it has greater power at large scales relative to small scales than in EC-Earth, and it has a single autocorrelation time of 6 h. Clear-sky radiation is not perturbed, and water and moist static energy are conserved column by column rather than globally. The standard deviation of the random pattern is 0.5, similar to that in the ECMWF scheme, except for gravity wave drag, for which the perturbations have standard deviation 0.42. The SKEBS implementation also differs in some details compared with that used in EC-Earth. For example, only wave numbers between 20 and 60 are perturbed, orographic drag is not taken into account in calculating the perturbations' standard deviations, the magnitude of the perturbations is reduced in the boundary layer, and stratosphere and perturbations are also added to the velocity potential.

2.2.3. NCAR Community Atmosphere Model, Version 4 (CAM4)

We used atmosphere-only integrations of CAM4 [*Gent et al.*, 2011], with details in *Christensen et al.* [2017]. The model uses a bulk mass flux deep convection parameterization based on the work of *Zhang and McFarlane* [1995]. We use a deterministic configuration ("CAM4 Det"), a configuration with a variant of SPPT activated ("CAM4 SPPT") and a configuration with a variant of SKEBS activated ("CAM4 SKEBS"). The model configurations all use the finite volume dynamical core and have a horizontal resolution of 0.9° × 1.25° (~100 km × 140 km at the equator) with 26 vertical levels, and all use the same parameters in the deterministic parameterization schemes. We use the final 10 years of the common period of each run (1991–2000), using a single model run for each configuration. SSTs and SICs are specified according to the data set of *Hurrell et al.* [2008].

The SPPT scheme is similar to that in EC-Earth but uses a random pattern with a horizontal decorrelation scale of 500 km and a decorrelation time of 6 h, so the large-scale and slowly varying parts of the pattern used in the ECMWF scheme are not present. However, it was noted in section 2.2.1 that most of the impact of stochastic physics in EC-Earth seems likely to be due to the 500 km-6 h scale in SPPT, so the important part of the stochastic schemes in EC-Earth is similar to the CAM4 SPPT scheme. The SKEBS implementation

works in a similar way to that in the ECMWF model, but potential temperature is perturbed as well as the streamfunction and the dissipation is assumed to be spatially and temporally constant and therefore so is the standard deviation of the perturbations. The standard deviation of the perturbations was made as large as it could be without changing the slopes of the potential and kinetic energy spectra of the stream function and potential temperature except for very high wave numbers.

2.3. Statistical Significance Assessment

For most of the diagnostics shown here, it has been verified that they generally vary little if they are plotted for the first and second halves of the time periods covered by each data set separately, by much less than would be required for our conclusions to be affected. Additionally, for the low-resolution EC-Earth data sets, it was checked that the results are similar to those shown for individual ensemble members. Therefore, for these diagnostics it is concluded that sampling variability is generally small compared to the signals seen in the data.

Diagnostics for which sampling uncertainty was found to be substantial were the differences in the mean precipitation between the stochastic and deterministic low-resolution model configurations (section 3.1) and ratios of rain rate frequency distributions in stochastic and high-resolution model configurations to those in the deterministic low-resolution configurations (section 3.2). The methods for assessing sampling uncertainty for those diagnostics are explained in the sections where the results are presented.

3. Results

3.1. Mean Precipitation

We first present the impact of stochastic physics on the time mean precipitation. Figures 1a, 1d, and 1g show the biases in annual mean precipitation of the low-resolution deterministic configurations of EC-Earth, the UM and CAM4 with respect to GPCP-Mon. Figures 1b, 1c, 1e, 1f, 1h, and 1i show the impacts of using stochastic physics and increasing the resolution in these models. With the exception of SKEBS in CAM4, the impact of using stochastic physics is a lot smaller than the biases of the deterministic models, so the stochasticity is not greatly changing this aspect of the models' behaviors.

Stippling in Figure 1 shows where differences are statistically significantly different from zero above the 95% level according to a Monte Carlo permutation test [*Efron and Tibshirani*, 1994]. For each pair of stochastic and deterministic model configurations, for each grid point separately, the combined set of annual mean precipitation values from both members of the pair was randomly separated into surrogate halves, and the difference between the means for each half was calculated. This was repeated 1000 times to calculate the probability that the magnitude of this difference would exceed the magnitude of the difference between the actual stochastic and deterministic model runs. Differences are deemed statistically significant at the 95% level if 5% or fewer of the differences between the surrogate halves of data are larger than the actual differences.

Figure 1b shows that in EC-Earth, stochastic physics increases mean precipitation in the tropical West Pacific and across southeastern Asia and decreases mean precipitation over the Maritime Continent. The effects in the tropical West Pacific are similar to those shown in the ECMWF seasonal forecasting system [*Weisheimer et al.*, 2014], as may be expected since the two models have similar atmospheric components, though the effects in EC-Earth are smaller. However, the impacts in other regions, such as the slight drying of Africa and wettening of tropical South America, are not very similar overall to those in the ECMWF model. The global mean root-mean-square error ($\overline{\text{RMSE}}$) of time-averaged precipitation is increased slightly from 0.78 mm/d to 0.80 mm/d. Increasing the horizontal resolution has a larger impact but does not reduce the biases overall, with the $\overline{\text{RMSE}}$ increasing to 0.90 mm/d (Figure 1c).

Figures 1e and 1f show that the UM Stoch Low and UM Det High configurations have mean precipitation mostly within 2 mm/d of that of UM Det Low, with the differences not generally being statistically significantly different from zero above the 95% level. Stochastic physics reduces the precipitation $\overline{\text{RMSE}}$ from 1.23 mm/d to 1.20 mm/d, and in UM Det High it is 1.13 mm/d (although note the UM Det High $\overline{\text{RMSE}}$ is calculated over a different time period, as given in section 2.2.2). CAM4 SPPT also only exhibits small differences in precipitation compared to CAM4 Det, with the $\overline{\text{RMSE}}$ increased from 1.02 mm/d to 1.05 mm/d (Figure 1h).

The lack of statistical significance of the differences at individual grid points may be because only one ensemble member is available for these runs. For the EC-Earth data sets, the impact of stochastic physics estimated from ensemble members chosen at random is similar to that estimated from all ensemble members and

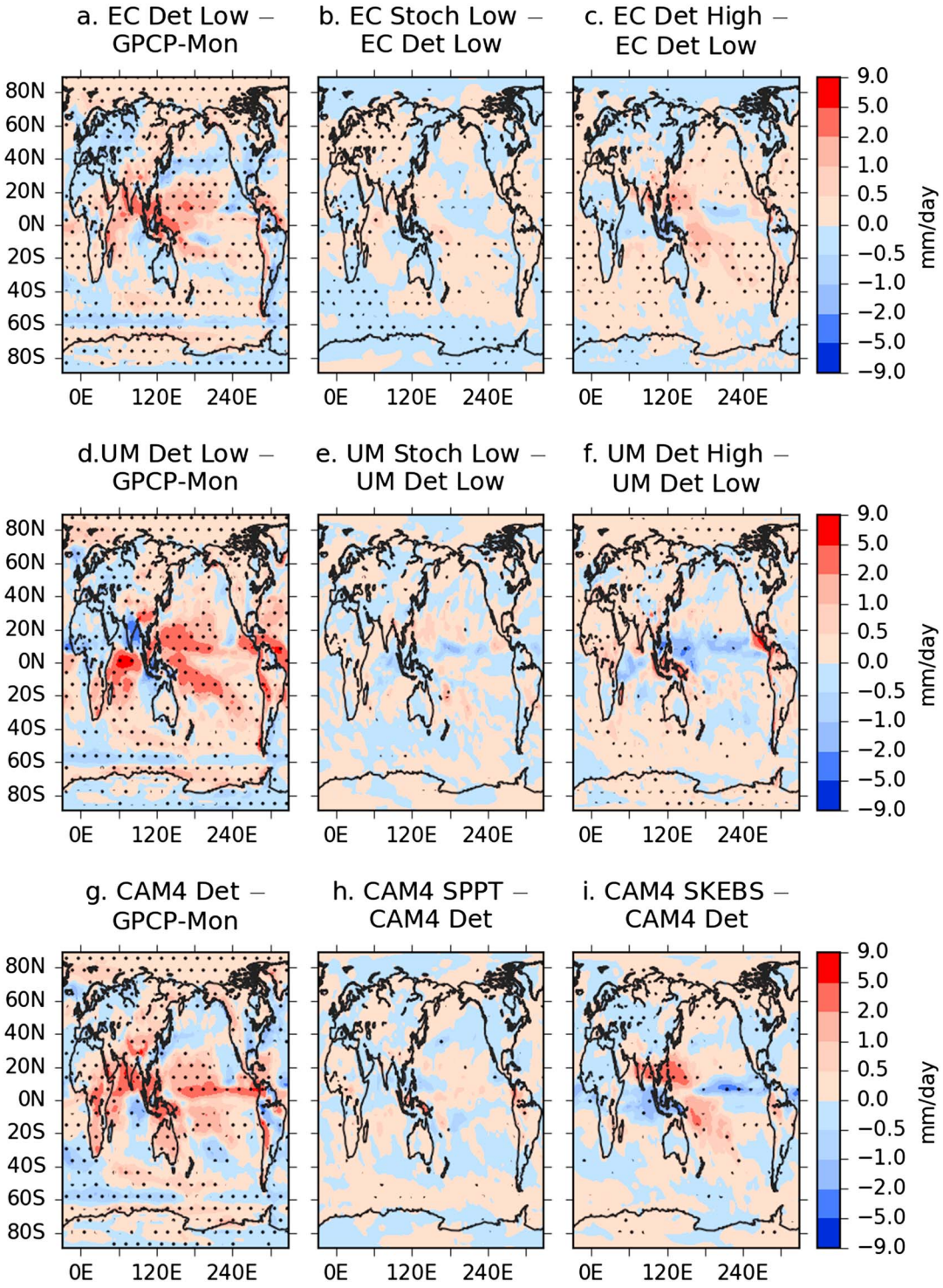

Figure 1. (a) The bias in annual mean precipitation in EC Det Low relative to GPCP-Mon. (b and c) The difference caused by using stochastic physics and increasing the resolution in EC-Earth. (d–f) Similar plots for the UM. (g) The bias in annual mean precipitation in CAM4 Det and (h and i) the difference produced by using SPPT and SKEBS in CAM4. Stippling shows where differences are statistically significant at the 95% level. The impact of stochastic physics and increasing resolution on the mean precipitation is generally less than 2 mm/d, except that of SKEBS in CAM4 which has a larger impact, and the impact varies between the models.

dissimilar overall to that in UM Stoch Low and CAM4 SPPT. This indicates that the differences between the diagnosed impacts in the different models cannot be explained by sampling variability and that the impacts of the stochastic physics schemes are actually different between EC-Earth and the UM and CAM4.

Figure 1i shows that SKEBS in CAM4 has a stronger impact than SPPT, causing changes in the mean precipitation by more than 3 mm/d in some locations, with increases over southern and eastern Asia and the southwestern tropical Pacific and decreases over the Maritime Continent and in the eastern tropical Pacific. This increases the size of model biases in some regions (e.g., southern Asia and the eastern Indian Ocean) and decreases them in others (e.g., the central and eastern tropical Pacific), without a clear overall improvement or deterioration globally (comparing Figures 1g and 1i). The precipitation $\overline{\text{RMSE}}$ is increased to 1.14 mm/d. The split between the precipitation maxima in the West Pacific Intertropical Convergence Zone is reduced by SKEBS (not shown).

Most of the changes in mean tropical precipitation are associated with convective precipitation (not shown). Of the stochastic schemes, only SKEBS in CAM4 substantially changes the grid-scale component of precipitation.

Therefore, the impact of stochastic physics schemes on the mean precipitation has significant dependence on the model and scheme used and varies between different regions. The following sections show that there are several impacts of stochastic physics on the precipitation variability that are found in different models and regions, and these therefore seem to be fairly independent of the impact of stochastic physics on the mean state.

3.2. Precipitation Rate Frequency Distributions

Here we examine the impact that stochastic physics has on the models' abilities to simulate realistic frequencies of rainfall of different intensities. Figures 2a, 2d, and 2g show the frequency distribution of different daily mean precipitation rates between 10°S and 10°N in GPCP, TRMM, and the different configurations of each model. Each row shows data for the different configurations of EC-Earth, the UM, and CAM4, respectively. We consider the 10°S–10°N equatorial band because this is a region where intense convection is important, and results are similar for individual latitudes in this range. Results for the wider tropical (23°S–23°N) band (not shown) are qualitatively similar but show some quantitative differences due to aggregating together regions where model biases and the impacts of the stochastic schemes are more different than those in the 10°S–10°N band, so the interpretation of the results is more complicated. (Note that these data for EC-Earth were also shown by *Davini et al.* [2017] and are included here for completeness, with this section being an expansion of their investigation.)

Figures 2b, 2e, and 2h show the frequencies of different rain rates in each model data set in the 10°S–10°N band as a fraction of those in GPCP. Figures 2c, 2f, and 2i show frequencies in the stochastic model configurations, and high-resolution configurations where applicable, as a fraction of those of the deterministic low-resolution configuration of each respective base model, to show the impact of stochastic physics and increasing model resolution.

Sampling variability was found to be substantial for the estimates of rainfall rate frequencies as a fraction of those in the deterministic low-resolution model configurations (Figures 2c, 2f, and 2i). Therefore, 95% confidence intervals were also calculated and plotted. These were calculated using a bootstrap method. First, for each data set, individual years of data were randomly sampled with replacement to create a surrogate data set. Then the ratios of rain rate frequencies between the relevant pair of surrogate data sets were calculated. This was repeated 1000 times to give an estimate of the distribution of the ratios associated with sampling uncertainty, and the confidence intervals were derived from these.

We discuss first the results pertaining to EC-Earth. Between 10°S and 10°N, the EC Det Low configuration has a frequency of relatively light precipitation rates below 15 mm/d that is too high by ~50% and has a too low frequency of moderate rates between 20 and 30 mm/d, by ~50% in the 25–30 mm/d interval (green curve in Figure 2b). For relatively heavy precipitation rates, higher than 30 mm/d, TRMM indicates much higher frequencies than GPCP, so the true frequencies are uncertain. EC Det Low has a frequency of rates above 30 mm/d that is only ~40% of that in GPCP and far below that in TRMM, however (Figure 2b), which is indicative of the simulated frequency being too low.

Stochastic physics has the effect of increasing the frequency of the lightest rain rates, decreasing the frequency of relatively moderate rain rates, and increasing the frequency of relatively heavy rain rates

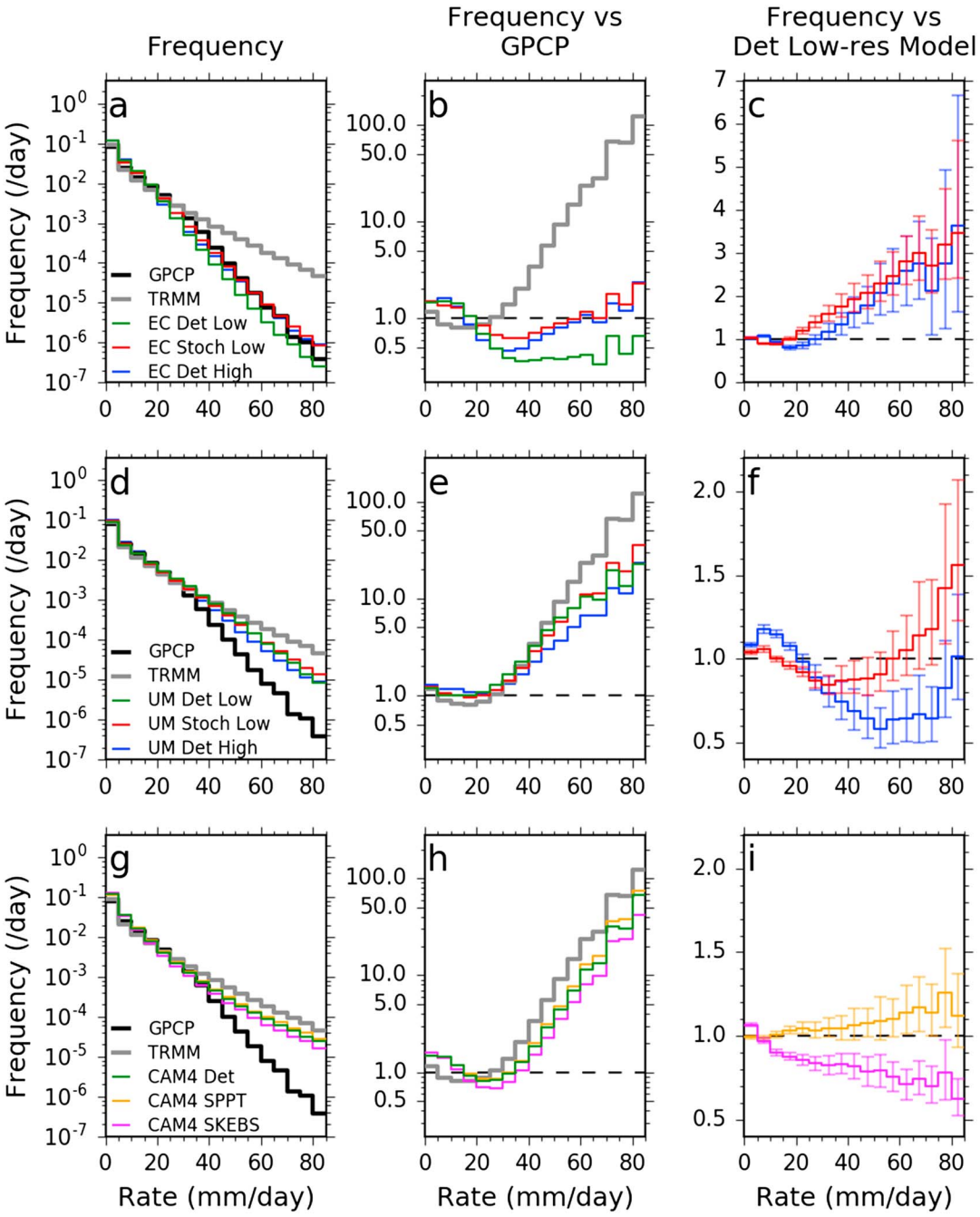

Figure 2. (a, d, and g) The frequency distributions of daily mean precipitation rates in the satellite and model data sets. Each row shows model data for a different family of model configurations. The precipitation has been averaged over 2.5° × 2.5° grid boxes, and data are aggregated for all grid boxes between 10°S and 10°N. Frequencies are shown for each 5 mm/d interval, excluding rates less than 0.1 mm/d. (b, e, and h) The frequencies of rain rates in the model data sets as a fraction of those in GPCP. (c, f, and i) The frequency as a fraction of that in the deterministic low-resolution model in each family of configurations. Vertical bars in Figures 2c, 2f, and 2i show the 95% confidence interval. Horizontal dashed lines show a ratio of 1. Note that the vertical axes in Figures 2a, 2b, 2d, 2e, 2g, and 2h are logarithmic and they have different ranges in Figures 2c, 2f, and 2i. Stochastic physics substantially raises the frequency of heavy precipitation events in all of the models, except for the SKEBS scheme in CAM4.

(red curves in Figures 2b and 2c). This matches the pattern expected from purely adding random perturbations to the precipitation [*Subramanian et al.*, 2017] as in the No Feedbacks conceptual model. The biases are mostly reduced—the frequency of rates between 5 and 15 mm/d is reduced by ~10% and the negative biases between 20 and 40 mm/d are approximately halved with respect to GPCP. Frequencies of heavier rain rates are increased by up to a factor of ~2–4 at the highest precipitation rates (Figures 2b and 2c), so that they are more in line with the observational estimates. The bias at light rain rates between 0.1 and 5 mm/d is slightly greater by ~5%, however. For comparison, increasing the resolution of the deterministic configurations gives a smaller improvement in the biases at rates between 5 and 40 mm/d and a similar increase in the frequency at higher rates (blue curves in Figures 2b and 2c).

Figure 3 shows similar diagnostics in a box over northern South America (10°S–5°N and 50°–75°W; shown in Figures A1a and A1b). In this region, the discrepancy between GPCP and TRMM is smaller (Appendix A), so it seems more clear here what the model biases are and whether the stochastic physics schemes are improving the model simulations. Note, however, that the observational uncertainty can be considerably larger than the difference between the GPCP and TRMM estimates. Part of the reason that the data sets agree better here may be that rain gauge data are used to calibrate their monthly means, but the estimates of daily totals are still made using different methods.

Figures 3b and 3c show that in this region, EC Det Low has a too high frequency of relatively light precipitation rates up to 20 mm/d, by ~40% between rates 10 and 15 mm/d. The simulated frequency of relatively heavy precipitation rates is much lower than that observed, by over 90% at rates above 30 mm/d. Stochastic physics has a large beneficial effect on these biases, approximately halving the excess of relatively light events between 5 and 15 mm/d (Figure 3b) and increasing the frequency of heavier precipitation rates above 20 mm/d from between a factor of ~2 and a factor of ~10–40 at the highest rates (Figure 3c). The stochastic physics schemes do not increase the frequency of rates above ~25 mm/d by enough to match those observed, however.

Conversely, increasing the resolution actually increases the biases in this region up to precipitation rates of 35 mm/d (compare the green and blue curves in Figure 3b). The frequencies of precipitation rates higher than this are beneficially increased by increasing the resolution, though by less than is achieved by using stochastic physics (Figure 3c).

The UM Det Low and CAM4 Det configurations have smaller positive biases between 10°S and 10°N in the frequency of relatively light precipitation rates below 15 mm/d and much more frequent heavy precipitation events, above 20 mm/d, than EC-Earth (Figures 2d, 2e, 2g, and 2h). In northern South America, the positive biases at rates below 15 mm/d are similar to those in EC-Earth and the frequencies of higher rates are again greater than in EC-Earth (Figures 3d, 3e, 3g, and 3h). The frequencies of high precipitation rates (above 35 mm/d) across all grid points between 10°S and 10°N are between those in GPCP and TRMM (Figures 2d, 2e, 2g, and 2h). In northern South America, however, the frequencies of high precipitation rates (above 25 mm/d) are much lower than observed, in common with EC-Earth, by up to ~80% and ~50% between rates 40 and 50 mm/d in the UM and CAM4, respectively.

In the UM, between 10°S–10°N stochastic physics increases the frequency of light precipitation rates below 10 mm/d by about 5%, decreases frequencies of relatively moderate rates between 15 and 55 mm/d by up to ~20% and increases the frequency of heavy rates between 60 and 80 mm/d by up to ~60% (Figure 2f). The pattern of there being an increase in the frequency of light rain rates, a decrease in the frequency of relatively moderate rain rates, and an increase in the frequency of heavy rain rates is similar to the impact in EC-Earth. This is also consistent with the No Feedbacks conceptual model. However, the range of rates whose frequencies are decreased is wider. The positive frequency bias increases at light rates between 0.1 and 10 mm/d and decreases at relatively moderate rates between 20 and 40 mm/d relative to that in the low-resolution deterministic configuration. Since the simulated distribution for rates above 40 mm/d for both the deterministic and stochastic low-resolution models is between that of GPCP and TRMM, it is not clear if stochastic physics is improving the distribution at these rates.

In northern South America, the stochastic physics in the UM has a qualitatively similar effect to that in EC-Earth, reducing the frequency of relatively moderate rain rates between 10 and 25 mm/d and increasing the frequencies of relatively heavy rates above ~25 mm/d, by more than ~40% at rates above 40 mm/d (Figure 3f). This substantially improves the simulated precipitation distribution in this region.

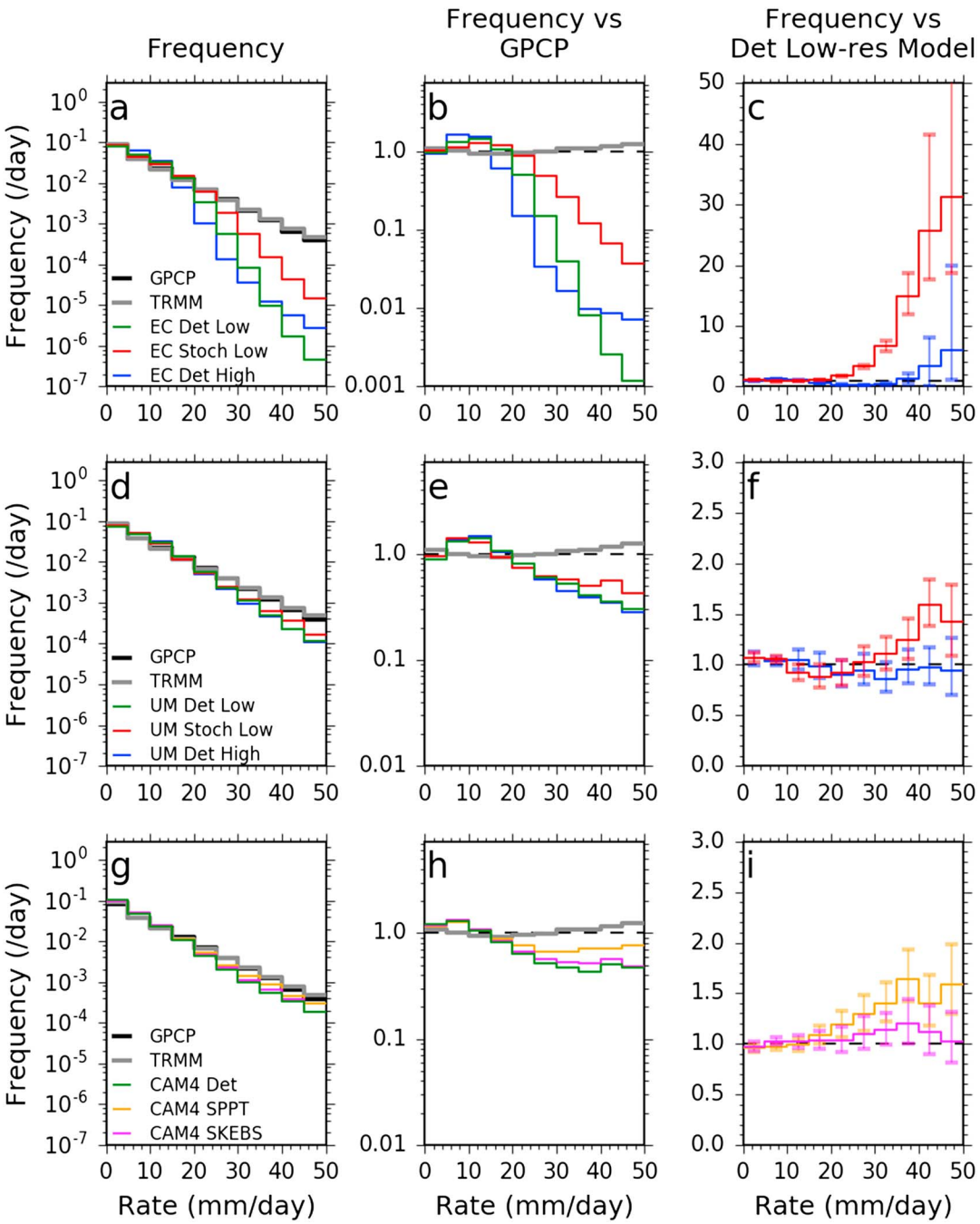

Figure 3. (a–i) As in Figure 2 but for grid boxes in the northern South America region (10°S–5°N and 50°–75°W) where the GPCP and TRMM data are in closer agreement. Note again that the vertical axes in Figures 3a, 3b, 3d, 3e, 3g, and 3h are logarithmic and they have different ranges in Figures 3b, 3e, and 3h and 3c, 3f, and 3i. Stochastic physics greatly increases the frequency of heavy precipitation events in all of the models, making the rain rate distribution more realistic in every case.

Unlike in EC-Earth, in the UM the impact of increasing horizontal resolution is not very similar to that of stochastic physics (compare the red and blue curves in Figures 2f and 3f). Stochastic physics performs better at reducing the deficit of high rain rates in northern South America (Figures 3e and 3f).

In CAM4, between 10°S–10°N, SPPT increases the frequency of moderate and heavy precipitation rates above 15 mm/d, by up to ~15% at rates above 60 mm/d (Figure 2i). This slightly reduces the negative frequency bias at relatively moderate rates between 20 and 30 mm/d. Conversely, SKEBS decreases frequencies for all rates

above 5 mm/d, by over 20% for rates above 60 mm/d—this may be associated with it decreasing the mean rainfall averaged between 10°S–10°N (Figure 1i). This shows that simply adding stochastic perturbations does not always increase the frequency of heavy precipitation events. The biases are reduced for rates between 5 and 15 mm/d and increased between 20 and 35 mm/d. The biases are also increased at the lowest rates, between 0.1 and 5 mm/d, since SKEBS increases the frequency by 6%.

In northern South America, the impact of SPPT in CAM4 is again to increase the frequencies of high rain rates, which reduces the model bias (Figures 3h and 3i), but by a larger amount than in the 10°S–10°N band, in common with the impact of stochastic physics in the other models. SKEBS does increase the frequency of heavy precipitation rates in this region, by a lesser amount than SPPT, in contrast to its impact in the whole 10°S–10°N band, showing that its impact is regionally variable.

It is also important to consider how similar results for the 10°S–10°N band are to those for equatorial regions at different longitudes. In EC-Earth, the impact of stochastic physics is similar at most longitudes (not shown). In the UM, its impact is also similar at most longitudes except over equatorial Africa and the tropical Atlantic, where frequencies of nearly all rain rates above 0.1 mm/d are increased, reducing model biases there. In CAM4, SPPT has a similar impact at all longitudes, with the strongest impacts being over the equatorial South American and African land areas. SKEBS in CAM4 increases the frequencies of most rain rates over equatorial South America, equatorial Africa, and in the tropical West Pacific and decreases frequencies of rain rates above ~5 mm/d elsewhere—these regional differences are perhaps associated with large changes in the mean rainfall in different regions caused by SKEBS (section 3.1).

In all of the models, there is a much smaller frequency of dry days (with less than 0.1 mm of precipitation) than in GPCP and TRMM, as in the models studied by *Stephens et al.* [2010]. This is not greatly improved by using stochastic physics or increasing the horizontal resolution in any case (not shown). Dry days make up 33% and 29% of the days in GPCP and TRMM, respectively, across all grid points between 10°S–10°N but only 8% of days in the EC Det Low and CAM4 Det configurations and 24% in UM Det Low. Stochastic physics causes absolute changes that are smaller than 2%. The stochastic schemes used in these models do not treat the triggering of convection as stochastic, and it could be interesting to test whether this diagnostic could be improved by using a scheme that did so [e.g., *Rochetin et al.*, 2014; *Peters et al.*, 2017].

To summarize, the stochastic physics schemes in EC-Earth and the UM slightly increase the frequency of the lightest rain rates, reduce the frequency of moderate rain rates, and increase that of heavy rain rates in most regions, consistent with the No Feedbacks conceptual model [*Subramanian et al.*, 2017]. In CAM4, the SPPT scheme increases the frequencies of all but the lowest rain rates, while the impact of SKEBS is regionally variable, which may be associated with its regionally variable impact on the mean precipitation (Figure 1i). Thus, feedbacks from changes in the statistics of the resolved variables seem important for the CAM4 results. In all of the models stochastic physics probably improves the rain rate distribution in northern South America, where GPCP and TRMM agree quite well on the distribution. In the whole 10°S–10°N zonal band, stochastic physics improves the simulated frequency of rain rates between 5 and 30 mm/d in EC-Earth and to a lesser extent that between 20 and 30 mm/d in the UM and CAM4. Observational uncertainty makes it difficult to determine if stochastic physics causes an improvement at higher rain rates, but it is more likely than not to cause an improvement in EC-Earth, which greatly underestimates the frequency of high rain rates compared to both GPCP and TRMM in its low-resolution deterministic configuration. The positive frequency bias at light rain rates is slightly increased in EC-Earth and the UM. There are large quantitative differences between the impact of stochastic physics in each model, however. Stochastic physics gives a larger improvement in the rain rate distributions in northern South America in EC-Earth and the UM than does increasing horizontal resolution, indicating that SPPT and SKEBS are representing important variability that is still not resolved even at a horizontal resolution of ~20–40 km.

3.3. Precipitation Autocorrelation

Another important aspect of tropical rainfall is its persistence, diagnosed here using the autocorrelation function. Figure 4 shows this averaged over all grid points between 10°S and 10°N in GPCP, TRMM, and the model data sets. In all model configuration families, the deterministic low-resolution model (green dashed curves) has a higher autocorrelation than that in GPCP and TRMM. In the cases of EC-Earth and the UM, the autocorrelation in the deterministic low-resolution model is greater than that in GPCP and TRMM by an amount considerably larger than the difference between GPCP and TRMM, and therefore, in these models the autocorrelation is likely to be too large.

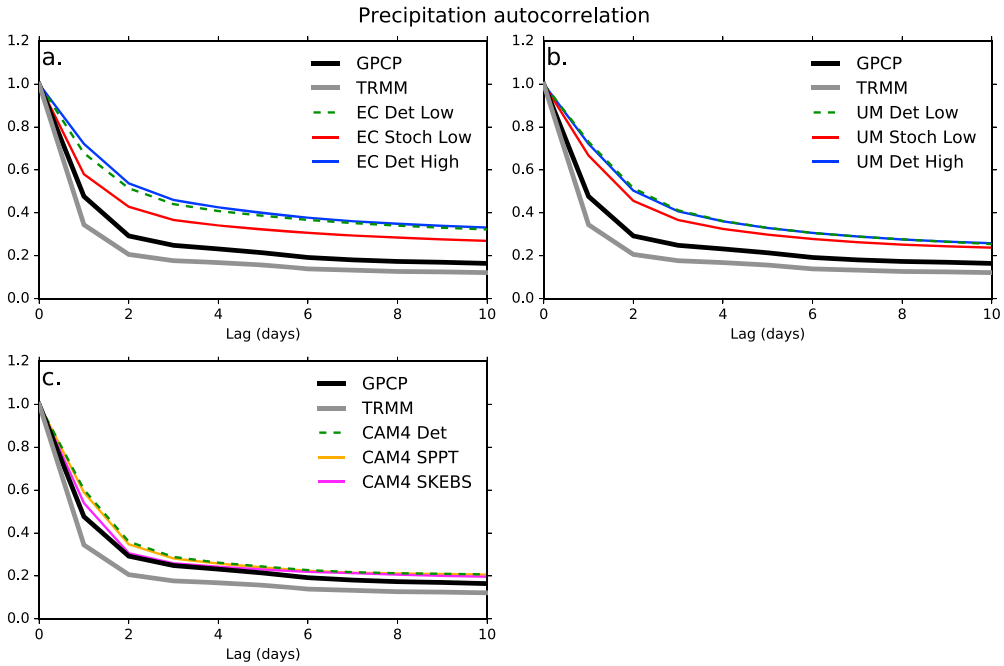

Figure 4. The mean autocorrelation between 10°S–10°N of daily mean precipitation averaged over 2.5° × 2.5° grid boxes in the satellite and model data sets. (a–c) Data for each family of model configurations along with TRMM and GPCP data. Most of the stochastic physics schemes reduce the autocorrelation, the exception being SPPT in CAM4, and this is very likely to be an improvement in EC-Earth (EC) and the UM.

Stochastic physics is very effective at reducing the autocorrelation in EC-Earth (Figure 4a) to values closer to those observed, and it has a more modest beneficial impact in the UM (Figure 4b). In CAM4, SKEBS reduces the autocorrelation slightly, while the SPPT scheme has very little effect (Figure 4c). The reduction in the autocorrelation is consistent with the No Feedbacks conceptual model, so there is not clear evidence of the importance of large-scale feedbacks from this diagnostic.

In EC-Earth and the UM, increasing the horizontal resolution has very little impact (Figures 4a and 4b), indicating again that stochastic physics is aiding with simulating the effects of variability that are not well resolved even at the high horizontal resolutions in these configurations.

The results are generally similar for different tropical regions (not shown). In each low-resolution deterministic model, the biases are typically larger over land. The lag 1 day autocorrelations in each model are positively biased everywhere, except over some parts of the tropical Pacific and Indian Oceans in CAM4 Det, where the autocorrelations are between those in GPCP and TRMM. Stochastic physics improves biases everywhere in EC-Earth and the UM. SKEBS in CAM4 improves biases except in places in the tropical Pacific where the sign of the bias is uncertain, where the changes are within the range spanned by the GPCP and TRMM data sets. SPPT in CAM4 has little effect anywhere.

Therefore, the persistence of rainfall on daily time scales is made more realistic by the stochastic physics in EC-Earth and the UM and by SKEBS in CAM4.

3.4. Precipitation Power Spectra

It is also important to know if stochastic physics has beneficial or detrimental effects on variability at different length and time scales. We diagnose this aspect of its effects by calculating space-time power spectra of precipitation following the method of *Wheeler and Kiladis* [1999] (but considering the absolute, not background removed, spectra and considering a larger range of wave numbers). Figure 5 shows diagnostics for the low-resolution configurations of the different models of the spectral power of precipitation, as a function of zonal wave number and frequency, averaged between 15°S–15°N in meridionally symmetric equatorial wave modes (results for the antisymmetric wave modes are very similar (not shown)).

Figures 5a–5c show the ratio of the power in the deterministic low-resolution configuration of each model to that in TRMM. Figures 5d–5f show similar ratios with respect to the power in GPCP. All of the models have less

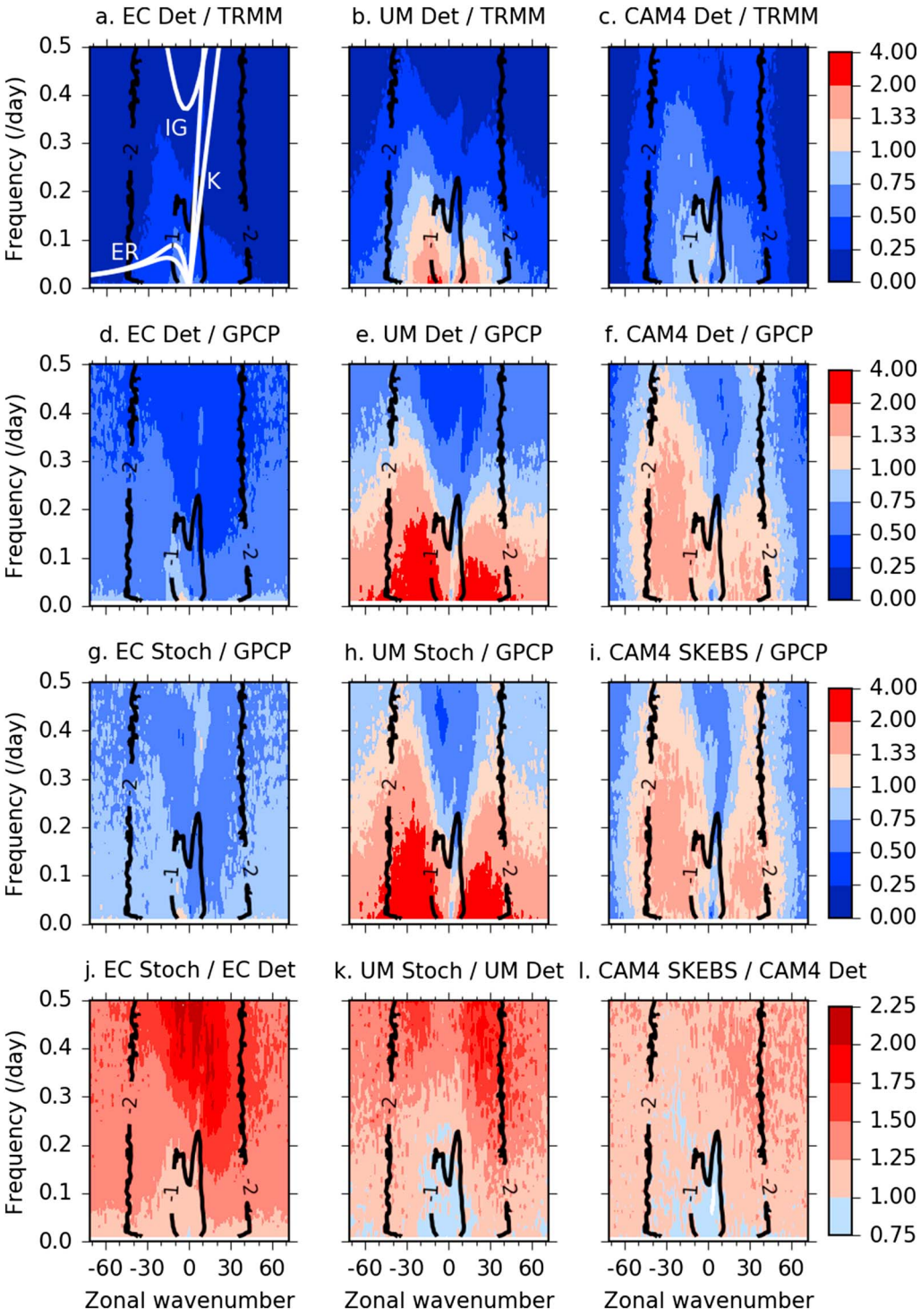

Figure 5. Filled contours show ratios between different data sets of the spectral power of precipitation in meridionally symmetric equatorial wave modes averaged between 15°S–15°N as a function of zonal wave number and frequency. Results relating (a, d, g, and j) to the low-resolution EC-Earth configurations, (b, e, h, and k) to the low-resolution UM configurations, and (c, f, i, and l) to the CAM4 Det and SKEBS configurations. Figures 5a–5c show the ratio of power in the deterministic configurations to that in TRMM, Figures 5d–5f the ratio of power in the deterministic configurations to that in GPCP, Figures 5g–5i the ratio of power in the stochastic configurations to that in GPCP, and Figures 5j–5l the ratio of the power in the stochastic configurations to that in the deterministic configurations. White lines in Figure 5a indicate theoretical dispersion curves for Kelvin waves (K), wave number 1 inertia gravity waves (IG), and equatorial Rossby waves (ER) with equivalent depths 12 m and 50 m. Stochastic physics is effective at increasing the power and reducing the model biases at high frequencies in EC-Earth and the UM, with SKEBS in CAM4 having a smaller effect.

power at frequencies above 0.2 per day at most wave numbers than is exhibited by both TRMM and GPCP. In EC-Earth the power is less than half that in both TRMM and GPCP at these high frequencies between about wave numbers −20 to 40, with smaller biases in the Kelvin wave part of the spectrum (Figures 5a and 5d). The power deficit reaches a similar relative size in the UM and CAM4 in some regions of the spectral space, though the biases are smaller overall (Figures 5b, 5c, 5e, and 5f).

Figures 5g–5i show the ratio of the power in the EC Stoch Low, UM Stoch Low, and CAM4 SKEBS configurations to that in GPCP. Figures 5j–5l show the ratios of the powers in the stochastic and deterministic configurations. In EC-Earth, stochastic physics is effective at reducing the biases (Figure 5g), increasing the power at nearly all frequencies and wave numbers. Its impact increases as frequency increases and it raises the power by over a factor of 2 for frequencies above ∼0.4 per day for most wave numbers between −10 and 10 (Figure 5j). Therefore, in this model, stochastic physics reduces the biases most where they are greatest in the deterministic low-resolution configuration.

Similarly in the UM, stochastic physics increases the power at high frequencies above ∼0.2 per day the most (Figures 5h and 5k), generally reducing the power deficit. Stochastic physics actually decreases power at most low frequencies below ∼0.1 per day for wave numbers smaller than ∼15, which reduces biases for most of these wave numbers and frequencies.

In CAM4, SKEBS increases the power at most wave numbers and frequencies, except at low wave numbers and frequencies like in the UM, but its effect is much smaller (Figures 5i and 5l). SPPT in CAM4 also increases the power at most wave numbers and frequencies, but by less than 20% for the large majority (not shown).

The impact of stochastic physics does not seem to project strongly onto equatorial wave modes in any model, where the dispersion curves for important wave modes are indicated by white lines in Figure 5a. The background-removed spectra are also not strongly affected (not shown). The impact is not clearly associated with the space and time scales of the stochastic schemes' random patterns alone. For example, in EC-Earth, the high-frequency components of the random pattern have decorrelation length scales of 500 km and 1000 km, corresponding to zonal wave numbers 80 and 40, respectively, whereas stochastic physics causes the biggest fractional increase in the power at wave numbers smaller than 20 (Figure 5j). If the spectra of the parameterized tendencies were known, it would be interesting to test whether the impact of stochastic physics on the spectra is similar to that expected in the No Feedbacks conceptual model, but this is not tested here. It seems notable, though, that there is a substantial increase in power at small wave numbers and high frequencies in EC-Earth and the UM (Figures 5j and 5k), which is suggestive of feedbacks carrying the influence of stochastic physics to length scales larger than those of weather systems that have similarly short time scales.

Increasing the horizontal resolution in EC-Earth and the UM does not have such beneficial impacts (not shown). The deterministic, high-resolution configuration of EC-Earth actually has less power and larger biases at high frequencies and also at high wave numbers than the EC Det Low configuration. UM Det High has increased power at high wave numbers and decreased power at low wave numbers at all frequencies relative to UM Det Low, which increases biases of the deterministic, low-resolution configuration in some parts of the spectrum and decreases biases in others.

Therefore, stochastic physics is effective at reducing the problem of there being too little power at high frequencies in the deterministic low-resolution configurations of the models evaluated here and also improves biases at most low frequencies for low wave numbers. It is more effective at reducing the biases than increasing the horizontal resolution in EC-Earth and the UM, though the stochastic configurations of the different models still have less power at high frequencies than in TRMM and GPCP.

3.5. Precipitation Standard Deviation

Another important consideration is how stochastic physics affects the total rainfall variability in different regions. Figure 6 shows diagnostics of the standard deviation of daily mean precipitation. The panels show, on descending rows, the differences in the standard deviation between the deterministic low-resolution configuration of each model and that in TRMM and GPCP, the differences between that in a stochastic low-resolution configuration and that in GPCP (again considering just the CAM4 SKEBS configuration of the CAM4 model), and the differences between that in the stochastic and deterministic low-resolution configurations.

EC Det Low has a lower standard deviation of precipitation in most equatorial regions than both TRMM and GPCP (Figures 6a and 6d). This bias is generally largest where the mean precipitation is largest (not shown).

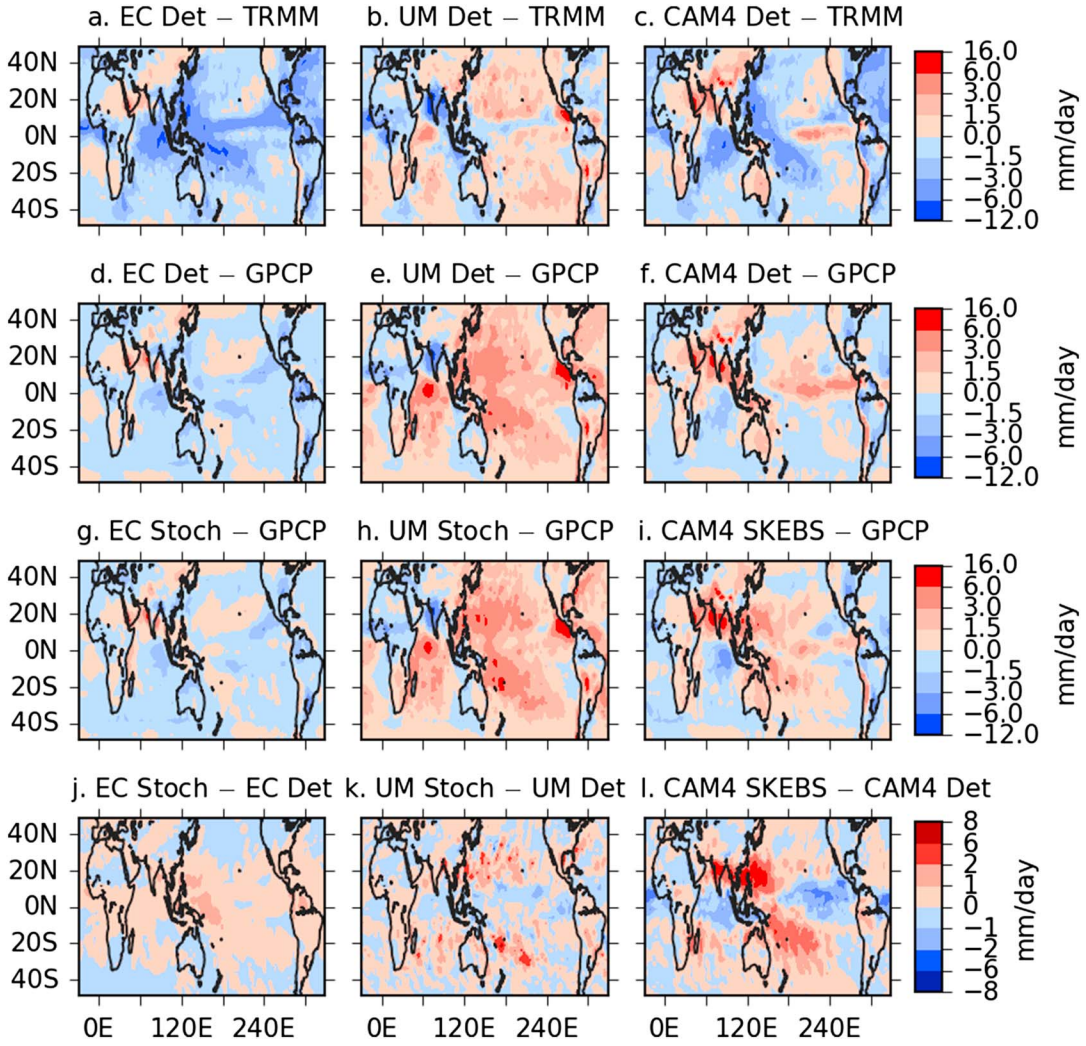

Figure 6. Filled contours show differences in the standard deviation (SD) of precipitation between different data sets. Results relating (a, d, g, and j) to the low-resolution EC-Earth configurations, (b, e, h, and k) to the low-resolution UM configurations, and (c, f, i, and l) to the CAM4 Det and SKEBS configurations. Figures 6a–6c show differences in the SD between the deterministic configurations and TRMM and Figures 6d–6f the differences with respect to GPCP. Figures 6g–6i show the differences in the SD between the stochastic configuration and GPCP and Figures 6j–6l the differences between the stochastic and deterministic configurations. Stochastic physics increases the SD and reduces biases in most equatorial regions in EC-Earth, while stochastic physics in the UM and SKEBS in CAM4 increase the SD mostly in subtropical regions, which improves some biases and worsens others.

Stochastic physics increases the standard deviation in most equatorial regions, particularly in the tropical West Pacific and over land (Figure 6j), generally substantially reducing the standard deviation biases in the equatorial regions, though not to zero, with biases of the stochastic configuration shown in Figure 6g. On average between 20°S and 20°N, stochastic physics increases the standard deviation by 0.33 mm/d, with the difference from the GPCP value (5.67 mm/d) changed to −0.52 mm/d from −0.85 mm/d. This impact from stochastic physics is what is expected from the No Feedbacks conceptual model.

UM Det Low has a higher standard deviation of precipitation than both GPCP and TRMM in subtropical Pacific regions, in the western Indian Ocean, over southern Africa and central South America (Figures 6b and 6e). It has a lower standard deviation than both GPCP and TRMM in equatorial land regions and over India. Stochastic physics reduces the standard deviation in most equatorial oceanic regions and increases it over equatorial Africa, equatorial South America, and in some subtropical regions, particularly in the West Pacific (Figure 6k). These impacts decrease standard deviation biases in some regions, such as in equatorial Africa and South America, and increase them in others, such as in the subtropical West Pacific. The substantial deficit in the

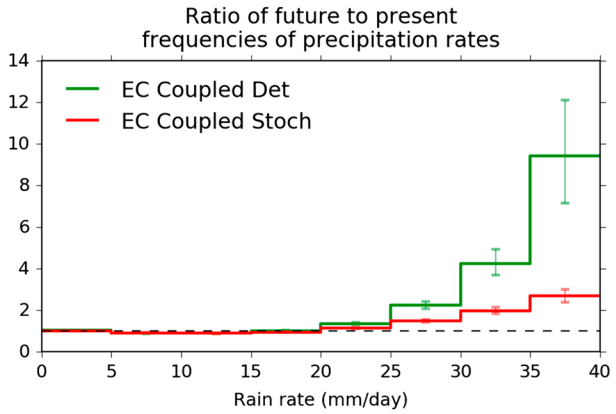

Figure 7. The ratio of frequencies of rain rates in the northern South America region (10°S–5°N and 50°–75°W) between future (2070–2099) and present-day (1980–2009) periods of coupled-ocean low-resolution EC-Earth integrations. The green and red curves show the ratios for the deterministic and stochastic model configurations, respectively. Vertical bars show the 95% confidence intervals. The greenhouse gas forcing followed the RCP8.5 Intergovernmental Panel on Climate Change scenario. The inclusion of stochastic physics greatly changes the simulated increase in the frequency of high precipitation rates in this region.

standard deviation over India is not much affected. The decrease in standard deviation in many regions near the equator is inconsistent with the No Feedbacks conceptual model, indicating that complex feedbacks from the resolved model variables are occurring. Stochastic physics decreases the standard deviation averaged over 20°S–20°N by 0.02 mm/d.

The standard deviation in the deterministic low-resolution configuration of CAM4 is in between that in TRMM and GPCP in many tropical regions, though it has a lower standard deviation than in both data sets in the southern Indian Ocean and in most of South America, and a higher standard deviation in southwest Asia and in the eastern equatorial Pacific (Figures 6c and 6f). The SKEBS configuration has a substantially larger standard deviation in southwest Asia and northeast of Australia and reduces it predominantly in the tropical eastern Pacific and the equatorial Indian Ocean (Figure 6l). This increases standard deviation biases in the Indian Ocean region and in the West Pacific and decreases them in the tropical eastern Pacific and in the Maritime Continent (comparing Figures 6f and 6i). The impact has a similar spatial pattern to the impact of SKEBS on the mean precipitation (Figure 1i), indicating that the effect of SKEBS on the precipitation variability is closely associated with its impact on the mean.

SPPT in CAM4 has a much smaller impact (not shown), increasing the standard deviation in the Maritime Continent and equatorial South America by up to ~2 mm/d and decreasing it in subtropical regions to the north and south of the Maritime Continent by a similar amount. SKEBS increases the standard deviation averaged over 20°S–20°N by 0.06 mm/d and SPPT increases it by only 0.01 mm/d.

These impacts can be contrasted with those of increasing horizontal resolution (not shown). Using high resolution in EC Earth decreases rather than increases the standard deviation of equatorial precipitation compared to the EC Det Low configuration, except in the western Pacific, so it does not help to correct the model standard deviation bias. The standard deviation is instead increased near southern Asia and in the southern subtropics between the western Indian Ocean and the eastern Pacific. In the UM, increasing the resolution decreases the standard deviation in much of the tropics except over India, central America, Australia, and near Madagascar, which does help to correct the positive standard deviation biases in the tropical Pacific and Indian Oceans shown in Figures 6b and 6d.

Overall, therefore, in EC-Earth stochastic physics has an impact similar to that expected from the No Feedbacks conceptual model—the standard deviation of precipitation is increased in most tropical regions, which reduces differences from observations. The impact of stochastic physics in the UM and CAM4 is more complex, however, and reduces some biases while increasing others. In EC-Earth and the UM, stochastic physics has quite a different impact to that of increasing horizontal resolution.

3.6. Impact on Projections of Future Tropical Precipitation Changes

As an example of how stochastic parameterizations may affect simulations of future climate change, Figure 7 shows the ratio of frequencies of rain rates in the northern South America region (10°S–5°N and 50°–75°W) between future (2070–2099) and present-day (1980–2009) time slices of coupled atmosphere-ocean low-resolution EC-Earth integrations, both for the deterministic and stochastic model configurations. The deterministic and stochastic coupled atmosphere-ocean configurations have similar precipitation rate distributions to those of the corresponding atmosphere-only configurations in this region (not shown). Both coupled models simulate a large fractional increase in the frequency of rain rates above 25 mm/d. This increase

is much larger in the deterministic model configuration, by more than a factor of 3 in the rain rate interval 35–40 mm/d. Vertical bars show the 95% confidence interval of the ratios, estimated using the same bootstrap method as in section 3.2, and these do not overlap at high rain rates. The projected changes also differ for the 10°S–10°N band, by ∼25% at heavy rain rates (not shown). Therefore, the use of stochastic physics strongly affects the simulated climate change response of tropical precipitation extremes and would strongly affect the quantitative attribution of observed extreme precipitation to climate change using this model. This does not mean that the change simulated by the stochastic model is more likely to be right in this case, since the climatology it simulates is far from adequate in northern South America (section 3.2), but it indicates that use of stochastic physics in models may be important for quantifying the impact of climate change on heavy tropical rainfall events.

4. Conclusions

We have presented analysis of the impacts of stochastic parameterization schemes on the variability of simulated tropical precipitation on daily to weekly time scales in three different atmospheric models: EC-Earth, the Met Office Unified Model (UM), and the Community Atmosphere Model, version 4 (CAM4). The results help to show the importance of small-scale variability that is not resolved in these models, at horizontal scales below ∼1°, for determining the statistics of tropical precipitation variability at this scale or larger. The stochastic physics schemes are variants of the stochastically perturbed physical tendencies scheme (SPPT) [*Buizza et al.*, 1999; *Palmer et al.*, 2009] and the stochastic kinetic energy backscatter scheme (SKEBS) [*Shutts*, 2005; *Berner et al.*, 2009]. The impacts of the two schemes were tested separately in CAM4, and they were combined in EC-Earth and the UM. For the schemes used in EC-Earth, however, results from experiments with the ECMWF seasonal forecasting model indicate that almost all of the impacts are due to SPPT, with most impact coming from the scale of the random pattern that matches that used in CAM4 (section 2.2.1).

Several improvements due to these schemes are found to be quite robust, being present in at least three of the four stochastic model configurations considered:

1. The frequency of the heaviest equatorial rainfall events is substantially increased in all configurations using SPPT (Figures 2c, 2f, and 2i, and 3c, 3f, and 3i). SKEBS in CAM4 increases the frequency of heavy rainfall events in northern South America, but not on average across the whole 10°S–10°N band, which may be associated with it substantially reducing the mean rainfall in this region (Figure 1i). Stochastic physics schemes improve the precipitation rate distribution in every case in northern South America, where observational uncertainty seems relatively low (Figures 3b, 3e, and 3h). For the 10°S–10°N band, observational uncertainty makes it difficult to determine if stochastic physics improves the simulated frequency of heavy rain rates, but it is more likely than not improving the precipitation rate distribution in EC-Earth (Figure 2b). In both regions, stochastic physics in EC-Earth and the UM reduces the frequency of more moderate rain rates and increases the frequency of light rain rates as well. In EC-Earth this reduces the positive bias in the simulated frequency of moderate rain rates between ∼5 and 15 mm/d in both regions, by ∼10% for the 10°S–10°N band (Figures 2b, 2c, 3b, and 3c).
2. The persistence of daily mean rainfall is reduced, except in the case of SPPT in CAM4 (Figure 4). This makes the simulated autocorrelation agree better with observational estimates. The improvement is particularly substantial in EC-Earth, with the bias in the 1 day lag autocorrelation being reduced by about 50% with respect to the GPCP estimate.
3. Variability is increased most at high frequencies above 0.3 per day, by up to about a factor of 2 in EC-Earth and the UM (Figures 5j and 5k), by up to ∼50% by SKEBS in CAM4 (Figure 5l) and by up to ∼20% by SPPT in CAM4 (not shown). All of the low-resolution deterministic models have less than half the power in both TRMM and GPCP in at least some wave numbers and frequencies (Figures 5a–5f), and the stochastic physics schemes reduce the deficit of spectral power at high frequencies.

Some impacts are less robust. For example, stochastic physics increases the temporal standard deviation of precipitation in most tropical regions in EC-Earth (Figure 6j), but it decreases the standard deviation in many tropical regions in the UM and CAM4 (Figures 6k and 6l). Its effect on the mean precipitation also depends strongly on the model and stochastic scheme being used (section 3.1 and Figure 1).

Overall, the stochastic schemes used in the models evaluated here give substantial improvements in some biases in tropical rainfall variability in all of the models without causing severe deteriorations in any of the

diagnostics we have used, when considering the tropics as a whole. This provides evidence that these stochastic schemes are performing reasonably at representing the effects of unresolved tropical variability and may also be useful for increasing the realism of tropical rainfall variability statistics in other climate models. Improvements given by more sophisticated stochastic schemes that are currently in development, mentioned in section 1, can be benchmarked against the improvements shown here. The results also indicate that variability that is unresolved in these models is important for determining the tropical rainfall statistics listed above.

Many of the impacts of stochastic physics differ in size between the models considered here, despite the stochastic physics schemes in the different models being quite similar. The impacts on the precipitation rate distributions, autocorrelation, and spectral power are largest in EC-Earth, with the impacts being considerably smaller in the UM and CAM4. This could be due to the impacts depending on the deterministic model structure or to differences between the details of the stochastic physics schemes used in each model. In order to tell how much difference each of these factors makes, it would be useful to test the impact of each scheme in one model or the impact of using exactly the same scheme in several models.

Since SPPT and SKEBS do not directly perturb precipitation, their effect on rainfall variability must come indirectly through their effect on grid box-scale variables. For example, if SPPT scales up the temperature and moisture tendencies given by the convection scheme, then the atmosphere will become more stable than it would have done, possibly resulting in less intense convection and precipitation at the next time step. If SPPT increased a drying tendency from the parameterization schemes, convection may similarly be subsequently inhibited. Conversely, when SPPT has the opposite effect, more intense precipitation may follow. This would have knock-on effects at later times and in other locations, making the overall impact of SPPT complex. SKEBS perturbations must affect rainfall by first affecting the large-scale state, which then affects the parameterizations' output.

The impacts of stochastic physics in EC-Earth seem quite consistent with the predictions of the No Feedbacks conceptual model, where stochastic physics is considered to only randomly perturb the resolved variables from which the precipitation is diagnosed, without giving rise to important feedbacks by otherwise changing the statistics of resolved variables. This does not necessarily mean that feedbacks do not still have an important role, however. The conceptual model is not consistent with all of the impacts of stochastic physics in the UM and CAM4, particularly the reduction in the precipitation standard deviation in some places (Figures 6k and 6l), showing that feedbacks due to changes in resolved variables are important in those models.

Stochastic physics produces some improvements that are not attainable by increasing horizontal resolution to ~16 km in EC-Earth and ~25–40 km in the UM (without changing the parameterizations). Using such high horizontal resolutions makes little improvement to the simulated precipitation rate distribution in northern South America (Figures 3b and 3e) or to the autocorrelation of precipitation (Figure 4), and it fails to increase spectral power at high frequencies (not shown). This provides evidence that stochastic physics can provide improvements in climate simulations that are not attainable simply by increasing the resolution this far and that unpredictable variability on smaller scales than these has an important influence on variability at larger scales, with resolutions as high as these still not likely to be used routinely for global climate modeling in the near future. This goes beyond previous results, relating to simulating midlatitude dynamics, suggesting that stochastic physics could give improvements that are only similar to those obtained by increasing horizontal resolution, rather than better [*Dawson and Palmer*, 2015]. This may be because the realism of parameterized convection is much more important in the tropics than in middle latitudes. This suggests that high-resolution regional models used in studies of changing tropical extremes [e.g., *Marengo et al.*, 2009; *Vizy and Cook*, 2012; *Crétat et al.*, 2014] could benefit from using stochastic parameterizations.

We also showed that the use of stochastic physics affects the simulated change of the frequency of extreme precipitation events in a future climate change scenario using the coupled atmosphere-ocean configurations of EC-Earth (section 3.6 and Figure 7). This suggests that a climate model's representation of unpredictable tropical variability will affect its projections of tropical climate change and that using stochastic parameterizations, based on a process understanding of observations and high-resolution models, could give a better simulation of the response to a forcing than using deterministic parameterizations. Uncertainty in how to best construct stochastic schemes should also be accounted for when assessing uncertainty in these projections.

The stochastic physics schemes improve certain biases in the deterministic model configurations but do not eliminate them in all models, such as there being too few heavy precipitation events in northern South America (Figures 3b, 3e, and 3h), the persistence of precipitation being too great (particularly in the UM; Figure 4), the spectral power at high frequencies being too low overall (Figures 5g–5i) and the standard deviation of tropical precipitation being too low in EC-Earth (Figures 6g). A possible reason for this is that SPPT and SKEBS are not fully representing the unpredictable component of tendencies in tropical convection, so the variance of this component is not large enough. This was also indicated by *Rodwell et al.* [2013, 2016] who found evidence using numerical weather prediction diagnostics that ECMWF ensemble weather forecasts, which use SPPT and SKEBS, do not have a large enough spread in certain convective situations, including in parts of northern South America.

4.1. Directions for Future Work

Other representations of subgrid-scale variability have been found to be more effective than the schemes tested here at increasing certain aspects of tropical variability. For example, superparameterization and the convection scheme of *Plant and Craig* [2008] were tested in version 5 of CAM (CAM5) by *Kooperman et al.* [2016] and *Wang et al.* [2016], respectively, and produced much larger increases in the frequency of heavy rainfall. *Goswami et al.* [2016] and *Peters et al.* [2017], respectively, show that variants of the stochastic multicloud model scheme increase the variance and decrease the autocorrelation of tropical precipitation in different GCMs by more than we found here. One speculative reason for this is that these parameterizations have been designed to represent the unpredictable component of variability at individual grid points, whereas SPPT and SKEBS were designed to insert perturbations that would grow rapidly to large scales, which is important for producing sufficient ensemble spread in weather forecasts. This could be a reason that SPPT and SKEBS do not wholly represent the unpredictable component of the atmospheric evolution tendency. Including additional representations of unpredictable variability on small scales may therefore be beneficial for simulating tropical variability in climate models. It would be interesting to do a direct comparison between the impacts of the schemes tested here and of those evaluated by *Kooperman et al.* [2016] and *Wang et al.* [2016]. Of course, improving the deterministic parts of parameterization schemes [e.g., *Jakob*, 2010] is also likely to be important for improving the representation of precipitation variability.

Stochastic physics was also not found to have much impact on the fraction of days that are dry in any model (section 3.2). Superparameterization did improve this statistic over land in CAM5 [*Kooperman et al.*, 2016]. It could be interesting to test if improved stochastic schemes could intermittently suppress convection to give better agreement with observations, for example, by introducing a stochastic convection trigger that could represent possible unresolved factors that unpredictably suppress convection.

Gaining a better understanding of the mechanisms by which stochasticity affects the rainfall variability is important for predicting how a given stochastic scheme will affect a model's simulations. A better understanding of the impact of SPPT could be found by seeing if perturbing particular variables on their own, such as specific humidity or temperature, can reproduce the effect of perturbing all variables together. Systematic testing in different models of the sensitivity to the correlation length and time scales of the random number patterns used would be very helpful. Applying the stochastic schemes in a single-column model could help to show which impacts depend on feedbacks from changes in resolved variables. Changing parameters of the deterministic schemes may help to reveal how the impacts depend on the climate of the deterministic configurations. Diagnostics of upscale growth of stochastic perturbations [e.g., *Selz and Craig*, 2015] may help with understanding how they affect the large-scale climatological state. Potential vorticity and diabatic tracer schemes [e.g., *Martínez-Alvarado et al.*, 2016] may be helpful for studying how stochastic perturbations affect particular processes relevant to tropical precipitation, such as the organization of deep convection.

Another topic for further analysis is understanding the impact stochastic physics has on the spatial structure of tropical precipitation. Visual inspection of snapshots of the tropical precipitation in the horizontal plane and Hovmöller diagrams in the different models indicates that stochastic physics is not strongly affecting the precipitation's spatial structure (not shown). *Peters et al.* [2017] showed a large impact of a stochastic multicloud model on the spatial structure, suggesting that SPPT and SKEBS are not able to reproduce this effect. A more detailed analysis would be useful.

A further implication of our results is that using stochastic physics may affect the simulation of wave fluxes into the tropical stratosphere and hence the simulation of the stratospheric quasi-biennial oscillation.

The influence of the tropics on midlatitude variability could also be affected if poleward wave fluxes into the extratropics are changed [*Branstator*, 2014]. These could form topics for future investigations.

Appendix A: Comparison of GPCP and TRMM Tropical Rainfall Estimates

Here the consistency of rainfall estimates in GPCP and TRMM are summarized in order to help to quantify observational uncertainty of rainfall statistics. Figure A1a shows the correlation between GPCP and TRMM daily rainfall in the tropics interpolated conservatively to a 2.5° × 2.5° grid. The correlation exceeds 0.8 in most places, the exceptions being located mostly in the central and eastern Pacific. The correlation is highest in northern South America, where it exceeds 0.9 at most points. This shows that GPCP and TRMM generally agree on which days and at which locations there is rainfall and on its relative intensity.

Figure A1b shows the root-mean-square (RMS) difference between GPCP and TRMM divided by the averages of their time means at each tropical location. The values exceed 1 in many oceanic locations, indicating that the differences between GPCP and TRMM daily rainfall estimates are large compared to their mean rainfall. The values over land are smaller, indicating better agreement between the estimates, possibly because calibration against rain gauge data is possible. The largest contiguous region with relatively small values of this quantity is northern South America, where RMS differences are 40–70% of the mean rainfall.

Figure A1c shows GPCP and TRMM time series of rainfall at a location in the tropical West Pacific that illustrate the reasons that the time series have a high correlation and also a large RMS difference. The time series agree well on the timing of rainfall, but TRMM generally estimates larger rainfall amounts when there is heavy

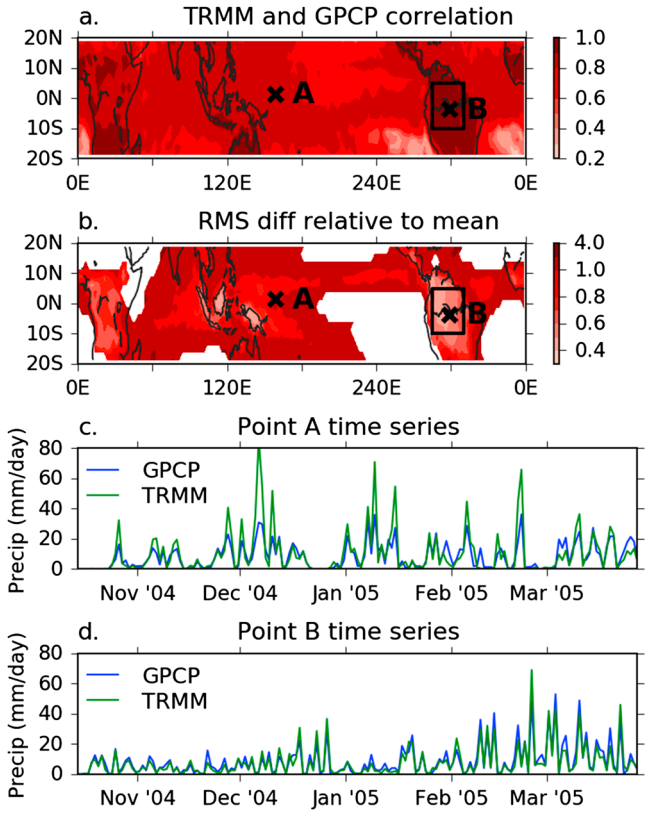

Figure A1. A comparison of GPCP and TRMM daily mean tropical rainfall data. (a) The correlation between GPCP and TRMM. (b) The root-mean-square difference between GPCP and TRMM divided by the averages of their time means at each grid point. Data are only shown at points where the average time mean rainfall in GPCP and TRMM exceeds 1.5 mm/d. The rectangles in Figures A1a and A1b show the northern South America region used in sections 3.2 and 3.6. (c) Time series of GPCP and TRMM rainfall at (158.75°E, 1.25°N), marked as cross "A" in Figures A1a and A1b. (d) Similar time series at (61.25°W, 3.75°S), marked as cross "B." The rainfall data have been conservatively interpolated to a 2.5° × 2.5° grid. The data sets are well correlated, but TRMM typically has greater extremes over ocean points, with the best agreement being over northern South America.

rainfall. Figure A1d shows similar time series at a point in northern South America, where the data sets agree better, consistent with the smaller RMS differences shown here in Figure A1b.

Acknowledgments
We thank D. Bundy for helping to produce the CAM4 simulations, A. Subramanian for stimulating discussions, and three anonymous reviewers for their constructive feedback. P.W. was supported by European Research Council grant 291406. The EC-Earth simulations were performed in the framework of the PRACE project Climate-SPHINX: S.C., P.D., and J.v.H. acknowledge PRACE for awarding access to the SuperMUC resource based in Germany at LRZ. S.C. and J.v.H. also acknowledge funding received from the PRIMAVERA project (European Commission under grant agreement 641727 of the Horizon 2020 research program). J.v.H. also acknowledges funding from the European Union's Horizon 2020 research and innovation program under grant agreement 641816 (CRESCENDO) and by the Italian project of Interest "NextData" of the Italian Ministry for Education, University and Research. Additionally, P.D. was supported by funding from the European Union's Horizon 2020 research and innovation program COGNAC under the European Union Marie Skłodowska-Curie grant agreement 654942. C.S. was supported by the Joint UK BEIS/Defra Met Office Hadley Centre Climate Programme (GA01101). A.W. was supported by the EU FP7 funded project SPECS, grant agreement 308378. The code and data used to perform the data analysis are available from Peter Watson (peter.watson@physics.ox.ac.uk) upon request.

References

Adler, R. F., et al. (2003), The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present), *J. Hydrometeorol.*, *4*(6), 1147–1167, doi:10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.

Bengtsson, L., M. Steinheimer, P. Bechtold, and J. Geleyn (2013), A stochastic parametrization for deep convection using cellular automata, *Q. J. R. Meteorol. Soc.*, *139*(675), 1533–1543.

Berner, J., F. J. Doblas-Reyes, T. N. Palmer, G. Shutts, and A. Weisheimer (2008), Impact of a quasi-stochastic cellular automaton backscatter scheme on the systematic error and seasonal prediction skill of a global climate model, *Philos. Trans. A. Math. Phys. Eng. Sci.*, *366*, 2561–2579, doi:10.1098/rsta.2008.0033.

Berner, J., G. J. Shutts, M. Leutbecher, and T. N. Palmer (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, doi:10.1175/2008JAS2677.1.

Berner, J., K. R. Fossell, S.-Y. Ha, J. P. Hacker, and C. Snyder (2015), Increasing the skill of probabilistic forecasts: Understanding performance improvements from model-error representations, *Mon. Weather Rev.*, *143*(4), 1295–1320, doi:10.1175/MWR-D-14-00091.1.

Berner, J., et al. (2017), Stochastic parameterization: Toward a new view of weather and climate models, *Bull. Am. Meteorol. Soc.*, *98*, 565–588, doi:10.1175/BAMS-D-15-00268.1.

Bouttier, F., B. Vié, O. Nuissier, and L. Raynaud (2012), Impact of stochastic physics in a convection-permitting ensemble, *Mon. Weather Rev.*, *140*(11), 3706–3721, doi:10.1175/MWR-D-12-00031.1.

Branstator, G. (2014), Long-lived response of the midlatitude circulation and storm tracks to pulses of tropical heating, *J. Clim.*, *27*(23), 8809–8826, doi:10.1175/JCLI-D-14-00312.1.

Buizza, R., M. Miller, and T. Palmer (1999), Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System, *Q. J. R. Meteorol. Soc.*, *125*, 2887–2908.

Christensen, H. M., J. Berner, D. R. B. Coleman, and T. N. Palmer (2017), Stochastic parameterisation and the El Niño–Southern Oscillation, *J. Clim.*, *30*, 17–38, doi:10.1175/JCLI-D-16-0122.1.

Cohen, B. G., and G. C. Craig (2006), Fluctuations in an equilibrium convective ensemble. Part II: Numerical experiments, *J. Atmos. Sci.*, *63*(8), 2005–2015, doi:10.1175/JAS3710.1.

Crétat, J., E. K. Vizy, and K. H. Cook (2014), How well are daily intense rainfall events captured by current climate models over Africa?, *Clim. Dyn.*, *42*(9–10), 2691–2711, doi:10.1007/s00382-013-1796-7.

Davies, L., C. Jakob, P. May, V. V. Kumar, and S. Xie (2013), Relationships between the large-scale atmosphere and the small-scale convective state for Darwin, Australia, *J. Geophys. Res.*, *118*, 11,534–11,545, doi:10.1002/jgrd.50645.

Davini, P., J. von Hardenberg, S. Corti, H. M. Christensen, S. Juricke, A. Subramanian, P. A. G. Watson, A. Weisheimer, and T. N. Palmer (2017), Climate SPHINX: Evaluating the impact of resolution and stochastic physics parameterisations in the EC-Earth global climate model, *Geosci. Model Dev. Discuss.*, *10*, 1383–1402, doi:10.5194/gmd-2016-115.

Dawson, A., and T. N. Palmer (2015), Simulating weather regimes: Impact of model resolution and stochastic parameterization, *Clim. Dyn.*, *44*, 2177–2193, doi:10.1007/s00382-014-2238-x.

Dorrestijn, J., D. T. Crommelin, A. P. Siebesma, H. J. J. Jonker, F. Selten, J. Dorrestijn, D. T. Crommelin, A. P. Siebesma, H. J. J. Jonker, and F. Selten (2016), Stochastic convection parameterization with Markov Chains in an intermediate-complexity GCM, *J. Atmos. Sci.*, *73*(3), 1367–1382, doi:10.1175/JAS-D-15-0244.1.

Efron, B., and R. J. Tibshirani (1994), *An Introduction to the Bootstrap*, CRC Press, Boca Raton, Fla.

Frenkel, Y., A. J. Majda, and B. Khouider (2012), Using the stochastic multicloud model to improve tropical convective parameterization: A paradigm example, *J. Atmos. Sci.*, *69*(3), 1080–1105, doi:10.1175/JAS-D-11-0148.1.

Gent, P. R., et al. (2011), The Community Climate System Model version 4, *J. Clim.*, *24*(19), 4973–4991, doi:10.1175/2011JCLI4083.1.

Goswami, B. B., B. Khouider, R. Phani, P. Mukhopadhyay, and A. Majda (2016), Improving synoptic and intra-seasonal variability in CFSv2 via stochastic representation of organized convection, *Geophys. Res. Lett.*, *43*, 1104–1113, doi:10.1002/2016GL071542.

Gregory, D., and P. R. Rowntree (1990), A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure, *Mon. Weather Rev.*, *118*(7), 1483–1506.

Huffman, G. J., R. F. Adler, P. Arkin, A. Chang, R. Ferraro, A. Gruber, J. Janowiak, A. McNab, B. Rudolf, and U. Schneider (1997), The Global Precipitation Climatology Project (GPCP) combined precipitation dataset, *Bull. Am. Meteorol. Soc.*, *78*(1), 5–20, doi:10.1175/1520-0477(1997)078<0005:TGPCPG>2.0.CO;2.

Huffman, G. J., R. F. Adler, M. M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind (2001), Global precipitation at one-degree daily resolution from multisatellite observations, *J. Hydrometeorol.*, *2*(1), 36–50, doi:10.1175/1525-7541(2001)002<0036:GPAODD>2.0.CO;2.

Huffman, G. J., D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, and E. F. Stocker (2007), The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, *J. Hydrometeorol.*, *8*(1), 38–55, doi:10.1175/JHM560.1.

Hung, M. P., J. L. Lin, W. Wang, D. Kim, T. Shinoda, and S. J. Weaver (2013), MJO and convectively coupled equatorial waves simulated by CMIP5 climate models, *J. Clim.*, *26*(17), 6185–6214, doi:10.1175/JCLI-D-12-00541.1.

Hurrell, J. W., J. J. Hack, D. Shea, J. M. Caron, and J. Rosinski (2008), A new sea surface temperature and sea ice boundary dataset for the community atmosphere model, *J. Clim.*, *21*(19), 5145–5153, doi:10.1175/2008JCLI2292.1.

Intergovernmental Panel on Climate Change (2014), *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 688 pp., Cambridge Univ. Press, Cambridge, U. K., and New York.

Jakob, C. (2010), Accelerating progress in global atmospheric model development through improved parameterizations: Challenges, opportunities, and strategies, *Bull. Am. Meteorol. Soc.*, *91*(7), 869–875, doi:10.1175/2009BAMS2898.1.

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

Klingaman, N. P., G. M. Martin, and A. Moise (2017), ASoP (v1.0): A set of methods for analyzing scales of precipitation in general circulation models, *Geosci. Model Dev.*, *10*, 57–83, doi:10.5194/gmd-2016-161.

Kober, K., A. M. Foerster, and G. C. Craig (2015), Examination of a stochastic and deterministic convection parameterization in the COSMO model, *Mon. Weather Rev.*, *143*, 4088–4103, doi:10.1175/MWR-D-15-0012.1.

Kooperman, G., M. Pritchard, M. Burt, M. Branson, and D. A. Randall (2016), Robust effects of cloud superparameterization on simulated daily rainfall intensity statistics across multiple versions of the Community Earth System Model, *J. Adv. Model. Earth Syst.*, *8*, 140–165, doi:10.1002/2015MS000574.

Lin, J. L., et al. (2006), Tropical intraseasonal variability in 14 IPCC AR4 climate models. Part I: Convective signals, *J. Clim.*, *19*(12), 2665–2690, doi:10.1175/JCLI3735.1.

Lin, J. W.-B., and J. Neelin (2003), Toward stochastic deep convective parameterization in general circulation models, *Geophys. Res. Lett.*, *30*(4), 1162, doi:10.1029/2002GL016203.

Lin, J. W. B., and J. D. Neelin (2000), Influence of a stochastic moist convective parameterization on tropical climate variability, *Geophys. Res. Lett.*, *27*, 3691–3694, doi:10.1029/2000GL011964.

Lin, J. W.-B., and J. D. Neelin (2002), Considerations for stochastic convective parameterization, *J. Atmos. Sci.*, *59*(5), 959–975, doi:10.1175/1520-0469(2002)059<0959:CFSCP>2.0.CO;2.

Madec, G., (2008), NEMO ocean engine, Tech. Rep., Institut Pierre-Simon Laplace (IPSL), France.

Marengo, J. A., R. Jones, L. M. Alves, and M. C. Valverde (2009), Future change of temperature and precipitation extremes in South America as derived from the PRECIS regional climate modeling system, *Int. J. Climatol.*, *29*(15), 2241–2255, doi:10.1002/joc.1863.

Martínez-Alvarado, O., S. L. Gray, and J. Methven (2016), Diabatic processes and the evolution of two contrasting summer extratropical cyclones, *Mon. Weather Rev.*, *144*(9), 3251–3276, doi:10.1175/MWR-D-15-0395.1.

Molteni, F., T. Stockdale, M. Balmaseda, G. Balsamo, R. Buizza, L. Ferranti, L. Magnusson, K. Mogensen, T. Palmer, and F. Vitart (2011), The new ECMWF seasonal forecast system (System 4), *ECMWF Tech. Memo. 656*, European Centre for Medium-Range Weather Forecasts, Reading, U. K.

Moss, R. H., et al. (2010), The next generation of scenarios for climate change research and assessment, *Nature*, *463*(7282), 747–756.

Ollinaho, P., S.-J. Lock, M. Leutbecher, P. Bechtold, A. Beljaars, R. M. Forbes, T. Haiden, R. J. Hogan, and I. Sandu (2017), Towards process-level representation of model uncertainties: Stochastically perturbed parametrisations in the ECMWF ensemble, *Q. J. R. Meteorol. Soc.*, *143*, 408–422.

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

Palmer, T., R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G. Shutts, M. Steinheimer, and A. Weisheimer (2009), Stochastic parametrization and model uncertainty, *ECMWF Tech. Memo. 598*, European Centre for Medium-Range Weather Forecasts, Reading, U. K.

Palmer, T. N. (2012), Towards the probabilistic Earth-system simulator: A vision for the future of climate and weather prediction, *Q. J. R. Meteorol. Soc.*, *138*, 841–861, doi:10.1002/qj.1923.

Peters, K., C. Jakob, L. Davies, B. Khouider, and A. J. Majda (2013), Stochastic behavior of tropical convection in observations and a multicloud model, *J. Atmos. Sci.*, *70*, 3556–3575, doi:10.1175/JAS-D-13-031.1.

Peters, K., T. Crueger, C. Jakob, and B. Möbis (2017), Improved MJO-simulation in ECHAM6.3 by coupling a Stochastic Multicloud Model to the convection scheme, *J. Adv. Model. Earth Syst.*, *9*, 193–219, doi:10.1002/2016MS000809.

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

Qian, T., A. Dai, K. E. Trenberth, and K. W. Oleson (2006), Simulation of global land surface conditions from 1948 to 2004. Part I: Forcing data and evaluations, *J. Hydrometeorol.*, *7*(5), 953–975, doi:10.1175/JHM540.1.

Reynolds, C. A., J. G. McLay, J. S. Goerss, E. A. Serra, D. Hodyss, and C. R. Sampson (2011), Impact of resolution and design on the U.S. Navy global ensemble performance in the tropics, *Mon. Weather Rev.*, *139*(7), 2145–2155, doi:10.1175/2011MWR3546.1.

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax (2007), Daily high-resolution-blended analyses for sea surface temperature, *J. Clim.*, *20*(22), 5473–5496, doi:10.1175/2007JCLI1824.1.

Rochetin, N., J.-Y. Grandpeix, C. Rio, and F. Couvreux (2014), Deep convection triggering by boundary layer thermals. Part II: Stochastic triggering parameterization for the LMDZ GCM, *J. Atmos. Sci.*, *71*(2), 515–538, doi:10.1175/JAS-D-12-0337.1.

Rodwell, M. J., et al. (2013), Characteristics of occasional poor medium-range weather forecasts for Europe, *Bull. Am. Meteorol. Soc.*, *94*, 1393–1405, doi:10.1175/BAMS-D-12-00099.1.

Rodwell, M. J., S. T. K. Lang, N. B. Ingleby, N. Bormann, E. Holm, F. Rabier, D. Richardson, and M. Yamaguchi (2016), Reliability in ensemble data assimilation, *Q. J. R. Meteorol. Soc.*, *142*, 443–454, doi:10.1002/qj.2663.

Saeed, F., A. Haensler, T. Weber, S. Hagemann, and D. Jacob (2013), Representation of extreme precipitation events leading to opposite climate change signals over the Congo basin, *Atmosphere*, *4*(3), 254–271, doi:10.3390/atmos4030254.

Sakradzija, M., A. Seifert, and A. Dipankar (2016), A stochastic scale-aware parameterization of shallow cumulus convection across the convective gray zone, *J. Adv. Model. Earth Syst.*, *8*(2), 786–812, doi:10.1002/2016MS000634.

Sanchez, C., K. Williams, and M. Collins (2016), Improved stochastic physics schemes for global weather and climate models, *Q. J. R. Meteorol. Soc.*, *142*, 147–159.

Sardeshmukh, P., C. Penland, and M. Newman (2003), Drifts induced by multiplicative red noise with application to climate, *Europhys. Lett.*, *63*(4), 498–504.

Selz, T., and G. C. Craig (2015), Simulation of upscale error growth with a stochastic convection scheme, *Geophys. Res. Lett.*, *42*, 3056–3062, doi:10.1002/2015GL063525.

Shutts, G. (2005), A kinetic energy backscatter algorithm for use in ensemble prediction systems, *Q. J. R. Meteorol. Soc.*, *131*(612), 3079–3102, doi:10.1256/qj.04.106.

Shutts, G. (2015), A stochastic convective backscatter scheme for use in ensemble prediction systems, *Q. J. R. Meteorol. Soc.*, *141*(692), 2602–2616, doi:10.1002/qj.2547.

Stephens, G., T. L'Ecuyer, R. Forbes, A. Gettlemen, J.-C. Golaz, A. Bodas-Salcedo, K. Suzuki, P. Gabriel, and J. Haynes (2010), Dreary state of precipitation in global models, *J. Geophys. Res.*, *115*, D24211, doi:10.1029/2010JD014532.

Subramanian, A., A. Weisheimer, T. Palmer, F. Vitart, and P. Bechtold (2017), Impact of stochastic physics on tropical precipitation in the coupled ECMWF model, *Q. J. R. Meteorol. Soc.*, *143*, 852–865.

Sušelj, K., T. F. Hogan, and J. Teixeira (2014), Implementation of a stochastic eddy-diffusivity/mass-flux parameterization into the Navy Global Environmental Model, *Weather Forecasting*, *29*(6), 1374–1390, doi:10.1175/WAF-D-14-00043.1.

Tiedtke, M. (1989), A comprehensive mass flux scheme for cumulus parameterization in large-scale models, *Mon. Weather Rev.*, *117*(8), 1779–1800, doi:10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2.

Titchner, H. A., and N. A. Rayner (2014), The Met Office Hadley Centre sea ice and sea surface temperature data set, version 2: 1. Sea ice concentrations, *J. Geophys. Res. Atmos.*, *119*, 2864–2889, doi:10.1002/2013JD020316.

Vancoppenolle, M., S. Bouillon, T. Fichefet, H. Goosse, O. Lecomte, M. A. Morales Maqueda, and G. Madec (2012), The Louvain-la-Neuve sea ice model, *Notes du Pole Modélisation*, *31*, 1–85.

Vizy, E. K., and K. H. Cook (2012), Mid-twenty-first-century changes in extreme events over northern and tropical Africa, *J. Clim.*, *25*(17), 5748–5767, doi:10.1175/JCLI-D-11-00693.1.

Walters, D., et al. (2017), The met office unified model global atmosphere 6.0/6.1 and JULES global land 6.0/6.1 configurations, *Geosci. Model Dev.*, *10*(4), 1487–1520, doi:10.5194/gmd-10-1487-2017.

Wang, Y., G. J. Zhang, and G. C. Craig (2016), Stochastic convective parameterization improves the simulation of tropical precipitation variability in the NCAR CAM5, *Geophys. Res. Lett.*, *43*, 6612–6619, doi:10.1002/grl.54561.

Ward, P. J., et al. (2015), Usefulness and limitations of global flood risk models, *Nat. Clim. Change*, *5*(8), 712–715.

Watson, P. A. G., H. M. Christensen, and T. N. Palmer (2015), Does the ECMWF IFS convection parameterization with stochastic physics correctly reproduce relationships between convection and the large-scale state?, *J. Atmos. Sci.*, *72*, 236–242, doi:10.1175/JAS-D-14-0252.1.

Weisheimer, A., S. Corti, T. Palmer, and F. Vitart (2014), Addressing model error through atmospheric stochastic physical parametrizations: impact on the coupled ECMWF seasonal forecasting system, *Philos. Trans. R. Soc. A*, *372*, 20130290.

Westra, S., H. J. Fowler, J. P. Evans, L. V. Alexander, P. Berg, F. Johnson, E. J. Kendon, G. Lenderink, and N. M. Roberts (2014), Future changes to the intensity and frequency of short-duration extreme rainfall, *Rev. Geophys.*, *52*, 522–555, doi:10.1002/2014RG000464.

Wheeler, M., and G. N. Kiladis (1999), Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber-frequency domain, *J. Atmos. Sci.*, *56*(3), 374–399, doi:10.1175/1520-0469(1999)056<0374:CCEWAO>2.0.CO;2.

Williams, P. D. (2012), Climatic impacts of stochastic fluctuations in air-sea fluxes, *Geophys. Res. Lett.*, *39*, L10705, doi:10.1029/2012GL051813.

Yonehara, H., and M. Ujiie (2011), A stochastic physics scheme for model uncertainties in the JMA one-week ensemble prediction system, *CAS/JSC WGNE Res. Act. Atmos. Oceanic Model.*, *41*(6), 9–10.

Zhang, G. J., and N. A. McFarlane (1995), Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian Climate Centre general circulation model, *Atmos. Ocean*, *33*(3), 407–446.