

Considerations for Stochastic Convective Parameterization

JOHNNY WEI-BING LIN*

Department of Atmospheric Sciences, University of California, Los Angeles, Los Angeles, California

J. DAVID NEELIN

Department of Atmospheric Sciences and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, Los Angeles, California

(Manuscript received 28 March 2001, in final form 23 July 2001)

ABSTRACT

Convective parameterizations in general circulation models (GCMs) generally only aim to simulate the mean or first-order moment of convection; higher moments associated with subgrid variability are not explicitly considered. In this study, an empirically based stochastic convective parameterization is developed that uses an assumed mixed lognormal distribution of rainfall, tuned with parameter values derived from observations, to control selected nonmean statistical properties of convection. Testing of this stochastic convective parameterization reveals that large-scale model dynamics interacts heavily with the convective parameterization, in ways such that the resulting output is fundamentally different from the input. This suggests stochastic parameterizations cannot be calibrated outside of a model's dynamical framework. Implications are discussed for the relative merits of the empirical approach versus another approach that introduces the stochastic process within the framework of the convective parameterization. Inclusion of the variance arising from unresolved scales by stochastic parameterization of convection is found to have a substantial impact upon atmospheric variability in the Tropics, including intraseasonal and longer timescales.

1. Introduction

In the atmosphere, it is reasonable to hypothesize that for a given large-scale temperature and moisture field, there is a contribution to the variability of convection that arises inherently from small-scale motions, but which are not well represented by large ensemble means. Though this process has relatively short correlation scales in both time and space, it can act as a noise forcing that potentially shows up at the large scales. Very little is understood, however, about the extent to which this hypothesis may be true. This present study seeks to gain some insight into the variability in convection that is not captured by ensemble means.

Because general circulation models (GCMs) have finite grid sizes, and thus must, in general, deterministically parameterize subgrid phenomena in terms of grid-level variables, these models implicitly assume that the unresolved phenomena can be considered in an en-

semble mean sense. As, on the whole, GCMs simulate the climatology of precipitation relatively well, this assumption seems justified. Work using a cloud resolving model also suggests that the assumption that subgrid convection can be considered in an ensemble mean sense is reasonably accurate (Xu et al. 1992). They find that the mean effects of convection closely follow the large-scale field, which is consistent with the notion that convection is fundamentally parameterizable.

At the same time, however, Xu et al. also find a substantial amount of convective "scatter" about the mean that is not determined by the large-scale environment. Such "scatter" in fields crucial to the simulation of convection, such as vertical motion, can have a standard deviation on the order of the mean values (Katzfey and Ryan 2000).

Simulations by GCMs of some of the other statistical features of precipitation besides the mean have been lacking. For example, through the use of a 3-hourly outgoing longwave radiation (OLR) measurement-based proxy for deep convective activity for winter 1984, Ricciardulli and Garcia (2000) find that the variance of tropical precipitation is substantial, and that the vast majority of the variance is contained at high frequencies (periods between 6 h and 2 days). Comparing this with a National Center for Atmospheric Research

* Current affiliation: Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado.

Corresponding author address: Dr. J. David Neelin, Department of Atmospheric Sciences, University of California, Los Angeles, 7127 Math Sciences, Los Angeles, CA 90095-1565.
E-mail: neelin@atmos.ucla.edu

(NCAR) Community Climate Model (CCM3) run with the Zhang–McFarlane convective parameterization (Zhang and McFarlane 1995), Ricciardulli and Garcia show the modeled variance of vertically integrated tropical deep convective activity is much less than seen in observations, and of the variance that is produced, the majority is *not* contained in high frequencies (except in the Amazon basin). However, the CCM3 simulation of mean convection is quite similar to observations. Ricciardulli and Garcia also find that the variance found in OLR observations is distributed broadly even down to small scales (zonal wavenumber $k > 60$), and that only 40% of the variance is found at large scales ($k < 15$).

Over the tropical ocean, the majority of daily precipitation values are of low intensity, with approximately 60% of all days having zero precipitation. A global precipitation frequency distribution (Lau et al. 1996) shows similar characteristics. GCMs, however, have a difficult time simulating the frequency of occurrence of low monthly precipitation (0–1 mm day⁻¹) rates; every one of the Atmospheric Model Intercomparison Project (AMIP) GCMs underestimates this component (Lau et al. 1996).

If the physical hypothesis described earlier is true (i.e., there exist variations in convection that may influence large-scale climate, but which are not well described by the ensemble mean), then current GCM parameterizations (which follow the large-scale and which appear unable to simulate variance associated with unresolved processes), may need to be improved to include such phenomena. (In a modeling framework, the ensemble mean quantities roughly correspond to resolved phenomena, while the physical processes not well described by the ensemble means correspond to unresolved variability.) Admittedly, because of the ways convective parameterizations are implemented in numerical models, in practice the parameterizations do include some of the variance usually associated with unresolved processes. For instance, numerical noise resulting from instantaneous adjustment [as in the Manabe et al. (1965) scheme], the use of “trigger functions,” and other nonlinearities, provide some such variance. This variance can be viewed as an accident or afterthought, although a “trigger function” approach could be viewed as a way of mimicking unresolved variance within a deterministic framework. Explicit representation of the unresolved higher moments of convection has been lacking.

In order to help quantify the importance of representing these deviations from the ensemble mean, and as an alternative to such a deterministic method of quantifying the importance of unresolved variance, we suggest that stochastic convective parameterizations can be developed that explicitly include the effects of the higher moments of convection. There is a long history of work related to representing parts of the climate system as stochastic processes (e.g., Leith 1975; Hasselmann 1976; Oerlemans 1979; Farrell and Ioannou 1994; Penland and Matrosova 1994; Eckert and Latif 1997; Blanke et al. 1997;

Kleeman and Moore 1997; Saravanan and McWilliams 1998; Majda et al. 1999; Timmermann and Lohmann 2000). In most cases, the stochastic process is not intended for use in GCM parameterizations, and in general the stochastic representation is not a “parameterization,” in the sense that the statistics are not closely linked to resolved processes. Buizza et al. (1999) model error in parameterized processes by a random multiplier sampled from a uniform distribution in the European Centre for Medium-Range Weather Forecasts Ensemble Prediction System, and find an increase in ensemble spread and improvement in probabilistic precipitation forecasts. However, Buizza et al. are not trying to represent a specific physical process, but rather uncertainty regarding parameterization processes in general. Yu and Neelin (1994) use a linearized primitive equation numerical model of the tropical atmosphere under moist convective adjustment to investigate the impact of white stochastic forcing in the temperature and ω equations, and find peak response at low frequencies and low wavenumbers. We aim to apply such stochastic methods to a parameterization operating in a nonlinear context.

We propose that methods of including unresolved variance through a stochastic parameterization may be grouped into two general approaches or classes:

- *Approach 1.* Stochastic processes introduced within the framework of the convective parameterization, informed by at least some of the physics that contribute to unresolved variance.
- *Approach 2.* Directly controlling the statistics of the overall convective heating by specifying a distribution as a function of model variables, with this dependence estimated empirically.

An important difference between the two is that in approach 1, the distribution (of such quantities as precipitation, days of zero precipitation, etc.) is not known in advance and is determined by interactions of the stochastic process with both the other elements of the convective parameterization and with the large-scale dynamics. In approach 2 there is an implicit assumption that the distribution, for example, of precipitation, is sufficiently independent of interactions with the large-scale such that it is reasonable to estimate outside the model framework. It is not clear in advance which approach is likely to be more fruitful, and so we endeavor to set up and test examples of each.

In approach 1, a variety of levels of modeling is foreseen. As an example of a highly complex version of a stochastic parameterization using approach 1, one could embed a randomly initiated simplified model of mesoscale systems [such as the archetypal mesoscale models proposed by Moncrieff (1981, 1992)] in a GCM convective parameterization. On the other end of the range of complexity, one could model buoyancy related variables as the sum of an ensemble mean component and a Gaussian distributed noise, within an existing parameterization.

Lin and Neelin (2000, hereinafter LN) develop such an elementary example of stochastic parameterization using approach 1 by adding a stochastic component to convective available potential energy (CAPE) related variables within an intermediate-level atmospheric model's existing Betts–Miller (1986) convective parameterization (in the present study this is referred to as the “CAPE scheme”). This implicitly assumes that the Betts–Miller scheme gives a plausible response of subdomains in a grid box to variable CAPE. The source of the stochastic component of convection is then modeled as arising from these subensembles, rectified by the nonlinearity of the convective elements. The random component is a first-order autoregressive process with an autocorrelation timescale τ_ξ , and is forced by Gaussian noise with mean zero and standard deviation σ_ξ .

Lin and Neelin (2000) find that results using this scheme depend strongly on τ_ξ , and that model simulations of total precipitation variance and the probability distribution function of precipitation improve with the use of the stochastic parameterization. Higher values of τ_ξ (20 min, 2 h, and 1 day cases were tested) yield results more closely matching observations. The addition of stochastic convection results in an increase in intraseasonal spectral power in low-wavenumber 850-hPa zonal wind anomaly.

In the present study, a first attempt at a stochastic convective parameterization using approach 2 is developed, where variance is added by tailoring the convective heating so that it reproduces certain statistical properties derived from observations. This type of empirically derived parameterization is similar in ways to the point process models (e.g., Eagleson 1978; Rodriguez-Iturbe et al. 1987) used in hydrology to represent temporal rainfall, while reproducing selected statistical properties. The choice of statistical properties is informed by work from the remote-sensing community, which is concerned with the extrapolation of total rainfall from satellite measurements (e.g., Kedem et al. 1990). Although the aim of the stochastic parameterization described in this study is to represent subgrid rainfall at the grid-scale level, (instead of extrapolating rainfall from grid-scale measurements), similar distributions and relationships would be expected to apply.

Beginning from such statistical approaches to representing rainfall and considering how they might be adapted to the problem of stochastic convective parameterization, we posit the following methodology (as an example of approach 2, which is to be tested):

- The distribution should have a mean value equal to the ensemble mean of convection generated by the traditional convective scheme.
- The shape of the distribution is based (at least in part) on observations.
- Parameters describing the distribution are expressed as functions of the ensemble mean of convection generated by the traditional convective scheme.
- The value of precipitation that is input into the model

dynamics is obtained by choosing randomly from the distribution.

It is also hypothesized that the statistical properties of most importance will be those that are poorly reproduced by current GCM convective parameterizations (e.g., frequency of zero precipitation, total variance). Thus, the distribution used in the scheme should aim to capture those properties. The particular implementation in this present study of a stochastic parameterization based on these principles is referred to as the “empirical lognormal scheme.” A description of other distributions that were tested in the course of this work and why we chose to use the lognormal distribution, is given at the end of section 3.

The assumptions one makes in formulating the empirical lognormal scheme may or may not work well. When evaluating a new parameterization methodology (such as described in approach 2), it can be instructive to choose a reasonable example of that methodology and work it all the way through, rather than focusing on the details of the parameter choices of the distribution. The validity of these assumptions and the usefulness of this approach can thus be tested. For such tests, an intermediate-level model is preferred to a full-scale GCM, due to computational expense.

In section 2, the atmospheric model used in this testing is described. The empirical lognormal scheme is described in detail in section 3. In section 5, the effects of the stochastic scheme on intraseasonal variability is described. Given that grid-scale precipitation data contains both small-scale and large-scale effects, one concern is how well a scheme at least partly based on such data will do in simulating convection; these results are presented in section 6. Sensitivities of various quantities (e.g., variance) are also examined. The conclusions focus on a critical examination of the validity of the posited assumptions that underlie this case study of a stochastic parameterization by approach 2. For instance, we are led to revise our presumption that a stochastic parameterization must preserve the input ensemble mean. We highlight the most substantial challenge to approach 2, which will have to be faced by any empirical parameterization: the assumption that the distribution can be determined without considering the interaction with model dynamics. A checklist of considerations to help guide future development is presented.

2. Model description

The numerical model used in this present study is version 2.1 of the Neelin–Zeng quasi-equilibrium tropical circulation model (QTCM1), a primitive equation-based atmospheric model that focuses on simulating the tropical atmosphere. For a full description of model formulation (for v2.0), see Neelin and Zeng (2000); Zeng et al. (2000) provide initial results of the model (using v2.1). The QTCM1 is more complex than a simple mod-

el, and includes full primitive equation nonlinearity and a radiative–convective feedback package. The model also has a simple land soil moisture routine, but does not include the effects of topography. Being simpler than a full-scale GCM, the QTCM1 is easier to diagnose and is computationally faster. For its convective parameterization, the QTCM1 uses the Betts–Miller moist convective adjustment scheme (Betts and Miller 1986), a scheme that is also used in some GCMs. The convective scheme is described more fully in section 3.

In the QTCM1, a Galerkin expansion is used to represent the vertical structure of atmospheric temperature, humidity, and velocity, in contrast with most full-scale GCMs, which use multiple homogeneous layers. As the vertical basis functions of the expansion are chosen using analytical solutions that assume convective quasi-equilibrium conditions, only a few need be retained. For temperature and humidity, the QTCM1 uses only a single vertical mode. For velocity, the QTCM1 uses a single baroclinic mode defined using the Galerkin expansion in temperature, as well as a separately defined barotropic velocity mode. Finite differencing is used in the horizontal with grid points every 5.625° in longitude and 3.75° in latitude.

The QTCM1 provides a reasonable simulation of tropical climatology and interannual variability (Zeng et al. 2000). An earlier version of the model was used to study Madden–Julian oscillation (MJO)-like tropical intraseasonal variability (Lin et al. 2000). The v2.1 release of the QTCM1 was also used in the stochastic convective parameterization study by LN.

The equation for the projection of tropospheric temperature onto the temperature basis function T_1 (in units that absorb heat capacity) has the form

$$\partial_t T_1 + [\text{L. S.}] = Q_c, \quad (1)$$

where [L. S.] denotes large-scale dynamics terms plus other parameterized quantities such as radiation and nonmoist convective turbulent heat transfer. Here Q_c is the vertically–averaged convective heating, which will be the object of our stochastic parameterization efforts. The vertical distribution of heating is determined implicitly in the QTCM1 framework and is unmodified in the current stochastic scheme. In the version of the Betts–Miller convective parameterization used by QTCM1, convective heating Q_c is given by Q_c^{BM} , where

$$Q_c^{\text{BM}} \propto \frac{1}{\tau_c} H(C_1) C_1. \quad (2)$$

The convective relaxation timescale τ_c has a value of 2 h, $H(C_1)$ is zero for $C_1 \leq 0$, and one for $C_1 > 0$, and C_1 is a measure of the CAPE, projected onto the model’s moisture and temperature basis functions, q_1 and T_1 . In the stochastic convective schemes, Q_c becomes a random variable whose distribution depends on the large-scale moisture and temperature related variables from the QTCM1 at any given time step.

3. Description of the empirical lognormal scheme

Being a highly nonlinear process, a full statistical characterization of precipitation requires consideration of a variety of measures. Which measures are chosen depends upon the purposes of the model. For instance, in a land surface water balance model (Eagleson 1978), the three important characteristics of rainfall whose statistics need to be properly simulated are time between storms, storm duration, and depth of water deposited by the storm. As the present study is an early attempt at studying the effects of the unresolved deviations from large-scale forced precipitation, it focuses on simulating two relatively gross statistical measures: precipitation variance and distribution of precipitation (including the percentage of time precipitation equals zero). It is hypothesized that much of the subgrid scale variance that shows itself at the grid scale will be high-frequency and low-magnitude. At the same time, one aims to preserve the mean precipitation simulated by the model.

In the empirical lognormal scheme, unlike the CAPE stochastic parameterization used in LN, the algorithm behind the Betts–Miller parameterization is assumed *not* to adequately represent the relationship between subgrid precipitation and the grid-scale variables but rather only the relationship between the ensemble mean precipitation and grid-scale variables. The results of the Betts–Miller calculated precipitation using (2) is used as input into a stochastic convection generator, which uses an empirically determined probability distribution to determine the value of precipitation that is seen by the model’s prognostic temperature and moisture equations.

The nonlinear nature of precipitation adds certain constraints to the choice of probability distribution in modeling rainfall. First, because precipitation is positive-only, the distribution cannot have negative values. Second, since a large percentage of the time there is no precipitation, there will need to be a discrete probability that precipitation is identically zero. Finally, the distribution needs to model a decaying probability of occurrence as intensity increases.

Following Kedem et al. (1990), precipitation is modeled in this study as following a mixed lognormal distribution, where nonzero intensities are described by a continuous lognormal distribution, and zero intensity is described by a discrete impulse probability. Although there is question as to whether the lognormal distribution appropriately describes conditional rain rate (e.g., Jameson and Kostinski 1999), such a distribution represents the gross features of rainfall nonlinearity, described above.

The cumulative distribution function (CDF) of this mixed distribution is (Kedem et al. 1990)

$$P(R \leq r) = G(r) = (1 - p)H(r) + pF(r), \quad (3)$$

where R is the area average rain rate (a random variable), and $1 - p$ equals P_0 , the probability that $R = 0$. In this study, rainfall is generally expressed in energy units (W

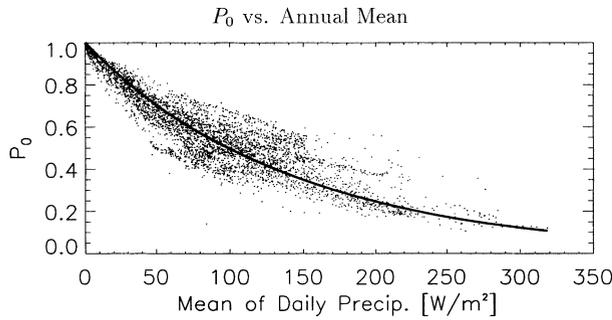


FIG. 1. Value of P_0 vs annual mean of the MSU daily data (in the period 1 Jan 1979–31 Dec 1995). The best-fit line follows (7).

m^{-2}). To convert to mm day^{-1} , divide by 28.2. The function $H(r)$ is zero for $r < 0$, and one for $r \geq 0$. The function $F(r)$ is the CDF for the lognormal distribution, whose probability density function $f(r) = F'(r)$, for $r > 0$, is given by (Kedem et al. 1990)

$f(r, \mu, \sigma)$

$$= \begin{cases} \frac{1}{r\sigma(2\pi)^{1/2}} \exp\left[-\frac{1}{2\sigma^2}(\ln r - \mu)^2\right], & r > 0 \\ 0, & r \leq 0, \end{cases} \quad (4)$$

where μ and σ are parameters describing the lognormal distribution (if Y is a lognormally distributed random variable, then X is a normally distributed random variable with mean μ and variance σ^2 for $Y = e^X$). The units of μ is $\ln(\text{W m}^{-2})$, while σ is unitless (Kedem et al. 1990). Parameter σ is also known as the shape parameter of the lognormal distribution. The mean and variance for R are defined as (Kedem et al. 1990)

$$E[R] = p \exp\left(\mu + \frac{\sigma^2}{2}\right) \quad (5)$$

$$\text{Var}[R] = p \exp(2\mu + \sigma^2)[\exp(\sigma^2) - p]. \quad (6)$$

The value of P_0 , the probability that $R = 0$, is parameterized as a function of Q_c^{BM} :

$$P_0 = \exp(-\mu_p Q_c^{\text{BM}}), \quad (7)$$

where $\mu_p = 0.007$ and is determined by a curve fit by eye of the relationship from observations (Fig. 1). The dataset is described in section 4a; all spatial points in the microwave sounding unit (MSU) dataset are shown that had missing values over less than 30% of the time series.

The empirical lognormal scheme assumes that the Betts–Miller convective parameterization provides an accurate estimate of the mean area average rain rate, that the grid box area is large enough so that it is statistically homogeneous in time and space, and thus $E[R] = Q_c^{\text{BM}}$, where Q_c^{BM} is the value of convective heating (Q_c) calculated by the Betts–Miller scheme based on grid scale C_1 . From this assumption, one can directly

parameterize Q_c as a function of Q_c^{BM} and σ , using (5) and (7). However, in order to simulate the effects of varying the autocorrelation timescale of convection, the randomly chosen convection is embedded in an autoregressive scheme. Thus, convective heating Q_c at time step t is modeled as

$$Q_c = \xi_t Q_c^{\text{BM}}, \quad (8)$$

where the factor ξ_t is scaled to have a mean of 1, and is described by

$$\xi_t = \epsilon_\xi \xi_{t-1} + (1 - \epsilon_\xi) y_t, \quad (9)$$

where ϵ_ξ is an autoregressive coefficient defined for a given σ and τ_ξ .

The characteristic timescale for ξ (τ_ξ) is defined as the time for the autocorrelation function of Q_c to fall to e^{-1} . Runs with values of $\epsilon_\xi = 0.356$, 0.844, and 0.986 are made, which correspond approximately to values of $\tau_\xi = 20$ min, 2 h, and 1 day, respectively.

The variable y_t is the random precipitation value chosen from the mixed lognormal distribution (3), for the current time step t . The distribution of y_t is scaled so that it has a mean of 1. From (5), we then derive a relationship between μ and σ for y_t for $E[R] = 1$:

$$\mu = \ln\left(\frac{1}{p}\right) - \frac{\sigma^2}{2}. \quad (10)$$

Thus, based on (7), (8), and (10), we have a stochastic parameterization of Q_c that is solely a function of Q_c^{BM} and σ .

Short et al. (1993) tally rain-rate statistics over a number of regions and conditions (monsoon and premonsoon rainfall in Darwin, Australia; Florida rain gauge data; and Global Atmospheric Research Programme (GARP) Atlantic Tropical Experiment (GATE) radar measurements in the tropical Atlantic Ocean), and at a few different sampling resolutions (e.g., 1 and 10 min). They find a linear relationship between the mean and standard deviation of conditional rain rate (i.e., the rain rate when there is rain), with a slope for standard deviation versus mean of 5/3. For a lognormally distributed conditional rain rate, using (5) and (6), this implies a constant σ of 1.153 (Short et al. 1993).

In LN, the variance amplitude of the stochastic forcing is set by matching model precipitation spectral power in a target frequency band (from +0.4 to +0.5 day^{-1} , chosen to approximate the amplitude of the white noise floor due to unrepresented convective processes). In this study, we considered using this method. However, in preliminary tests, the highest values of target frequency band spectral power we could obtain in the $\tau_\xi = 1$ day case, without numerical problems, was approximately 5–10 times less than seen in the MSU observations. The reason for this appears to be partly because the low frequency variability responds strongly to the stochastic input for this value of τ_ξ , even more so than in LN. Instead of choosing the amplitude of the white noise input as in

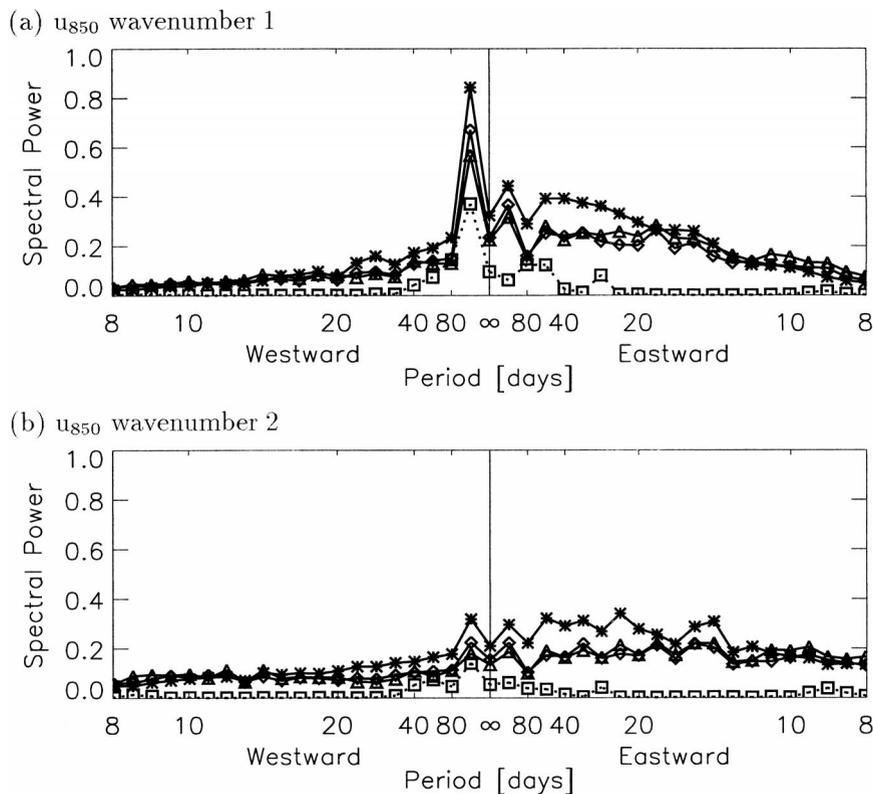


FIG. 2. 850-hPa zonal wind spectral power of daily mean anomalies for (a) wavenumber 1 and (b) wavenumber 2 in an equatorial band from 5.625°N to 5.625°S . Runs are shown for τ_{ξ} equals 20 min (triangle), 2 h (diamond), and 1 day (asterisk). A control run with deterministic Q_c is shown by the dotted line (square). Units of power are $(\text{m s}^{-1})^2$. Std dev of the spectral power estimator is 10%.

LN, we examine the response of the model to a stochastic scheme where values of σ are the same, and thus where the variance of the stochastic input term y_i (in the absence of model feedbacks) is fixed and given by (6). In the empirical lognormal scheme used in this study, all model runs are set to $\sigma = 4$, which is approximately the highest level of σ usable for all experiments.

In principle, this stochastic parameterization will provide input values of variance in accordance to (6), while preserving the value of the mean. In practice, to prevent unrealistic values of Q_c , a maximum cutoff value of $50\,000 \text{ W m}^{-2}$ is applied (which is approximately the record measured 1-day rainfall in a hurricane). Values greater than this cutoff, which are reasonably infrequent, are set to 0. One expects, however, that this will skew the distribution of Q_c such that the mean will be underestimated. This can be verified by running the model using the climatological value of Q_c taken from a control run as the value for Q_c^{BM} . At the same time, one also expects climatology is dominated by large-scale processes, and thus will be robust. Analysis of these issues are found in sections 7 and 8a. A minimum cutoff is also applied to account for possible machine precision limitations; values less than 10^{-5} W m^{-2} are set to 0.

Because the maximum cutoff value has this effect on the mean Q_c represented by the empirical lognormal scheme it affects the first point in the posited methodology: preservation of the mean given by the traditional scheme by the stochastic parameterization. Sensitivity experiments can be made by attaching a constant scaling factor α to (8) such that the empirical lognormal scheme is altered from (8) into

$$Q_c = \alpha \xi_i Q_c^{\text{BM}}. \quad (11)$$

The maximum cutoff for Q_c of $50\,000 \text{ W m}^{-2}$ and the minimum cutoff of 10^{-5} W m^{-2} are still applied, as described in section 3. The value of α can be used to change the mean in an aggregate sense, when it is affected by the maximum cutoff. It can also be used to test impacts in the model climatology of not preserving the mean in the non-mean-preserving stochastic parameterization. Changing the scale factor is similar to changing $1/\tau_c$, since the two are more or less formally equivalent. In section 8, sensitivity tests with different α will reveal surprises concerning how stochastic convection interacts with the large-scale.

Besides the mixed lognormal distribution, other distributions were tested. A mixed exponential distribution did not provide much improvement over the mixed log-

normal in terms of the tail decrease with large values and the necessity of imposing a maximum cutoff value. A mixed beta distribution was also tried, because the cumulative distribution function converges to one at a finite value. However, computing the beta distribution at each time step and grid point in space ended up being computationally too expensive. The mixed lognormal is presented here largely because of the precedent in the remote sensing literature, and because the main results appear to depend more strongly on dynamical feedbacks than on the particular distribution chosen.

Model runs are conducted both with and without model dynamics. In the runs without model dynamics, the stochastic parameterization is run using climatological Q_c from the control run as the stochastic parameterization's input value of Q_c^{BM} .

4. Data and analysis methods

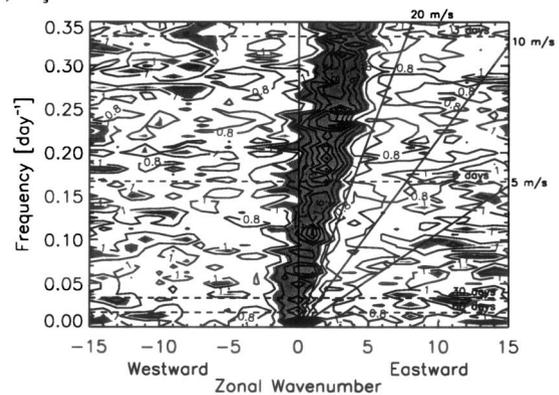
a. Data

The ideal dataset for calculating the precipitation statistics relevant for development (and evaluation) of a stochastic convective parameterization should cover both land and ocean, consist of aggregated measurements (that approximate GCM grid cells) versus point measurements, span a relatively long time period to enable the spectral analysis results to be reasonably significant, and have fine enough resolution such that individual convective complexes as well as mesoscale systems are resolved. A dataset that meets all these criteria is not found. For the purposes of the present study, which represents an initial sensitivity test of the approach 2 stochastic parameterization methodology, the Spencer (1993) daily precipitation estimates [calculated from passive radiometer measurements by the MSU carried on the Television Infrared Observational Satellite-N (TIROS-N) and National Oceanic and Atmospheric Administration (NOAA) series of polar orbiting satellites] is an appropriate compromise. Although the MSU estimates only include oceanic regions, the dataset is relatively long, extending from 1 January 1979 through 31 October 1996. Climatology is calculated as the long-term mean from 1979 to 1995. The grid spacing of the dataset is $2.5^\circ \times 2.5^\circ$ latitude and longitude. If the sensitivity tests suggest that this stochastic parameterization approach may be fruitful, one could imagine future work using additional datasets to produce a more refined estimate of the distribution.

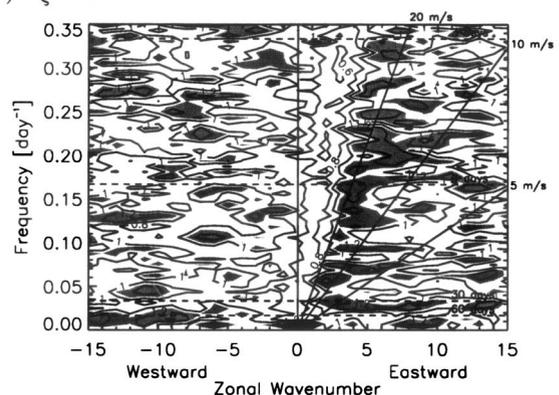
b. Analysis methods

Three cases of τ_ξ are examined: 20 min (which is the model time step), 2 h, and 1 day. A control run is made using the deterministic convective parameterization described in (2). Model runs 46 years (each year consisting of twelve 30-day months) in length are conducted using climatological sea surface temperature (SST) as the low-

(a) $\tau_\xi = 20$ min



(b) $\tau_\xi = 2$ hrs



(c) $\tau_\xi = 1$ day

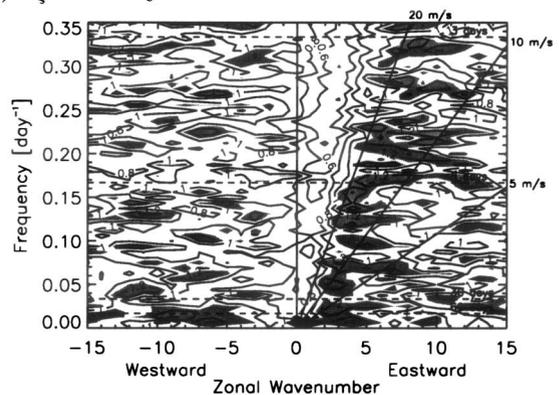


FIG. 3. Spectral power for daily mean anomalies of OLR normalized by an estimate of background spectral power for τ_ξ equals (a) 20 min, (b) 2 h, and (c) 1 day in an equatorial band from 5.625°N to 5.625°S . Areas with values greater than 1.1 are shaded. Contour interval is 0.2. Diagonal lines denote constant phase speed.

er boundary forcing. The SST used is from the Reynolds and Smith (1994) blended dataset. The first year of simulation is discarded to remove any residual spinup.

Daily means of model variables are used in all analysis. Calculations of climatology, variance, and estimates of the probability distribution function (pseudo-PDF) of the model runs use the first 10 years of the

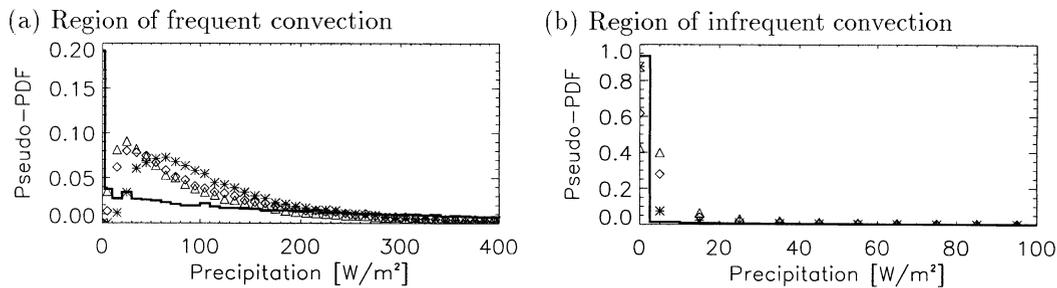


FIG. 4. Pseudo-PDF of observed (MSU) and model daily mean precipitation in (a) a region of frequent convection (model 5.625°N, 180°–202.5°E; MSU 5°N, 180°–202.5°E), and (b) a region of infrequent convection (model 9.375°S, 241.875°–275.625°E; MSU 10°S, 242.5°–275°E). Bin size of probability estimate is 10 W m⁻². MSU is the thick solid line. Model runs shown are for τ_ξ equals 20 min (triangle), 2 h (diamond), and 1 day (asterisk). Note (a) and (b) have different x- and y-axis scales.

analysis period; wavenumber–spectral power calculations use all 45 years of the analysis period. Calculations of variance for the MSU observations use the timeseries extending from 1 January 1979 to 31 December 1995, and use anomalies obtained by subtracting the 1979–95 long-term mean.

The anomalies used in calculating variance and spectral power for the model runs are calculated at each grid point by removing daily mean deviations from a spline fit to the monthly climatology over the respective analysis periods. No detrending is applied. Pseudo-PDFs are calculated by counting the number of values that fall inside a bin, and dividing by the total number of values in the time series. Plots of pseudo-PDFs show the pseudo-PDF value for the middle of the bin. The spectral power estimates use the “summing” method (Press et al. 1989, 465–466), with a nonoverlapping bin group size of $K = 101$, to control the error of the estimator, and a Hanning window to control frequency leakage. This bin group size yields a bandwidth of 0.006 23 day⁻¹, which implies the longest period that can be interpreted physically on the spectral power plots is 160 days.

5. Effects of the empirical lognormal scheme on intraseasonal variability

Runs using the empirical lognormal scheme described in section 3 ($\alpha = 1$) are conducted for the three autocorrelation timescales $\tau_\xi = 20$ min, 2 h, and 1 day. Figure 2 shows wavenumbers 1 and 2 spectral power for 850-hPa zonal wind in an equatorial band, for the three τ_ξ cases. The addition of stochastic noise produces noticeable effects on eastward propagating intraseasonal variability. In wavenumber 1, one sees preferential enhancement of lower frequency variability. In wavenumber 2, the enhancement effects are more broadly spread in frequency. Higher values of τ_ξ generally produce higher levels of intraseasonal variability spectral power. A westward peak is seen in wavenumber 1 spectral power, for which we do not have an explanation. The enhancement of atmospheric variability also extends to lower frequencies (i.e., 160 days in the spectral power figures). While

the spectral power at longer timescales is just as sensitive to the autocorrelation timescale as that at intraseasonal timescales, it raises the possibility that nonresolved variance may contribute in a nonnegligible manner even to very low frequency climate variability.

Figure 3 shows the ratio of OLR spectral power to an estimate of the background power [using a methodology from Wheeler and Kiladis (1999, hereinafter WK) with slight modifications] in an equatorial band for the three τ_ξ cases. The most important differences in methodology are described by Lin et al. (2000). Values of the ratio greater than 1.1 are taken as exceeding the 95% significance level. Although Fig. 3 in the present study is structured to be similar to Fig. 3 in WK, extreme caution should be exercised in any comparison since the methodology assumes that the model representation of background power is such that the model signal’s relationship to model background is similar to the relationship between the observed signal and observed background power. Indications are that this may not be true for the model. Thus, the areas in Fig. 3 showing significant power to background ratio should *not* be interpreted as demonstrating the intraseasonal variability in the model has a certain type of structure, but only that the variability *when compared to the model’s background*, has a certain structure.

From Fig. 3, one sees the low-frequency signal is concentrated in wavenumber 1. The magnitude of the signal is not as high as in WK or Lin et al. (2000). In the version of the model used in Lin et al., extratropical disturbances were stronger and the evaporation–wind feedback tended to be stronger, either of which may account for a stronger MJO-like signal in that version. Comparing results over several wavenumbers for different τ_ξ (Figs. 3a–c), increasing τ_ξ tends to be associated with a slower phase speed in the eastward propagating signal. Phase speed from dry dynamics alone is faster than phase speed with moist dynamics, since the moist dynamics requires time for such feedbacks to be set up. At longer autocorrelation time, one might expect that large random deviations are less frequent, which would enable more time for organized convective feed-

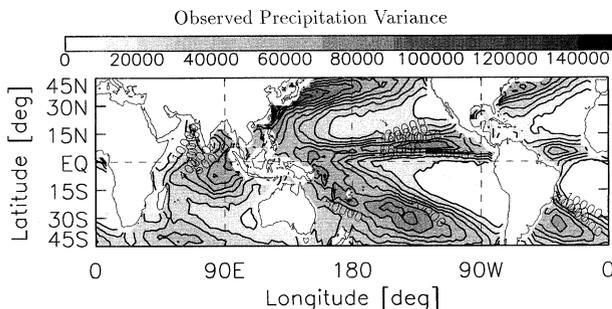


FIG. 5. Variance of daily mean precipitation for MSU observations. Units of $(\text{W m}^{-2})^2$. Contour interval is 10 000.

backs to act. Thus, we speculate that for $\tau_\xi = 20$ min, compared to the cases when $\tau_\xi = 2$ h and $\tau_\xi = 1$ day, the stochastic convection acts at a timescale such that a dry dynamical response is favored, while at the longer τ_ξ , the increased autocorrelation time enables the moist response to make a larger contribution.

6. Simulation of precipitation by the empirical lognormal scheme

a. Probability distribution function

Figure 4 shows the pseudo-PDF for observed and model generated daily mean precipitation, for tropical regions of frequent and infrequent convection. In the frequent convection region (Fig. 4a), the peak of pseudo-PDF is shifted towards low magnitude days for lower values of τ_ξ . Compared to observations, the model overestimates the probability of days less than approximately 200 W m^{-2} . In the infrequent convection region (Fig. 4b), higher values of τ_ξ do a better job in simulating the observed pseudo-PDF. In comparison, the CAPE scheme in LN, for $\tau_\xi = 1$ day, simulates the pseudo-PDF much better than the results in Fig. 4 (see appendix A for results using the CAPE scheme).

b. Variance

Figure 5 shows total variance of observed MSU daily precipitation. Figure 6 shows total variance of model daily precipitation for the three τ_ξ cases. Immediately one notices that the model underestimates peak observed variance by at least a factor of 4. The CAPE scheme in LN also underestimates observed variance. Still, variance simulated by the model shows some of the gross features seen in observations, such as placing the maximum variance in regions of high climatological convection.

In theory, both schemes should allow one to set arbitrary levels of variance. In practice, numerical constraints limit how much stochastic variance can be specified. The variance shown in Fig. 6 is about the maximum variance that can be obtained, with the empirical lognormal scheme, without the model encountering

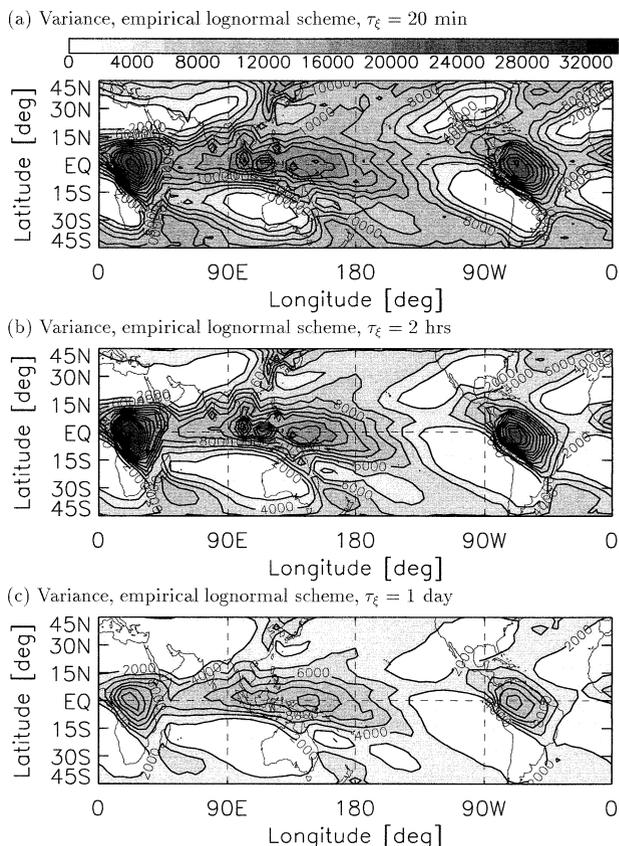


FIG. 6. Variance of daily mean precipitation for model runs using the empirical lognormal scheme ($\alpha = 1$), with τ_ξ equals (a) 20 min, (b) 2 h, and (c) 1 day. Units of $(\text{W m}^{-2})^2$. Contour interval is 2000. All panels in this figure share the same color bar.

problems with numerical stability (presumably associated with occasional large excursions of heating).

As seen in Fig. 6, the formulation used for implementing the autocorrelation timescale (9) results in less variance with higher τ_ξ , since at higher τ_ξ , the weight in front of the stochastic input term y , in (9) decreases.

When precipitation variance (Fig. 6) is compared to the spectral power plots for u_{850} in Fig. 2, one finds that while the largest level of precipitation variance is at lowest value of τ_ξ , the largest value of spectral power associated with intraseasonal variability is at the largest value of τ_ξ . This lack of a one-to-one correspondence between precipitation and 850-hPa zonal wind total variance is also seen when comparing precipitation anomaly spectral power (not shown) and 850-hPa zonal wind anomaly spectral power (Fig. 2). This suggests that the coupling between convection and circulation may be more nuanced than suggested by the oscillating heating source mechanism that has been proposed as an explanation for the MJO (e.g., Chang 1977).

c. Climatology

Figure 7 shows January and July precipitation climatology for the control run. Figure 8 shows January

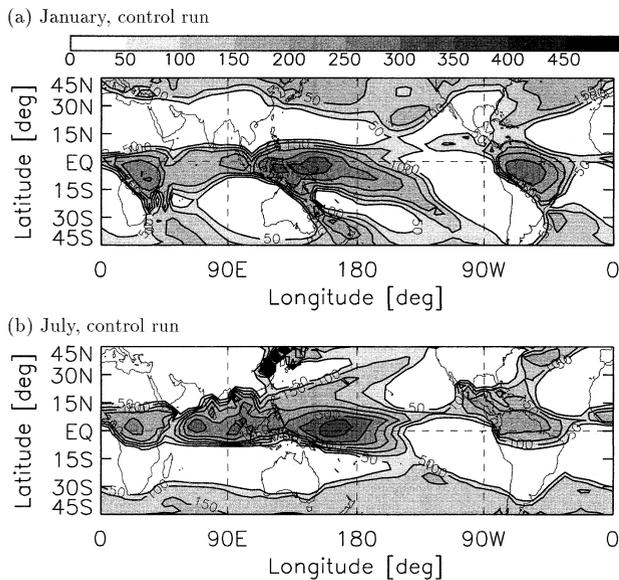


FIG. 7. (a) Jan and (b) Jul precipitation climatology ($W m^{-2}$) for the control run. Contour interval is 50. All climatology plots in this paper share the same color bar.

and July precipitation climatology for model runs using the empirical lognormal scheme for $\tau_{\xi} = 1$ day ($\alpha = 1$). Comparable figures for $\tau_{\xi} = 20$ min and 2 h are similar, except regions with the most extreme precipitation are larger for the $\tau_{\xi} = 1$ day case.

Compared with the control run, the climatology for the stochastic case is generally similar, although it has lower maxima in tropical convergence zones, and has a spatial structure that is more diffused. A variety of causes may contribute to producing such a climatology. The implementation of a maximum precipitation cutoff (described in section 3) in the stochastic parameterization may be removing the large values needed to balance the low-magnitude values to reproduce the mean set by the standard Betts–Miller convective scheme. It is also possible that the stochastic component itself affects climatology, producing less rainfall in areas that would otherwise receive more. Comparing these results to a case where the stochastic parameterization operates without model dynamics may help us understand how the empirical lognormal scheme affects climatology (see next section).

7. The empirical lognormal scheme, without model dynamics

The stochastic parameterization is run without model dynamics by setting the input value of Q_c^{BM} in the empirical lognormal scheme to the value of precipitation climatology from the control run. Figure 9 shows precipitation climatology for such a run ($\alpha = 1$). Only the case for $\tau_{\xi} = 1$ day is shown; climatology for $\tau_{\xi} = 20$ min and 2 h are nearly identical, at a contour interval of $50 W m^{-2}$.

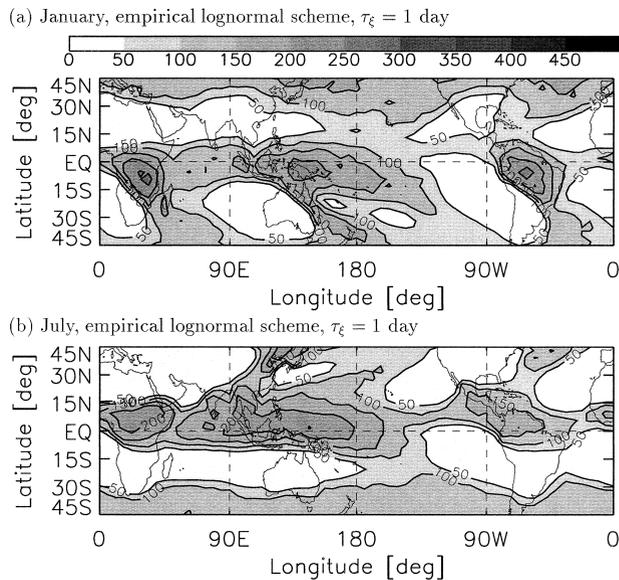


FIG. 8. (a) Jan and (b) Jul precipitation climatology ($W m^{-2}$) for the empirical lognormal scheme ($\alpha = 1$), for $\tau_{\xi} = 1$ day. Contour interval is 50. All climatology plots in this paper share the same color bar.

Although the precipitation climatology without model dynamics (Fig. 9) retains a structure with a tropical maximum, the peak magnitude of the climatology is noticeably lower than in the case with model dynamics (Fig. 8), and in turn is much lower compared to the control run (Fig. 7). Since in the case without model dynamics the cutoff is the only cause of reducing the mean of precipitation, such a cutoff decreases clima-

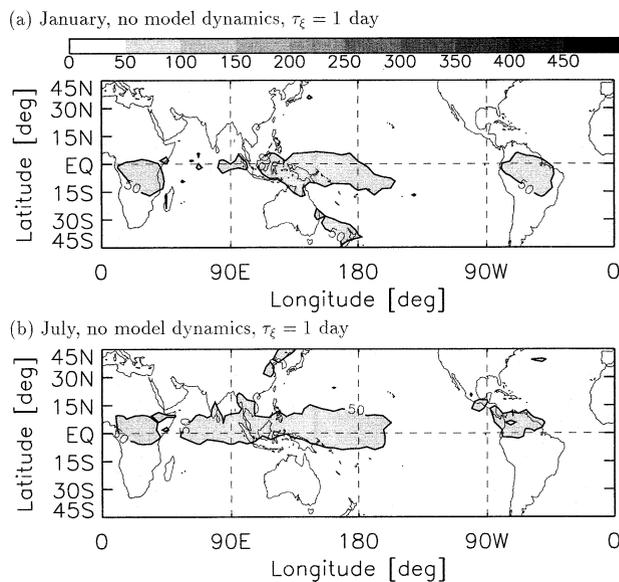


FIG. 9. (a) Jan and (b) Jul precipitation climatology ($W m^{-2}$) for the empirical lognormal scheme ($\alpha = 1$), without model dynamics, for $\tau_{\xi} = 1$ day. Contour interval is 50. All climatology plots in this paper share the same color bar.

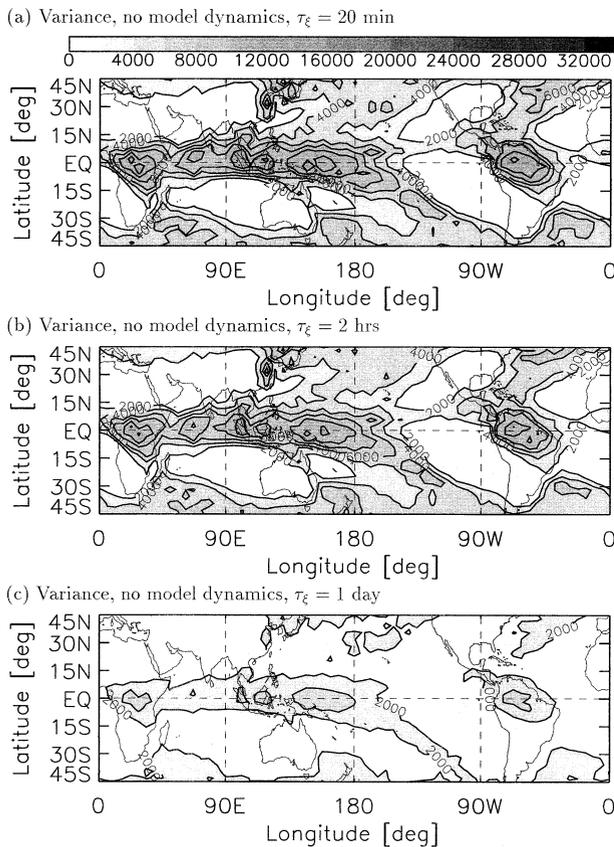


FIG. 10. Variance of daily mean precipitation for model runs using the empirical lognormal scheme ($\alpha = 1$), with no model dynamics, with τ_ξ equals (a) 20 min, (b) 2 h, and (c) 1 day. Units of $(\text{W m}^{-2})^2$. Contour interval is 2000. All panels in this figure share the same color bar.

tology quite substantially. At the same time, as the case with model dynamics shows, the model dynamics works to increase the climatology to much higher values.

Figure 10 shows total variance of daily mean precipitation for the empirical lognormal scheme ($\alpha = 1$) with model dynamics removed. As one would expect, and as in the case with model dynamics included, higher values of τ_ξ have lower total variance. Variance is lower, though, when model dynamics are removed, implying that the model dynamics act to increase variance that is provided to the system.

One outstanding question has been the importance of the shape of the probability distribution, both for low-magnitude days as well as the tail of the distribution (the latter of which is altered by the maximum cutoff). The experiments described in this section, however, suggest that interactions with model dynamics may be more important than distribution shape. Further experiments will test the sensitivity of climatology and other quantities, to help elucidate the relative importance of deterministic versus stochastic processes.

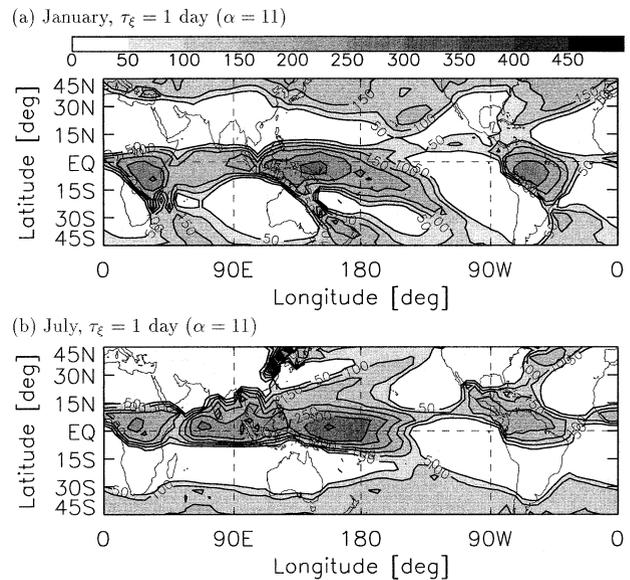


FIG. 11. (a) Jan and (b) Jul precipitation climatology (W m^{-2}) for the empirical lognormal scheme, $\alpha = 11$, for $\tau_\xi = 1$ day. Contour interval is 50. All climatology plots in this paper share the same color bar.

8. Sensitivity of the empirical lognormal scheme to a scaling factor

In section 7 it was seen that while the maximum precipitation cutoff (discussed in section 3) substantially decreases the magnitude of the precipitation climatology, model dynamics work against the reduction. In this section, a case is considered where the precipitation climatology produced by the stochastic parameterization matches the control run climatology, in an aggregate sense. This is done by setting α in (11) to 11, which is a value that will produce a climatology that matches that given in the control run (Fig. 7). Runs are made with and without model dynamics, using $\alpha = 11$. In the case without model dynamics, the stochastic scheme uses precipitation climatology from the control run as the input value for Q_c^{BM} , as described in section 3.

a. Climatology

Figure 11 shows precipitation climatology for the empirical lognormal scheme, with $\alpha = 11$, with model dynamics. Comparable figures for $\tau_\xi = 20$ min and 2 h are similar, except that regions with the most extreme precipitation are larger for $\tau_\xi = 1$ day. Figure 12 shows precipitation climatology for the same case, except with model dynamics removed (as described in section 3). For the runs without model dynamics, only the case with $\tau_\xi = 1$ day is shown; climatology for $\tau_\xi = 20$ min and 2 h are nearly identical, at a contour interval of 50 W m^{-2} .

In comparing Fig. 11 and Fig. 12, one finds that precipitation climatology is the same using the stochastic parameterization with and without model dynamics, and

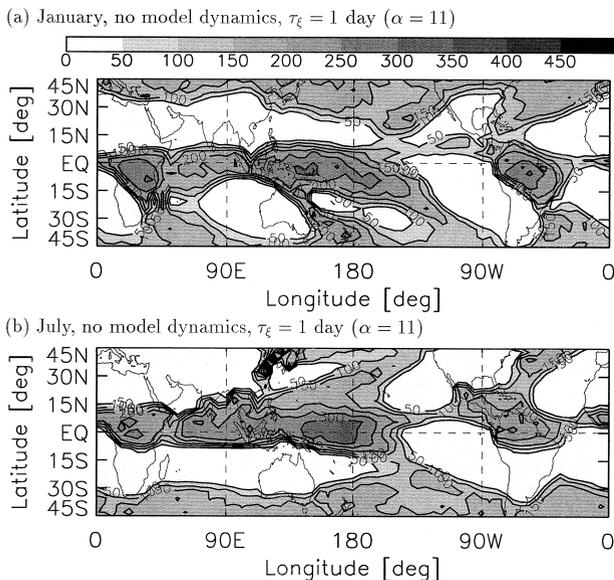


FIG. 12. (a) Jan and (b) Jul precipitation climatology (W m^{-2}) for the empirical lognormal scheme, no model dynamics, $\alpha = 11$, for $\tau_\xi = 1$ day. Contour interval is 50. All climatology plots in this paper share the same color bar.

that both reproduce the climatology of the control run. This implies that inclusion of a stochastic component to precipitation does not necessarily substantially alter the distribution or magnitude of climatological convection.

That the precipitation climatology for the empirical lognormal scheme, with $\alpha = 11$, nearly reproduces the climatology from the control run also suggests that precipitation climatology is fundamentally a function of the model, as opposed to the presence of a stochastic component. This also suggests that tests be made for the deterministic case with a different τ_c , since τ_c is more or less formally identical to α in the empirical lognormal scheme (11).

Figure 13 shows precipitation climatology from the deterministic $\tau_c = 22$ h run. This climatology is nearly the same as that for the empirical lognormal scheme, with $\alpha = 1$, with model dynamics (Fig. 8; the similarity is clearest when comparing to the case of $\tau_\xi = 1$ day). At the same time, when $\alpha = 11$, the climatology is almost identical to the deterministic case when $\tau_c = 2$ h (Fig. 7). Since Q_c is proportional to $1/\tau_c$, in the deterministic case this is equivalent to multiplying Q_c by 11. This implies that in both cases using the empirical lognormal scheme ($\alpha = 1$ and 11), the climatology is insensitive to the inclusion of stochastic noise itself, but rather is more sensitive to changes that effectively act as changes to τ_c . Because CAPE in the Tropics so easily adjusts to balance with large-scale dynamics, the effect after interaction with dynamics of a scaling factor in the Q_c term (whether it be α or $1/\tau_c$) is more nuanced than might be expected.

Note that in the case of deterministic $\tau_c = 22$ h, total variance, as well as spectral power separated into wave-

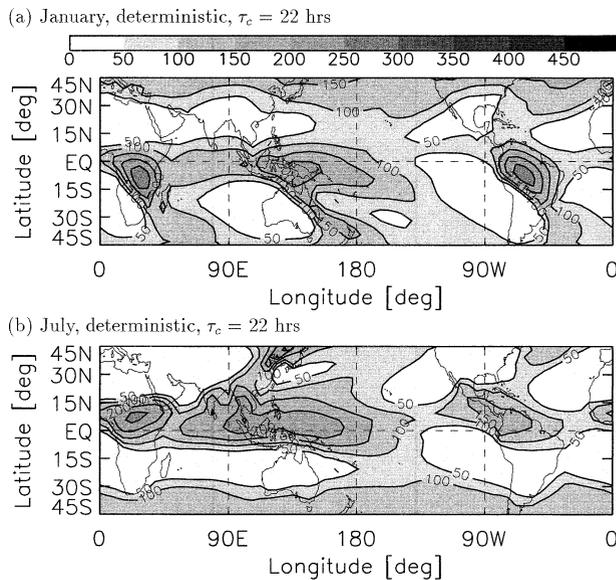


FIG. 13. (a) Jan and (b) Jul precipitation climatology (W m^{-2}) for deterministic run with $\tau_c = 22$ h for this figure should be compared to Fig. 8 (see text). Contour interval is 50. All climatology plots in this paper share the same color bar.

numbers (not shown), is decreased substantially. Thus, the effect of a scaling factor in a deterministic case is different than in a stochastic case. For the stochastic case, in comparing the runs with the scaling factor changed from $\alpha = 11$ to $\alpha = 1$, one finds a slight decrease in total variance of precipitation and an increase in low-wavenumber 850-hPa zonal wind anomaly spectral power at intraseasonal timescales. In the deterministic case, for a comparable change in the Q_c scaling factor (i.e., $\tau_c = 2$ h to $\tau_c = 22$ h) total variance (not shown) is decreased to nearly negligible levels, and 850-hPa zonal wind anomaly spectral power (not shown) spectral power shows a general decrease in low wavenumber power with period between 20 and 80 days. This result is consistent with previous analysis of the response of the tropical atmosphere to changing τ_c , as represented in deterministic linear models (Yu and Neelin 1994; Neelin 1997). In those models, increasing τ_c acts to suppress higher wavenumbers by allowing the tropical atmosphere to deviate from quasi-equilibrium and thus become more convectively damped.

These results suggest the following ideas about precipitation climatology and the inclusion of a stochastic component to convection:

- Climatology is fundamentally determined by large-scale model processes. Changes to the stochastic convection parameterization that would appear in an off-line test (no dynamical feedbacks) to affect the mean precipitation turn out to be of secondary importance when feedbacks are included.
- The climatology of the system tends toward a state

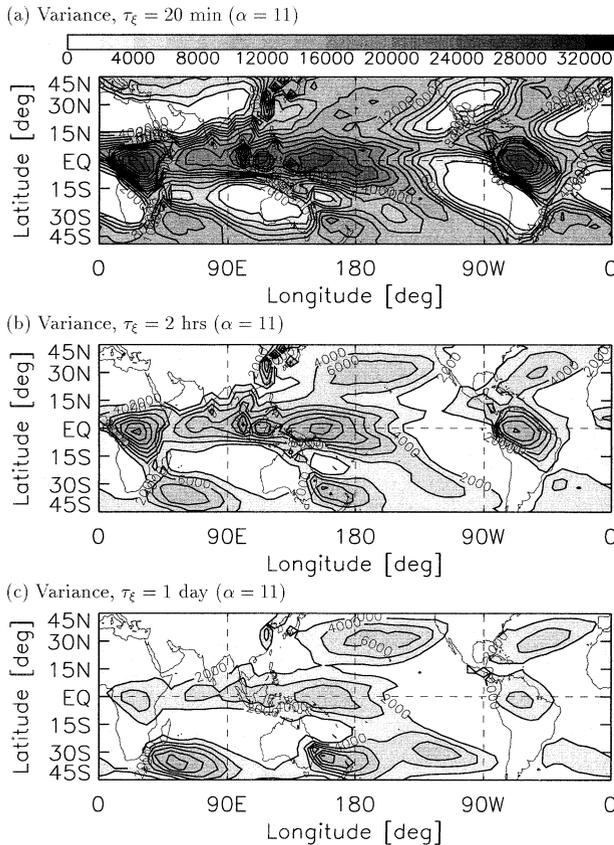


FIG. 14. Variance of daily mean precipitation for model runs using the empirical lognormal scheme, $\alpha = 11$, with τ_ξ equals (a) 20 min, (b) 2 h, and (c) 1 day. Units of $(W m^{-2})^2$. Contour interval is 2000. All panels in this figure share the same color bar.

that is set by the choice of parameters unrelated to the presence of stochastic noise (such as τ_c).

- The method with which one parameterizes the stochastic nature of convection can change the precipitation climatology of the model in a way similar to changing convective timescale τ_c .
- In general one expects stochastically driven variations in the presence of nonlinearity to affect the mean state, but (perhaps surprisingly) this mechanism appears secondary to the processes listed above.

Having explained some of the effects on climatology, and seen that the stochastic parameterization itself does not have much impact on the behavior of climatology, we examine the impact of $\alpha = 11$, as well as model dynamics, on total variance and pseudo-PDF, quantities that should be more influenced by the fluctuating nature of a stochastic scheme versus a deterministic scheme.

b. Variance

Tropical precipitation total variance for the $\alpha = 11$ case with model dynamics (Fig. 14), for $\tau_\xi = 20$ min, is comparable to the $\alpha = 1$ case. For $\tau_\xi = 1$ day, the

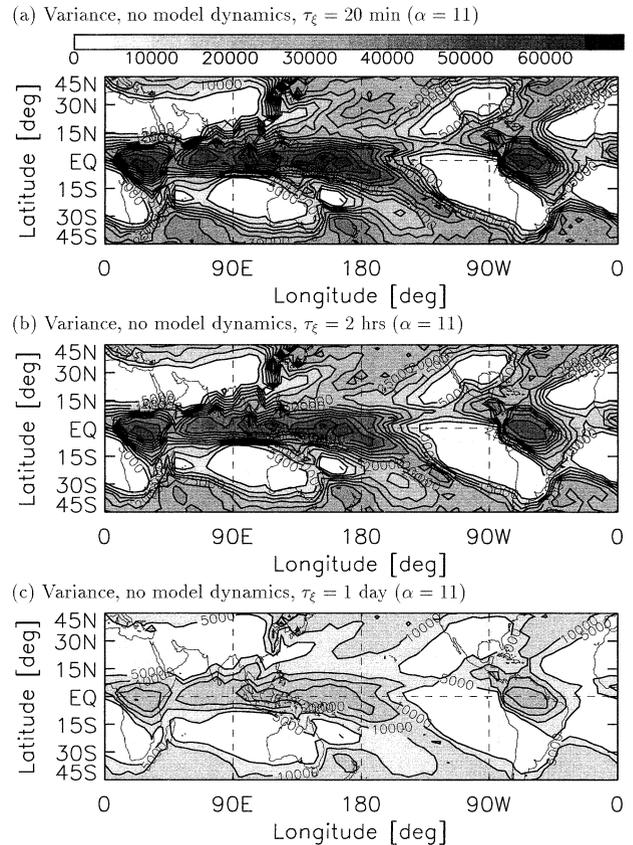


FIG. 15. Variance of daily mean precipitation for model runs using the empirical lognormal scheme, no model dynamics, $\alpha = 11$, with τ_ξ equals (a) 20 min, (b) 2 h, and (c) 1 day. Units of $(W m^{-2})^2$. Contour interval is 5000 (which is different from other model run variance figures). All panels in this figure share the same color bar.

$\alpha = 11$ case has tropical precipitation variance weaker than the $\alpha = 1$ case; the variance maximum no longer remains even in the Tropics.

Comparing the $\alpha = 11$ case with model dynamics to the $\alpha = 11$ case without model dynamics, one finds quite different behavior than in the $\alpha = 1$ case (with and without model dynamics). Figure 15 shows total precipitation variance for $\alpha = 11$, with model dynamics removed. One finds that total variance produced by the stochastic scheme in absence of model dynamics is quite substantial. And yet, this variance does not translate into very much total variance when it is filtered through model dynamics. In fact, the model dynamics appears to act in a manner *opposite* to how it behaves in the $\alpha = 1$ case, where the variance with model dynamics is higher than without model dynamics. We do not have an explanation for this behavior, but it clearly underlines the strong modifications that occur by interaction with model dynamics.

c. Probability distribution function

Figure 16 shows the pseudo-PDF for a region of frequent convection and a region of infrequent convection,

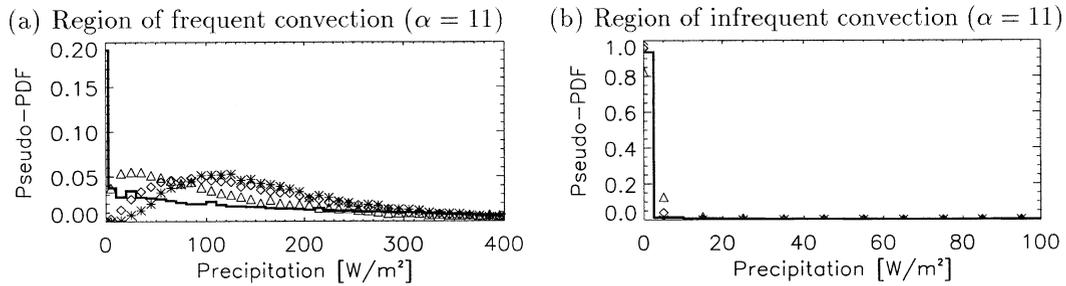


FIG. 16. Pseudo-PDF of observed (MSU) and model daily mean precipitation in a (a) region of frequent convection (model 5.625°N, 180°–202.5°E; MSU 5°N, 180°–202.5°E), and a (b) region of infrequent convection (model 9.375°S, 241.875°–275.625°E; MSU 10°S, 242.5°–275°E). Bin size of probability estimate is 10 W m⁻². MSU is the thick solid line. Model runs shown are for $\alpha = 11$ for τ_ξ equals 20 min (triangle), 2 h (diamond), and 1 day (asterisk). Note (a) and (b) have different x- and y-axis scales.

for the empirical lognormal scheme with $\alpha = 11$. The pseudo-PDF for the empirical lognormal scheme with $\alpha = 11$, without model dynamics, is not shown.

Compared to the empirical lognormal scheme with $\alpha = 1$, the $\alpha = 11$ pseudo-PDF seems to better match the observed pseudo-PDF, both in areas of frequent and infrequent convection. As one would expect, setting $\alpha = 11$ generally decreases the frequency of low-magnitude days. Yet, in comparison to the $\alpha = 1$ runs, the

strength of the intraseasonal variability is much less (see section 8d), suggesting that a better match of the observed distribution is not enough by itself to enhance intraseasonal variability.

In the $\alpha = 11$ (as well as the $\alpha = 1$) case, the probability of zero values in regions of frequent convection is substantially underestimated, and thus could probably be improved. As it stands, we do not know whether or not it is a leading order effect. In the $\alpha = 1$ case, the

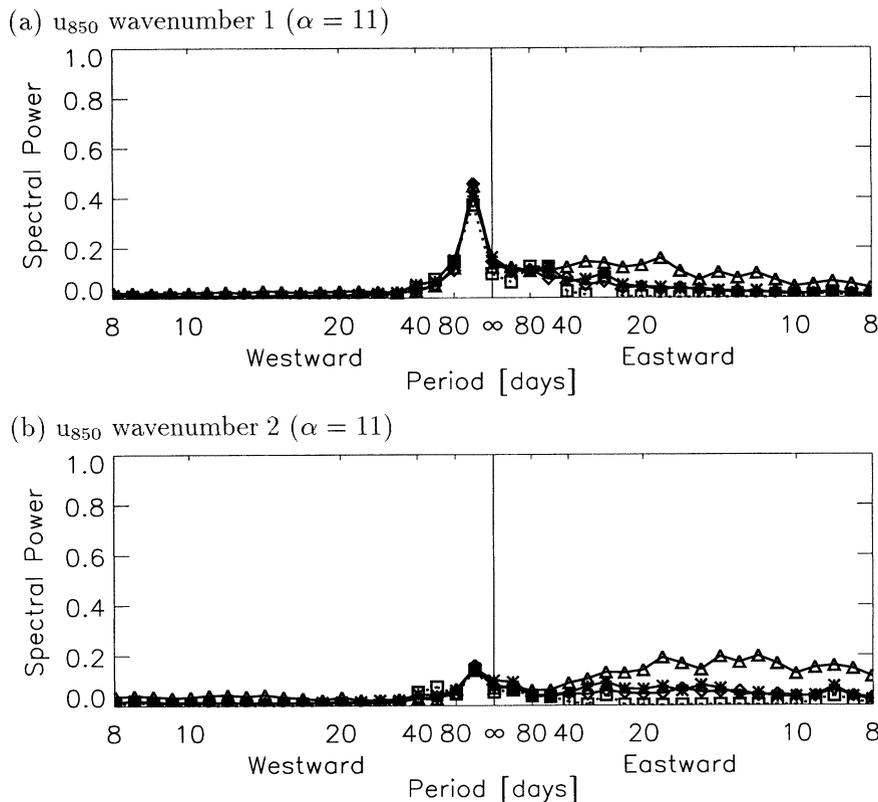


FIG. 17. Spectral power for 850-hPa zonal wind (a) wavenumber 1 and (b) wavenumber 2 in an equatorial band from 5.625°N to 5.625°S. Runs for $\alpha = 11$ are shown for $\tau_\xi = 20$ min (triangle), 2 h (diamond), and 1 day (asterisk). A control run is shown by the dotted line (square). Units of (m s⁻¹)². Std dev of the spectral power estimator is 10%.

pseudo-PDF of the stochastic convection without model dynamics (not shown) is quite different from with model dynamics, particularly at low-magnitude days. For $\alpha = 1$, model dynamics also acts to alter the distribution of low-magnitude days (not shown). Because model dynamics can substantially change the distribution of Q_c , one must iterate around the interaction with model dynamics to obtain a match with observations.

d. Spectral analysis

Figure 17 shows spectral power for daily mean anomalies of 850-hPa zonal wind for the empirical lognormal scheme with $\alpha = 11$. The impact of the stochastic scheme on increasing spectral power of intraseasonal variability is much less. Of what effect there is on spectral power, the largest occurs for the $\tau_\xi = 20$ min case. This is in contrast to the $\alpha = 1$ case, where $\tau_\xi = 1$ day has the largest intraseasonal response. For the $\tau_\xi = 20$ min case, however, the spectral power for $\alpha = 11$ is only somewhat less than for $\alpha = 1$. For no obvious reason, something in the interaction with dynamics is filtering the variance when a scaling factor is applied (as in the $\alpha = 11$ case).

Analysis similar to that used to produce Fig. 3 is also conducted for $\alpha = 11$. The results (not shown) are similar to Fig. 3 and likewise show a reduction of the phase speed of the eastward propagating signal with increasing τ_ξ .

9. Discussion and conclusions

When developing convective parameterizations that seek to include the effects of unresolved convective processes upon grid-scale variables, two different approaches can be taken. In approach 1, the behavior and physical effects of unresolved processes (i.e., effects associated with subgrid convective motions and mesoscale systems) are modeled. Thus, it is not assumed that the overall distribution of convective heating is known in advance. The scheme used in LN provides an elementary example of such a modeling approach. In that study, it was assumed that unresolved effects in buoyancy related variables (the model's formulation for CAPE) could be described as a first-order autoregressive red noise process. This is, in the simplest sense of the term, a "model" of the physics involved in the unresolved processes, and thus the effect of this noise upon total convective heating is not known beforehand. The extent to which the scheme in LN can simulate the observed probability distribution of convective heating is thus a measure of the validity of the assumptions LN make in the "model" of the unresolved physics. On the other end of the spectrum in this first class of approaches to stochastic convective parameterizations would be a randomly initiated simplified model of a mesoscale system embedded within a GCM grid cell.

In approach 2, a stochastic parameterization is formulated that does not seek to explicitly model the effects

of unresolved processes per se, but rather seeks to just control the key higher-moment statistics of the convective heating that the model sees. Theoretically, this approach would lend itself most readily to implementation in a full-scale GCM; the stochastic scheme would be calibrated offline, to match the key statistical traits observed in actual precipitation, and then implemented into the GCM. Another way to think about the difference between these two approaches to stochastic parameterization is that in approach 1, one parameterizes the physics of the unresolved processes using a stochastic framework, while in approach 2, one parameterizes the statistics of the convective heating, using a stochastic framework. The empirical lognormal scheme used in this study illustrates the latter approach for a model of intermediate complexity.

As in LN, this present study suggests that the inclusion of unresolved variance in the form of a stochastic convective parameterization affects intraseasonal variability. When considered further, the results of these two studies also suggest a checklist of considerations that appear important in formulating a stochastic convective parameterization:

- *The influence of a stochastic parameterization on intraseasonal variability may be very different from the influence on the climatology.* To leading order, climatology appears to be only modestly affected by the inclusion of a stochastic component (beyond ways that are comparable to changing the deterministic parameterization). Intraseasonal variability, on the other hand, appears to be sensitive to the presence of noise, as well as to details of its parameterization.
- *Effects of autocorrelation time differ with context.* Changing the autocorrelation time of a stochastic process that acts on elements within a convective scheme (such that the process is operated on by the nonlinearities of the convective scheme, as well as the nonlinearities of the model dynamics) has a substantial effect. There appears to be less sensitivity to changes in autocorrelation time of a stochastic process that acts outside the scheme (such that the autocorrelation time is operated on by the nonlinearities of the model dynamics alone).
- *Preservation of the mean vertically integrated convective heating is not crucial.* We initially postulated that a useful property of the stochastic parameterization would be that it yield the same mean as a related deterministic scheme. The results suggest that even if the convective parameterization (without model dynamics) does not reproduce climatology, once the stochastic parameterization is implemented in an atmospheric model, the model dynamics will tend to adjust the mean towards a climatology intrinsic to the model, for given parameter settings.
- *Heating strongly interacts with the large-scale, sometimes unexpectedly.* Prior to this study, we hypothesized that an empirical stochastic parameterization could be calibrated offline, and then implemented in

the atmospheric model. The results of this study suggest this is not a workable strategy since heating and large-scale dynamics are so intertwined. This is particularly dramatic when total variance and probability distribution are considered. Parameters affecting the probability distribution of heating need to be chosen in the context of interactive model dynamics.

During this study, attention was paid to finding a probability distribution that would properly represent the key features of stochastically varying convection, while at the same time would satisfy numerical constraints that inevitably enter into modeling efforts. The major constraints included ensuring model precipitation did not reach values that would cause numerical problems and limiting the computational expense required for inverting a distribution. We note the caveat that the distribution used in this study has deficiencies (especially in representation of the probability that precipitation equals zero) that may be important. In the end, however, the results suggest that large-scale effects of model dynamics so strongly affect the behavior of convection in the model that details of the distribution are not as important as was believed.

Implementation of the empirical lognormal example of stochastic parameterization approach 2 suggests that the atmosphere (as simulated in this model) selectively and heavily modifies any input variance. The main postulated merit of approach 2 is the hope that one might be able to parameterize the unresolved variance separately from considerations of model dynamics. The results of this study suggests such direct control of the distribution of convective heating ignores important feedbacks, since interactions with model dynamics so strongly influence effects of the stochastic convection. Yet, as seen in LN, even a simple example of approach 1 to stochastically parameterizing unresolved variance (where the physics of the convective elements are modeled stochastically) yields a probability distribution of precipitation that more closely matches observations than the approach used in the present study. Accordingly, approach 1 is recommended.

Perhaps more importantly, the results of approach 2 strongly support the LN approach 1 results suggesting that the effort of developing stochastic convective pa-

rameterizations is worthwhile. Results from experiments using both the physical modeling and empirical approaches indicate that large-scale intraseasonal variability is affected. While less conclusive, there are indications that atmospheric contributions at longer time-scales are likewise affected. Results from this present study suggest unresolved variance strongly interacts with large-scale dynamics, even if the precise characteristics of its representation are crudely formulated. The overall conclusion of both approaches is to provide modeling evidence that a substantial percentage of the large-scale variance observed in the atmosphere (not only in precipitation, but also in winds) might be the response of the large-scale to smaller-scale variability. It appears that the higher-moment effects (as opposed to just the mean) of convective scale and mesoscale motions may be important to the climate system, and that modeling these effects may be a useful—and challenging—task to accomplish.

Acknowledgments. This research was partially supported by National Science Foundation Grant ATM-0082529 and National Oceanic and Atmospheric Administration Grant NA16GP2003. During final preparation and revisions of the manuscript, JWJ was supported by a visiting fellowship from the Cooperative Institute for Research in Environmental Sciences (CIRES) at the University of Colorado, Boulder. This study benefited from discussions with C. Chou, J. Gao, B. Mapes, J. Meagher, H. Su, and N. Zeng. MSU precipitation data were obtained from the National Center for Atmospheric Research data archives and from the NOAA-CIRES Climate Diagnostics Center.

APPENDIX

Pseudo-PDF for the CAPE Scheme

Figures A1 and A2 show pseudo-PDFs in a region of frequent tropical convection for the CAPE scheme used in LN, for the $\tau_{\xi} = 1$ day case. Pseudo-PDFs of MSU observations are also shown, for comparison. These plots can be compared to model runs using the empirical lognormal scheme (Figs. 4 and 16).

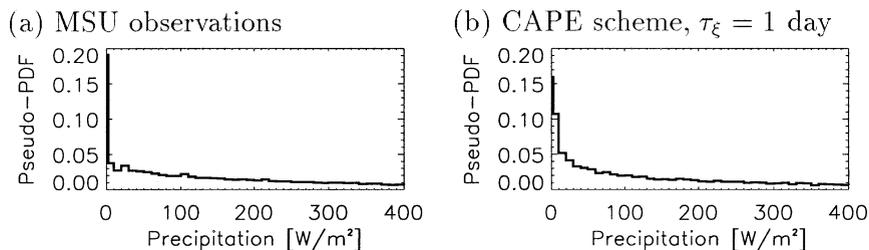


FIG. A1. Pseudo-PDF of daily mean precipitation in region of frequent tropical convection for (a) MSU (5°N , 180° – 202.5°E during the period 1 Jan 1979–31 Dec 1995), and (b) model runs using the CAPE scheme for $\tau_{\xi} = 1$ day (5.625°N , 180° – 202.5°E for 10 model years). Bin size for both pseudo-PDFs is 10 W m^{-2} .

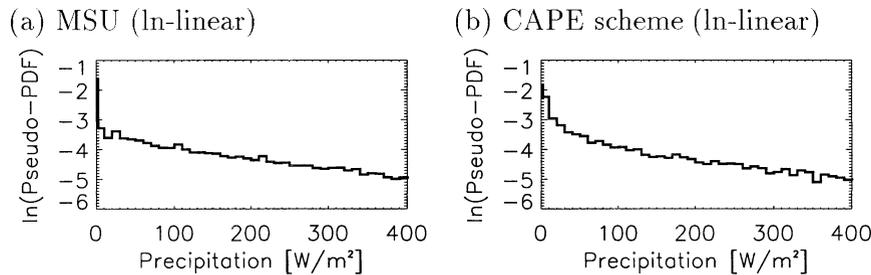


FIG. A2. Same as Fig. A1, except the natural logarithm of pseudo-PDF is plotted on the y axis. Bin size for both pseudo-PDFs is 10 W m^{-2} .

REFERENCES

- Betts, A. K., and M. J. Miller, 1986: A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX, ATEX and Arctic air-mass data sets. *Quart. J. Roy. Meteor. Soc.*, **112**, 693–709.
- Blanke, B., J. D. Neelin, and D. Gutzler, 1997: Estimating the effect of stochastic wind stress forcing on ENSO irregularity. *J. Climate*, **10**, 1473–1486.
- Buizza, R., M. Miller, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908.
- Chang, C.-P., 1977: Viscous internal gravity waves and low-frequency oscillations in the tropics. *J. Atmos. Sci.*, **34**, 901–910.
- Eagleson, P. S., 1978: Climate, soil, and vegetation: 2. The distribution of annual precipitation derived from observed storm sequences. *Water Resour. Res.*, **14**, 713–721.
- Eckert, C., and M. Latif, 1997: Predictability of a stochastically forced hybrid coupled model of El Niño. *J. Climate*, **10**, 1488–1504.
- Farrell, B. F., and P. J. Ioannou, 1994: A theory for the statistical equilibrium energy spectrum and heat flux produced by transient baroclinic waves. *J. Atmos. Sci.*, **51**, 2685–2698.
- Hasselmann, K., 1976: Stochastic climate models. Part I: Theory. *Tellus*, **28**, 473–485.
- Jameson, A. R., and A. B. Kostinski, 1999: Fluctuation properties of precipitation. Part V: Distribution of rain rates—Theory and observations in clustered rain. *J. Atmos. Sci.*, **56**, 3920–3932.
- Katzfey, J. J., and B. F. Ryan, 2000: Midlatitude frontal clouds: GCM-scale modeling implications. *J. Climate*, **13**, 2729–2745.
- Kedem, B., L. S. Chiu, and G. R. North, 1990: Estimation of mean rain rate: Application to satellite observations. *J. Geophys. Res.*, **95**, 1965–1972.
- Kleeman, R., and A. M. Moore, 1997: A theory for the limitation of ENSO predictability due to stochastic atmospheric transients. *J. Atmos. Sci.*, **54**, 753–767.
- Lau, K.-M., J. H. Kim, and Y. Sud, 1996: Intercomparison of hydrologic processes in AMIP GCMs. *Bull. Amer. Meteor. Soc.*, **77**, 2209–2227.
- Leith, C. E., 1975: Climate response and fluctuation dissipation. *J. Atmos. Sci.*, **32**, 2022–2026.
- 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.
- , —, and N. Zeng, 2000: Maintenance of tropical intraseasonal variability: Impact of evaporation-wind feedback and midlatitude storms. *J. Atmos. Sci.*, **57**, 2793–2823.
- Majda, A. J., I. Timofeyev, and E. V. Eijnden, 1999: Models for stochastic climate prediction. *Proc. Natl. Acad. Sci.*, **96**, 14 687–14 691.
- Manabe, S., J. Smagorinsky, and R. F. Strickler, 1965: Simulated climatology of a general circulation model with a hydrological cycle. *Mon. Wea. Rev.*, **93**, 769–798.
- Moncrieff, M. W., 1981: A theory of organized steady convection and its transport properties. *Quart. J. Roy. Meteor. Soc.*, **107**, 29–50.
- , 1992: Organized convective systems: Archetypal dynamical models, mass and momentum flux theory, and parameterization. *Quart. J. Roy. Meteor. Soc.*, **118**, 819–850.
- Neelin, J. D., 1997: Implications of convective quasi-equilibrium for the large-scale flow. *The Physics and Parameterization of Moist Atmospheric Convection*, R. K. Smith, Ed., Kluwer Academic, 413–446.
- , and N. Zeng, 2000: A quasi-equilibrium tropical circulation model—Formulation. *J. Atmos. Sci.*, **57**, 1741–1766.
- Oerlemans, J., 1979: A model of a stochastically driven ice sheet with planetary wave feedback. *Tellus*, **31**, 469–477.
- Penland, C., and L. Matrosova, 1994: A balance condition for stochastic numerical models with application to the El Niño–Southern Oscillation. *J. Climate*, **7**, 1352–1372.
- Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1989: *Numerical Recipes in Pascal*. Cambridge University Press, 750 pp.
- Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimal interpolation. *J. Climate*, **7**, 929–948.
- Ricciardulli, L., and R. R. Garcia, 2000: The excitation of equatorial waves by deep convection in the NCAR Community Climate Model (CCM3). *J. Atmos. Sci.*, **57**, 3461–3487.
- Rodriguez-Iturbe, I., D. R. Cox, and V. Isham, 1987: Some models for rainfall based on stochastic point processes. *Proc. Roy. Soc. London*, **410A**, 269–288.
- Saravanan, R., and J. C. McWilliams, 1998: Advective ocean–atmosphere interaction: An analytical stochastic model with implications for decadal variability. *J. Climate*, **11**, 165–188.
- Short, D. A., D. B. Wolff, D. Rosenfeld, and D. Atlas, 1993: A study of the threshold method utilizing raingage data. *J. Appl. Meteor.*, **32**, 1379–1387.
- Spencer, R. W., 1993: Global oceanic precipitation from the MSU during 1979–91 and comparisons to other climatologies. *J. Climate*, **6**, 1301–1326.
- Timmermann, A., and G. Lohmann, 2000: Noise-induced transitions in a simplified model of the thermohaline circulation. *J. Phys. Oceanogr.*, **30**, 1891–1900.
- Wheeler, M., and G. N. Kiladis, 1999: Convectively coupled equatorial waves: Analysis of clouds and temperature in the wave-number–frequency domain. *J. Atmos. Sci.*, **56**, 374–399.
- Xu, K.-M., A. Arakawa, and S. K. Krueger, 1992: The macroscopic behavior of cumulus ensembles simulated by a cumulus ensemble model. *J. Atmos. Sci.*, **49**, 2402–2420.
- Yu, J.-Y., and J. D. Neelin, 1994: Modes of tropical variability under convective adjustment and the Madden–Julian oscillation. Part II: Numerical results. *J. Atmos. Sci.*, **51**, 1895–1914.
- Zeng, N., J. D. Neelin, and C. Chou, 2000: A quasi-equilibrium tropical circulation model—Implementation and simulation. *J. Atmos. Sci.*, **57**, 1767–1796.
- Zhang, G. J., and N. A. McFarlane, 1995: Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian Centre General Circulation Model. *Atmos.–Ocean*, **33**, 407–446.