Reconciling non-Gaussian Climate Statistics with Linear Dynamics

Prashant D. Sardeshmukh and Philip Sura

Climate Diagnostics Center

CIRES, University of Colorado

and Earth System Research Laboratory, NOAA

Boulder, Colorado

Journal of Climate

Revised 15 July 2008

Abstract

Linear stochastically forced models have been found to be competitive with comprehensive nonlinear weather and climate models at representing many features of the observed covariance statistics and at predictions beyond a week. Their success seems at odds with the fact that the observed statistics can be significantly non-Gaussian, which is often attributed to nonlinear dynamics. The stochastic noise in the linear models can be a mixture of state-independent ("additive") and linearly state-dependent ("multiplicative") Gaussian white noises. It is shown here that such mixtures can produce not only symmetric but also skewed non-Gaussian probability distributions if the additive and multiplicative noises are correlated. Such correlations are readily anticipated from first principles. A generic stochastically generated skewed (SGS) distribution can be analytically derived from the Fokker-Planck equation for a single-component system. In addition to skew, all such SGS distributions have power-law tails, and a striking property that the (excess) kurtosis K is always greater than 1.5 times the square of the skew S. Remarkably, this K-S inequality is found to be satisfied by circulation variables even in the observed multi-component climate system. A principle of "Diagonal Dominance" in the multi-component moment equations is introduced to understand this behavior.

To clarify the nature of the stochastic noises (turbulent adiabatic versus diabatic fluctuations) responsible for the observed non-Gaussian statistics, a long 1200-winter simulation of the northern winter climate is generated using a dry adiabatic atmospheric general circulation model forced only with the observed long-term winter-mean diabatic forcing as a constant forcing. Despite the complete neglect of diabatic variations, the model reproduces the observed K-S relationships, and also the spatial patterns of the skew and kurtosis of the daily

tropospheric circulation anomalies. This suggests that the stochastic generators of these higher moments are mostly associated with local adiabatic turbulent fluxes. The model also simulates fifth moments that are approximately 10 times the skew, and probability densities with power-law tails, as predicted by the linear theory.

1. Introduction

Although the governing equations for weather and climate evolution are obviously nonlinear, in many contexts the evolution of anomalies (departures from a background state) is well approximated by linear equations of the form

$$\frac{dx}{dt} = A\,x + f_{ext} + B\,\eta - D \qquad (1)$$

where $x(t)$ is an N-component anomaly state vector, $f_{ext}(t)$ is an N-component external forcing vector, η is an M-component noise vector of independent Gaussian white noises with zero mean, $A(t)$ and $B(t)$ are N**x**N and N**x**M matrices, and $D(t) = \langle B\eta \rangle$ is an N-component expected mean noise forcing vector (which we retain for future reference even though it is strictly zero here). The equations for all classical free and forced linear wave dynamics in the climate system may be cast in this form, as may also those for the evolution of small-amplitude forecast errors, important in modern data assimilation techniques. The traditional use of linear models in such contexts has advanced both basic understanding and practical applications, but has stopped short of claiming that observed full-amplitude anomalies also obey such equations. Evidence has steadily accumulated, however, to support even this latter stronger claim, especially for "coarse-grained" anomalies averaged over various time and space scales. This evidence has come from a wide range of studies demonstrating the approximate linearity of the global climate response to combinations of radiative forcings (e.g. Knutson et al 2006, and references); the approximate linearity of the global atmospheric response to tropical SST changes (e.g., Barsugli and Sardeshmukh 2002, Schneider et al 2003, Barsugli et al 2006); the

approximately linear dynamics of seasonal tropical SST anomalies (Penland and Sardeshmukh 1995); the competitiveness of simple linear seasonal forecast models with global coupled climate models (Saha et al 2006); the approximately linear evolution of weekly-averaged atmospheric circulation anomalies (Winkler et al 2001, Newman and Sardeshmukh 2008); and the competitiveness of Week 2 and Week 3 linear forecast models with comprehensive numerical weather prediction (NWP) models (Winkler et al 2001, Newman et al 2003). Even on the time scales of daily weather, linear stochastically forced (LSF) models of the form (1), although not as accurate as NWP models for daily predictions, are realistic enough to capture many features of the second-order statistics of observed synoptic variability, such as the geographical structures of eddy variances and covariances and momentum and heat fluxes (e.g., Farrell and Ioannou 1995, Hall and Sardeshmukh 1998, Whitaker and Sardeshmukh 1998, DelSole 2004).

In light of these and many other studies, the relevance of LSF dynamics even in the chaotic nonlinear climate system seems undeniable. Indeed without it, the plethora of diagnostic studies of weather and climate variations (and of weather and climate model errors) based on linear regressions and correlations would have limited value. The basic premise in (1) concerning the dynamics of "coarse-grained" anomalies is that the coarse-grained nonlinear tendency terms, associated primarily with fluxes by unresolved eddies, can in principle be linearly parameterized in terms of the coarse-grained anomalies, and the unparameterized remainder can be treated as stochastic white noise. It is important to recognize that the matrix A in (1) is therefore in general not that obtained by directly linearizing the governing equations but also includes such linear flux parameterizations, and the matrix B accounts for the amplitude and correlation structure of the

unparameterized remainder as a "stochastic parameterization". Procedures for estimating these matrices directly from data, as well as for testing the validity of (1), are called Linear Inverse Modeling (Penland 1989), and are discussed in detail in Penland and Ghil (1993), Penland and Matrosova (1994), Penland and Sardeshmukh (1995) and Winkler et al (2001).

An attractive feature of the LSF approximation (1) is that the moment equations for the evolving probability density functions (PDFs) are closed, i.e. equations for the higher-order moments involve moments of the same or lower order. Specifically, the equations for the first moment $<x>(t)$ and the second moment $C(t) = <x\,x^T>$ are

$$\begin{aligned} \frac{d}{dt} <x> &= A <x> + f_{ext} \\ \frac{d}{dt}\, C &= A\,C + C\,A^T + Q + <x> f_{ext}^T + f_{ext} <x^T> \end{aligned} \tag{2}$$

where $Q = B\,B^T$, and angle brackets $<\ >$ denote expected values. The centered second moment $C' = <x'x'^T>$ of the departures $x' = x - <x>$ is related to C as $C' = C - <x><x^T>$. Note that these linear equations for $<x>$ and C are applicable to the moments of the marginal (i.e. unconditional) PDF $p(x)$ as well as the conditional PDF $p(x(t)|x(0))$ of $x(t)$ given $x(0)$, that is, to the moments of the observed as well as forecast probability distributions. They can therefore be used to model and predict those PDFs without having explicitly to generate ensembles of integrations as in traditional ensemble forecasting techniques using GCMs. In the simplest scenario in which A, B, and f_{ext} are constant, the stationary solutions to (2) may be written as

$$
\begin{aligned}
< x > \; &= \; - A^{-1} f_{ext} \\
\frac{dC}{dt} \; &= \; 0 \; = \; A\,C \; + \; C\,A^T \; + \; Q + \; < x > f_{ext}^T \; + \; f_{ext} < x^T >
\end{aligned}
\tag{3a}
$$

for the marginal moments, and

$$
\begin{aligned}
\hat{x}'(t) \; &\equiv \; < x'(t) \mid x'(0) > \; = \; e^{At} x'(0) \\
\hat{C}(t) \; &\equiv \; < (\hat{x}' - x')\,(\hat{x}' - x')^T > \; = \; C' - e^{At} C' e^{A^T t}
\end{aligned}
\tag{3b}
$$

for the conditional (i.e. forecast) moments of the departures x' of x from the mean forced anomalous state <x> in (3a). If x is Gaussian, then given that Gaussian PDFs are characterized completely by their first and second moments, these equations (and more generally, equations (2)) provide a complete description of system variability and predictability. This, together with the fact that LSF models are also typically of vastly lower dimension (N < 100, M < 100) than GCMs provides a strong incentive to determine what aspects of the real nonlinear climate system can and cannot be captured by such models.

LSF models (1) are consistent with Gaussian statistics. The reasons for this are two-fold. First, because of averaging, the PDFs of the coarse-grained anomalies for which (1) is appropriate are approximately Gaussian. This is a direct consequence of the Central Limit Theorem, which dictates that the probability distribution of an average of a sufficiently large number (in practice, often less than 30) of independent and identically distributed but not necessarily Gaussian variables is approximately Gaussian. The statistics of monthly and longer averages in the climate system are indeed approximately Gaussian (Stephenson et al 2004, Penland and Sardeshmukh 1995). Second, since any linear

combination of Gaussian variables is also strictly a Gaussian variable, the dynamics of such variables are consistent with LSF dynamics. Thus if $x(t)$ in (1) is Gaussian, and $f_{ext}(t)$ is either constant or Gaussian, then dx/dt and therefore $x(t+\Delta t)$ is Gaussian, and a dynamical evolution consistent with Gaussian statistics is attained. In a nonlinear system in which A depends on x, dx/dt is not Gaussian even if $x(t)$ is Gaussian, and an evolution consistent with Gaussian statistics is not guaranteed.

Gaussian statistics thus imply LSF dynamics. But do non-Gaussian statistics necessarily imply nonlinear dynamics? In particular does the existence of skewness, demonstrating an asymmetry in the statistics of opposite-signed anomalies, necessarily establish the nonlinearity of the underlying dynamics? This is our primary concern in this paper. The issue is not only of fundamental but also practical interest. For example, even if LSF models are competitive with nonlinear GCMs at representing observed second-order statistics and second-order measures of forecast performance such as r.m.s. errors and anomaly correlations, one may still wonder if they can remain so at representing the higher-order moments of the marginal and forecast probability distributions. In particular, one may wonder if they are capable of representing the tails of the marginal and forecast distributions, and therefore the likelihood of extreme weather and climate events.

The issue would be moot if the PDFs of the observed circulation were Gaussian. As mentioned above, the PDFs of monthly and longer-term averages are almost Gaussian, but the PDFs of the less coarse-grained weekly averages are appreciably non-Gaussian (Sura et al 2005, and references), and those of daily averages are even more so. Figure 1 shows the skew S and (excess) kurtosis K of the observed daily-averaged 300 mb

vorticity in the northern winters of 1970-1999. Both quantities are large in the hemispheric jet stream waveguide (Hoskins and Ambrizzi 1993, Borges and Sardeshmukh 1995, Branstator 2002) and have a coherent geographical structure. (For reasons that will become clearer below, detailed assessments of the statistical significance of such higher moments are not a major concern of this paper. We note in passing that similar patterns of K and S were obtained using subsets of the data and also by White (1980) using a different and smaller dataset). Figure 2 displays the results of Figure 1 in the form of a scatter plot. A remarkable tendency toward a parabolic relationship between K and S is evident. Similar remarks may also be made concerning the non-Gaussian character of, for instance, sea surface temperature (SST) variability in the eastern tropical Pacific: the PDFs are almost Gaussian for 3-month averages (Penland and Sardeshmukh 1995), appreciably non-Gaussian for monthly averages (Hannachi et al 2003), and substantially non-Gaussian for daily averages, with a similar remarkable tendency toward a parabolic K-S relationship (Sura and Sardeshmukh 2008).

LSF models of the form (1) *can* generate non-Gaussian statistics, but only if $f_{ext}(t)$ is non-Gaussian. One may expect some non-Gaussianity, for example, from slow non-Gaussian variations of natural and anthropogenic radiative forcings. Similarly, insofar as tropical SSTs may be considered as "forcing" the extratropical circulation, the PDFs of the extratropical circulation may be influenced by the non-Gaussianity of the tropical SSTs. Such mechanisms, however, do not solve the problem of explaining the non-Gaussianity of x but merely shift it to explaining the non-Gaussianity of f_{ext}. Since our concern here is with the implications of non-Gaussian statistics for the linearity or nonlinearity of the

internal system dynamics, we will henceforth ignore such external generators of non-Gaussian variability and assume that f_{ext}. is either constant or Gaussian.

Our goal in this paper is to show that non-Gaussian statistics can be reconciled with LSF dynamics through a relatively minor extension of (1): by allowing the stochastic forcing amplitude matrix B to depend linearly on the system state. Specifically, we consider the implications of its elements B_{im} being of the form

$$B_{im}(x,t) = G_{im}(t) + \sum_j E_{ijm}(t)\, x_j \quad . \tag{4a}$$

In the following, we will refer to the stochastic forcing $B\eta$ as being purely "additive" if E is zero, and purely "multiplicative" if G is zero. If G_{im} and E_{ijm} are nonzero but not for the same noise component η_m, we have an uncorrelated mixture of additive and multiplicative noises. If G_{im} and E_{ijm} are both nonzero for some noise components η_m, we have correlated additive and multiplicative ("CAM") noise. An important aspect of CAM noise forcing is that its expected mean "noise-induced drift" D is not zero in (1), but as shown below in section 2, is related to the noise parameters as

$$D_i(t) = \frac{1}{2}\sum_j \sum_m E_{ijm}(t) G_{jm}(t) \quad . \tag{4b}$$

We will show how CAM noise occurs naturally in a quadratically nonlinear dynamical system, such as the climate system, through terms involving x and rapidly decorrelating system components that may be approximated as noise.

We will see that additive plus uncorrelated multiplicative noise can produce symmetric non-Gaussian, but not asymmetric (i.e. skewed) PDFs. To generate asymmetric PDFs, one must have CAM noise. We will also see how this necessitates modifying the moment equations (2) due to the so-called "noise-induced drift" (of which *D(t)* represents the expected mean), but the equations remain linear and closed as before. However, since the modified stochastic forcing can generate non-Gaussian statistics, those moment equations no longer provide a complete description of system variability and predictability. We will seek some insight into the higher moments by investigating the simplest LSF system with CAM noise, a 1-d system (with N=1, M=2) of the form

$$\frac{dx}{dt} = Ax + b\,\eta_1 + (Ex+g)\,\eta_2 - \frac{1}{2}Eg \quad , \tag{5}$$

in which A, b, E and g are all scalar constants, and the last term on the right hand side explicitly represents the mean noise-induced drift as in (4b). Note that A is negative in (5), and without loss of generality, b and E are positive; however g can be positive or negative. We will show that the higher moments of this system are interrelated in a remarkably simple way, and can account for the parabolic *K-S* relationship in Figure 2. Further, we will derive the full PDF of this system from the corresponding Fokker-Planck equation and show that in addition to being skewed, it has power-law tails, whose existence is also often associated with nonlinear dynamics. We will then demonstrate the relevance of this generic univariate "Stochastically-Generated-Skewed (SGS)" PDF even in the real multivariate climate system.

Finally, we will seek to clarify the physical nature of the stochastic noise responsible for the observed non-Gaussian circulation statistics. We are especially interested in determining if it is associated primarily with adiabatic or diabatic noise, i.e. with turbulent adiabatic fluxes or rapid diabatic forcing variations. To this end we will examine a long 108000-day perpetual winter simulation (equivalent to 1200 90-day winters) generated by Sardeshmukh and Sura (2007) using a dry adiabatic GCM forced only with the observed time-mean diabatic forcing as a constant forcing. We will assess to what extent this constant-forcing simulation captures the non-Gaussian statistics shown in Figs 1 and 2.

The paper is organized as follows. We begin in section 2 with a derivation of the moment equations for LSF systems with the extended stochastic forcing (4), and highlight the necessity of CAM noise to generate skew. In this context we also note an inconsistency in the explanation of the skew of weekly-averaged circulation anomalies offered by Sura et al (2005) in terms of pure multiplicative noise. In Section 3 we discuss how the existence of CAM noise may be justified in the climate system with quadratic nonlinearities and "slow" and "fast" system components. Section 4 follows with a detailed analysis of the generic 1-d system (5) with CAM noise. Section 5 presents results from the long adiabatic GCM simulation and compares them with observations. To understand the remarkable consistency of the observational and GCM-simulated higher-order statistics with those of the generic 1-d system, we introduce in section 6 a principle of increasing "diagonal dominance" in the higher-order moment equations of multi-component LSF systems. Concluding remarks, including a brief discussion of how the 1-d approximation

may be exploited to estimate the probabilities of extreme weather and climate anomalies, follow in section 7.

2. Moment equations for the extended system

For any dynamical system of the form

$$\frac{dx}{dt} = \mathcal{A}(x,t) + \mathcal{B}(x,t)\,\eta \quad , \tag{6}$$

where $\mathcal{A}$ is an N-component vector, $\mathcal{B}$ is an N**x**M matrix, and η is an M-component vector of Gaussian white noises (in the Stratanovich sense) that are independent and delta-correlated in time as $<\eta_m(t)\eta_m(t')> = \delta(t-t')$, the Fokker-Planck Equation (FPE) for the evolution of the probability density $p(x,t \mid x_0,t_0)$ may be written

$$\frac{\partial p}{\partial t} = -\sum_i \frac{\partial}{\partial x_i}\left[\left(\mathcal{A}_i + \frac{1}{2}\sum_j\sum_m \frac{\partial \mathcal{B}_{im}}{\partial x_j}\mathcal{B}_{jm}\right)p\right] + \frac{1}{2}\sum_i\sum_j\sum_m \frac{\partial^2}{\partial x_i \partial x_j}\left(\mathcal{B}_{im}\mathcal{B}_{jm}p\right). \tag{7}$$

For convenience we will henceforth utilize Einstein's notational convention of assuming summation over repeated indices; the cumbersome summation signs then become redundant. This enables (7), for instance, to be written in the more compact form

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x_i}\left[\left(\mathcal{A}_i + \frac{1}{2}\frac{\partial \mathcal{B}_{im}}{\partial x_j}\mathcal{B}_{jm}\right)p\right] + \frac{1}{2}\frac{\partial^2}{\partial x_i \partial x_j}\left(\mathcal{B}_{im}\mathcal{B}_{jm}p\right) \quad . \tag{8}$$

For a linear stochastically forced system with $\mathcal{A}(x,t) = A(t)x + f_{ext}(t) - D(t)$ as in (1) and $\mathcal{B}(x,t) = B(x,t)$ as in (4), the FPE is

$$\begin{aligned}\frac{\partial p}{\partial t} &= -\frac{\partial}{\partial x_i}\left\{\left[\left(A_{ij} + \frac{1}{2}E_{ikm}E_{kjm}\right)x_j + \cancel{\frac{1}{2}E_{ijm}G_{jm}} + \left(f_{ext}\right)_i - \cancel{D_i}\right]p\right\} + \frac{1}{2}\frac{\partial^2}{\partial x_i \partial x_j}\left(B_{im}B_{jm}p\right) \\ &= -\frac{\partial}{\partial x_i}\left\{\left[M_{ij}x_j + \left(f_{ext}\right)_i\right]p\right\} + \frac{1}{2}\frac{\partial^2}{\partial x_i \partial x_j}\left(B_{im}B_{jm}p\right)\end{aligned} \tag{9}$$

where $M_{ij} = A_{ij} + 0.5E_{ikm}E_{kjm}$. From this the equations for the first two (marginal as well as conditional) moments of x may be derived as

$$\begin{aligned}\frac{d}{dt}<x> &= M<x> + f_{ext} \\ \frac{d}{dt}C &= MC + CM^T + \tilde{Q} + <x>f_{ext}^T + f_{ext}<x^T>\end{aligned} \tag{10}$$

where $\tilde{Q}_{ij} = G_{im}G_{jm} + E_{ikm}C_{kl}E_{jlm}$. One can now see why D in (1) must be of the form (4b), since any other choice results in $<x>$ being nonzero even in the absence of external forcing.

Note that equations (10) are of identical form to (2), except that A and Q are replaced by M and $\tilde{Q}$. In the simplest scenario in which A, G, E and f_{ext} are constant, the solutions to (10) are also identical to those in (3a) and (3b), with A and Q again replaced by M and $\tilde{Q}$, except that the equation for the growth of the forecast error covariance $\hat{C}(t)$ (the second equation in (3b)) can no longer be expressed in the same elegant analytic form.

The extension to state-dependent noise in (4) thus preserves the linear and closed character of the equations for the first and second (and also higher-order) moments. Crucially, the extended system *still responds linearly to external forcing*, and *the prediction of the expected future state given an initial state is still a linear prediction*. This extension is thus completely consistent with all the accumulated evidence cited in the previous section in support of the LSF approximation. However, it now also allows for the representation of non-Gaussian statistics, especially of odd moments such as skew. This is most easily understood by revisiting (4). For pure additive or multiplicative noise, or for any uncorrelated mixture of the two, the magnitude of the stochastic forcing B is symmetric with respect to the sign of x; there is therefore no mechanism in (1) to generate skew in the absence of external forcing. For CAM noise, however, the magnitude of B is not symmetric with respect to the sign of x. This introduces an asymmetry in (1) and can generate skewed statistics even in the absence of external forcing.

To account for skewed statistics in a linear framework, one therefore needs CAM noise. In this context it is interesting to recall the study of Sura et al (2005), who proposed an explanation of the skew of observed wintertime 7-day running mean tropospheric circulation anomalies in terms of pure multiplicative noise. Specifically, they sought to understand the departure from Gaussianity of the joint PDF of the two dominant EOFs of 750 mb streamfunction (see Figure 3) in terms of the statistics of the least damped eigenmode of the barotropic vorticity equation linearized about the long-term mean flow (Borges and Sardeshmukh 1995), if the mode is steadily forced and stochastically

damped (i.e. if stochastic perturbations are introduced in its damping rate). Their model may be expressed in our notation as

$$\frac{dx}{dt} = A\,x + f_{ext} + Ex\eta \quad , \tag{11}$$

where x is a 2-component vector whose components x_1 and x_2 represent the amplitudes of the real and imaginary parts of the least damped eigenmode, A is a constant $2\mathbf{x}2$ matrix with elements $[\ -r\ -\omega\ |\ \omega\ -r\ ]$ corresponding to the real and imaginary parts of the corresponding eigenvalue, and E *is* a $2\mathbf{x}2$ identity matrix multiplied by the amplitude of the scalar Gaussian white noise η, Sura et al experimented with various ad hoc choices of f_{ext} before finding a joint PDF of x_1 and x_2 whose departures from Gaussianity capture the essence of the observed departures, as shown in Figure 3. The result is indeed remarkable, especially given the simplicity of (11). However, it is not internally consistent. Briefly, the model cannot generate skew without f_{ext}, since (11) is then exactly symmetric with respect to the sign of x, but with a nonzero f_{ext}, the expected mean anomaly $<x>$ cannot be zero, because (10) implies

$$< x > \ = \ -M^{-1}\ f_{ext}\ , \qquad \text{where} \quad M \ = \ [A + 0.5E^2] \quad . \tag{12}$$

Such a pure multiplicative noise model cannot therefore explain the skew of *centered* anomalies with zero mean, as in the observational panel of Figure 3. The difficulty does not arise in a CAM noise model, since it can generate skew even in the absence of external forcing.

3. Justification of CAM noise

The existence of CAM noise can be anticipated in any quadratically nonlinear dynamical system with "slow" and "fast" system components, usually associated with relatively long and short correlation time scales. To see this, consider the evolution equation for the full state vector X in the form

$$\frac{dX_i}{dt} = L_{ij}X_j + N_{ijk}X_jX_k + F_i \quad , \tag{13}$$

where X_i is the *i*th component of *X*, the first and second terms on the right are the linear and quadratically nonlinear adiabatic tendencies (in which we include linear and quadratically nonlinear damping terms), and all other tendencies are represented by the "external" forcing F_i. Writing X_i as a sum of mean and anomaly parts, $X_i = \overline{X}_i + X_i'$, the equation for the anomalies may be expressed as

$$\frac{dX_i'}{dt} = [L_{ij} + (N_{ijk} + N_{ikj})\overline{X}_k]\, X_j' \;+\; N_{ijk}(X_j'X_k' - \overline{X_j'X_k'}) \;+\; F_i' \quad .$$

Let *X* consist of slow components *x* and fast components *y*, so that $X'^T = [x'^T \;\; y'^T]$ and $\overline{X}^T = [\overline{x}^T \;\; \overline{y}^T]$. Then the equation for the anomalous slow components may be written

$$\begin{aligned}\frac{dx_i'}{dt} \;=\;& [L_{ij} + (N_{ijp} + N_{ipj})\overline{y}_p]\, x_j' \\ &+\; [(N_{ijp} + N_{ipj})x_j' + \{L_{ip} + (N_{ijp} + N_{ipj})\overline{x}_j\}]\, y_p' \;-\; (N_{ijp} + N_{ipj})\, \overline{x_j'y_p'} \\ &+\; N_{ipq}(y_p'y_q' - \overline{y_p'y_q'}) \;+\; N_{ijk}\,(x_j'x_k' - \overline{x_j'x_k'}) \;+\; f_i'\end{aligned} \tag{14}$$

Now let $N_{ijp} + N_{ipj} = E_{ijp}$, $L_{ip} + (N_{ijp} + N_{ipj})\overline{x}_j = G_{ip}$, and $(N_{ijp} + N_{ipj})\overline{x'_j y'_p} = D_i$. Further, approximate $N_{ipq}(y'_p y'_q - \overline{y'_p y'_q})$ as $H_{ij} x'_j + G_{ir} z_r$, where H and G are linear operators and z is a noise vector, and where the index r uniquely identifies each combination of p and q. Finally, let $L_{ij} + (N_{ijp} + N_{ipj})\overline{y}_p + H_{ij} = A_{ij}$. Then neglecting the nonlinear terms involving $x'x'$ in (14), we obtain

$$\frac{dx'_i}{dt} = A_{ij} x'_j + (E_{ijp} x'_j + G_{ip})\, y'_p - D_i + G_{ir} z_r + f'_i \qquad (15)$$

Dropping the primes on x' for clarity, defining $f_{ext} = f_i'$, and also defining an extended noise vector $\eta^T = [y'^T \; z^T]$ with M = P+R components, we finally arrive at the form

$$\frac{dx_i}{dt} = A_{ij} x_j + (E_{ijm} x_j + G_{im})\, \eta_m - \frac{1}{2} E_{ijm} G_{im} + (f_{ext})_i \qquad (16)$$

which is identical to (1) with B and D as in (4).

It is important to appreciate that it is approximating the fast variables y' and z in (15) as *stochastic noise* η in (16) that enables the mean noise-induced drift D_i in (15) to be represented as in (4b), and to close the moment equations as in (10). Otherwise the slow-fast variable separation inherent in (6) and the FPE are not valid. The conditions under which the components of a dynamical system may be separable into slow and "fast enough" components in this sense, as well as procedures for classifying specific system components as such, have been the subject of many theoretical and empirical studies (e.g. Khas'minskii 1966, Papanicolaou and Kohler 1974, Hasselmann 1976, Penland 1996,

Winkler et al 2001, Majda et al 2003, Gardiner 2004, Franzke et al 2005). Our intention is not to pursue a similar specific classification here but merely to highlight how CAM noise occurs naturally in a quadratically nonlinear system with a slow-fast separation of time scales. Indeed the above considerations make it easier to justify CAM noise than either pure multiplicative noise or uncorrelated additive and multiplicative noise.

In the climate system, the quadratically nonlinear tendency terms are mostly associated with adiabatic fluxes, and the "external" forcing term F in (13) represents a combination of internal diabatic interactions and truly external forcing. In reality, therefore, F also depends on the system state. For small perturbations around $\bar{X}$, we may write

$$F_i' \;=\; F_i(X) - \bar{F}_i \;\simeq\; (f_{ext})_i \;+\; \left.\frac{\partial F_i}{\partial X_j}\right|_{\bar{X}} X_j' \;+\; \frac{1}{2}\left.\frac{\partial^2 F_i}{\partial X_j \partial X_k}\right|_{\bar{X}} (X_j' X_k' - \overline{X_j' X_k'}) \qquad (17)$$

Decomposing X' into slow and fast components and following a similar development to that from (14) onward, one can see how F' can also give rise to CAM noise components. We will present evidence in section 5 that this source of CAM noise is, however, relatively minor compared to that associated with rapidly varying adiabatic fluxes for the generation of the non-Gaussian statistics shown in Figures 1 and 2.

4. A generic 1-D linear system with CAM noise

Having seen that the extended stochastic forcing (4) provides a mechanism for a linear system to have non-Gaussian statistics, we now address the issue of what specific types of non-Gaussian statistics it can generate. This can be helpful in discriminating between

this and nonlinear dynamical mechanisms of non-Gaussian variability. We are especially interested in determining if CAM noise can account for the parabolic *K-S* relationship in Fig 2. To this end we consider the simplest possible 1-d system with CAM noise, as in (5). The integrated FPE for that system may be written

$$[\, Mx \,]\, p \;=\; \frac{1}{2}\,\frac{d}{dx}\,[\,(E^2x^2 + 2Egx + g^2 + b^2)\,p\,] \quad , \tag{18}$$

where $M = A + 0.5\,E^2 < 0$. Equation (18) may be used to obtain expressions for the *n*-th order moments $< x^n >$ by multiplying by x^{n-1} and integrating over *x*. For the first two moments, this yields $< x > \;=\; 0$ and $< x^2 > \;=\; \sigma^2 \;=\; -(g^2 + b^2)/(2M + E^2)$. For the higher moments, we obtain

$$\left[M + \left(\frac{n-1}{2}\right)E^2\right] < x^n > \;=\; -\left(\frac{n-1}{2}\right)\left[2Eg < x^{n-1} > + (g^2 + b^2) < x^{n-2} >\right] . \tag{19}$$

Note that in general, $\left[M + 0.5(n-1)E^2\right]$ is negative for $< x^n >$ to exist. Using (19) to obtain expressions for $< x^3 >$ and $< x^4 >$, and remembering that the skew *S* and excess Kurtosis *K* are defined as $S \;=\; < x^3 > /\sigma^3$ and $K \;=\; < x^4 > /\sigma^4 \;-3$, gives after some manipulation,

$$K \;=\; \frac{3}{2}\left[\frac{1+\alpha}{1+(3/2)\alpha}\right] S^2 \;+\; 3\left[\frac{1+(1/2)\alpha}{1+(3/2)\alpha} \;-\; 1\right] \;\geq\; \frac{3}{2}\,S^2 \quad , \tag{20}$$

where $\alpha = E^2 / M < 0$. The inequality in (20) exists because both ratios involving α are greater than 1 (see also Sura and Sardeshmukh 2008). Note that the *equality* $K = 1.5\, S^2$ is satisfied only for $E = 0$, in which case it reduces to a triviality $0 = 0$.

Thus, *regardless of the model parameters*, K exceeds 1.5 S^2 in a 1-d LSF system with CAM noise. This is a simple and specific prediction of the character of non-Gaussian variability. And remarkably, the points in Figure 2 satisfy this K-S inequality with the same parabolic dependence of K on S, albeit with a small negative bias.

One can similarly make specific predictions for the other moments using (19). For example, the fifth moment of x must satisfy

$$\mu_5 \equiv \frac{< x^5 >}{\sigma^5} \quad \begin{cases} > \; 10S + 3S^3 \;\text{ for }\; S > 0 \\ < \; 10S + 3S^3 \;\text{ for }\; S < 0 \end{cases} \tag{21}$$

We will attempt to verify this prediction of the 1-d model in section 5.

Finally, we consider the stationary probability density of the 1-d system (5), obtained by solving for $p(x)$ in (18) as

$$p(x) \;=\; \frac{1}{\mathcal{N}} \left[(Ex+g)^2 + b^2 \right]^{\frac{1}{\alpha}-1} \exp\left[-\frac{2g}{\alpha b} \arctan\left(\frac{Ex+g}{b} \right) \right], \tag{22}$$

where $\mathcal{N}$ is a normalization constant which ensures that $p(x)$ integrates to unity (See also van Kampen 1981, Muller 1987). This PDF is clearly skewed if $g \neq 0$, and has a unique

maximum at $x_{\max} = Eg/(M - E^2)$. We will refer to it as the generic "Stochastically-Generated-Skewed (SGS)" probability density function. In addition to skew, the SGS distribution also has power-law tails, since for large magnitudes of x we have

$$p(x) \sim x^{2\left(\frac{1}{\alpha}-1\right)} \exp\left[-\frac{2g}{\alpha b} \arctan\left(\frac{Ex}{b}\right)\right] \sim x^{2\left(\frac{1}{\alpha}-1\right)} \exp\left(\mp \frac{\pi g}{\alpha b}\right), \qquad (23)$$

where the $-(+)$ sign in the exponential factor obtains for large positive (negative) x. Remembering that α is negative, this means that if g is positive, p is greater for large positive x than for large negative x, but has the same exponent of the power-law tail. We will attempt to verify this specific prediction of the 1-d model also in the next section.

Before ending this section, we note that (5) is the simplest but not the only possible stochastically forced 1-d system with deterministic linear dynamics that can have non-Gaussian statistics. The most general such system is the linear system

$$\frac{dx}{dt} = Ax + \sum_m \sqrt{[(E_m x + g_m)^2 + c_m x]}\ \eta_m - \frac{\beta}{2} + f_{ext} \qquad (24)$$

perturbed by so-called "radical noise", where $\beta = \sum\left(E_m g_m + 0.5 c_m\right)$, and where to avoid confusion we have discontinued the use of the Einstein summation convention. Equation (5) is a special case of (24) with $c_m \equiv 0$, $f_{ext} \equiv 0$, and 2-d noise with components η_1 and η_2. Defining $E^2 = \sum E_m^2$, $G^2 = \sum g_m^2$, and $M = A + 0.5E^2 < 0$, the integrated Fokker-Planck equation (18) is modified in this more general case to

$$\left[Mx + f_{ext}\right] p \;=\; \frac{1}{2}\,\frac{d}{dx}\,[\;(E^2x^2 + 2\beta x \;+ G^2)\,p\;] \qquad (25)$$

Note that $< x > \;=\; -M^{-1} f_{ext}$, so the system still responds linearly to external forcing, and the prediction of the expected future state given an initial state is still a linear prediction. To describe the dynamics of centered anomalies with zero mean, we set $f_{ext} = 0$, as in (5). Following a development very similar to that from (19) to (22), one can then derive similar equations for the relationships between the moments, and also solve (25) directly for the PDF (see Appendix). Importantly, the *K*-*S* relationship $K \;\geq\; 1.5\,S^2$ and the $\mu_5 - S$ relationship (21) remain valid in all cases, but the PDF can differ from (22). Our chief motivation for introducing this more general process (24) here, however, is to point out that if there is no multiplicative noise, i.e. if $E \equiv\; 0$, then the *K*–*S* relationship becomes an *equality* $K \;\equiv\; 1.5\,S^2$, which, unlike (20), is valid even for nonzero values of *K* and *S*.

If both $E \equiv\; 0$ and $G \equiv\; 0$, the solution of (25) with $f_{ext} = \beta / 2$ is a Gamma pdf with a shape parameter 1/2 and a scale parameter $-\beta / M$. This result may come as a surprise, since Gamma PDFs are usually associated with the squares of Gaussian variables. Nonetheless, it could have been readily anticipated from (1). In the 1-d case, for constant model parameters and no external forcing in (1), the equation for the square of the state variable in (1) can be cast in the form (24) with $E = 0$, $G = 0$, and $f_{ext} = \beta / 2$.

Although (24) incorporates interesting extensions of the simple model (5) *in the 1-d case*, we have not pursued them further in this study even though the deterministic dynamics

remain linear. This is mainly because it is difficult to justify the relevance of full-fledged "radical noise" in the *N-d* climate system from first principles. As the previous section showed, it is easier to justify CAM noise, which is a special case of radical noise, given the importance of quadratic nonlinearities in the *N-d* climate system. One can nonetheless imagine the general 1-d linear model (24) being useful in many other contexts than the one considered in this study.

5. Results from a long dry adiabatic GCM simulation with constant forcing

Are the skewness and kurtosis of the daily 300 mb vorticity anomalies shown in Fig. 1 due to CAM noise, and if so, are they associated primarily with turbulent adiabatic or diabatic forcing fluctuations? To clarify this, we examine a long 1200-winter simulation of the northern winter climate generated by Sardeshmukh and Sura (2007) using a dry adiabatic atmospheric general circulation model forced only with the observed long-term winter-mean diabatic forcing as a constant forcing. The model has a T42 spatial discretization in the horizontal and 5 levels in the vertical, and is exactly of the form (13), but with a prescribed *constant* forcing $\overline{F}_i$ estimated from observations. Despite the complete neglect of forcing variations F_i', the model reproduces many second-order statistics of the observed atmospheric circulation variability as described fully in Sardeshmukh and Sura (2007). Figure 4, shown in an identical format to Figure 1 but constructed from the model output, shows that the model also captures many features of the observed skewness and kurtosis of the daily 300 mb vorticity anomalies. There are some notable areas of discrepancy, especially over the Atlantic sector; to what extent this is due to the neglect of F_i' or the coarse spatial resolution of the model is unclear at

present. Nonetheless, the generally successful simulation of the essential character of the observed *S* and *K* fields despite the model's neglect of transient diabatic forcing suggests that the stochastic generators of these higher-order vorticity moments are mostly associated with turbulent adiabatic fluxes. Furthermore, the fact that the model also captures the observed parabolic *K*-*S* relationship in Fig 5, consistent with the 1-d theory, suggests a dominant role for the *local* adiabatic fluxes in generating these moments.

One advantage of examining a long 1200-winter simulation is that one can have much greater confidence in the statistical significance of the higher-order statistics. Figure 6 attempts to verify the relationship (21) between the fifth moments and skewness predicted by the 1-d theory. (In view of the enormous sampling uncertainties involved, we did not attempt to do this with our 30-winter observational dataset). Results are shown for the simulated daily 300 mb vorticity as well as the 500 mb geopotential height anomalies at all northern hemispheric gridpoints. Figure 6 clearly bears out the prediction of the 1-d theory even in the multi-component GCM simulation, which again highlights not only the relevance of CAM noise but also the dominance of the *local* stochastic dynamics in generating the higher moments. This is a powerful validation of the linear 1-d theory.

Figure 7 verifies another major prediction of the 1-d theory, that the PDFs must have power-law tails. Here we also attempted a comparison with observations, at two north Pacific locations of the largest skew of 300 mb vorticity and 500 mb heights. Our hope was that the relatively large deviations from Gaussianity at those locations might generate

more statistical confidence in the character of the estimated PDF tails; however we did not attempt to put error bars on those tails. The observational PDFs in the left panels of Fig 7 do appear to have power law tails, at least on the “fat” tail side. The right panels show the corresponding PDFs from the model simulation, now with error bars. They clearly have power-law tails, that remarkably have the same slope as the observed for the vorticity PDF, and only a slightly steeper slope than the observed for the geopotential height PDF. The model’s power-law tail extends on the "fat" tail side to values of x up to 7 standard deviations. On the “thin” tail side, the probability densities are so low as to be a challenge to estimate even from an 1200-winter long simulation. Still, a hint of a power-law dependence, at least for the 300 mb vorticity, is evident in the lower right panel of Figure 7, and with the same slope as on the “fat” tail side, as predicted by the linear 1-d theory.

6. A principle of “Diagonal Dominance” in the higher-order moment equations

The success of the local 1-d model (5) in explaining the essential character of the observed and GCM-simulated non-Gaussian statistics may come as a surprise, given the obvious importance of nonlocal dynamics in the multivariate climate system. The key point, however, is that this success applies to the understanding and simulation of the *higher-order* non-Gaussian statistics and power-law tails. We argue below that the 1-d model (5) becomes progressively better at representing the higher-order statistics of multivariate systems through a principle of increasing “diagonal dominance” in the higher-order moment equations.

Diagonal Dominance refers to the progressively greater importance of the *self-correlation* terms in the higher-order moment equations of multivariate systems. Consider, for illustrative purposes, (1) and (4) with $f_{ext} = 0$ and time-independent A, G, E, and D, and also in a space in which each system component is normalized by its standard deviation so that its marginal PDF is a standard Gaussian with unit variance. The equation for the n-th moment $< x_i^n >$ of the i-th component of x, obtained by multiplying (1) by x_i^{n-1} and taking expectation values, is then

$$\begin{aligned} \frac{1}{n}\frac{d}{dt} < x_i^n > \; &= \; 0 \\ = \quad & A_{ii} < x_i^n > \; + \sum_{\mathrm{j}\neq \mathrm{i}} A_{ij} < x_i^{n-1} x_j > \\ + \sum_{\mathrm{m}} G_{im} < x_i^{n-1} \eta_m > \; + \; & \sum_{\mathrm{m}} \sum_{\mathrm{j}} E_{ijm} < x_i^{n-1} x_j \eta_m > \; - \; < x_i^{n-1} > D_i \end{aligned} \tag{26}$$

where for added clarity we have discontinued the use of the Einstein summation convention. The first term on the right hand side of (26) is a self-correlation term. The second term involves correlations between powers of x_i and other system components, whose magnitudes are generally smaller than unity and become small for large n. To see this, consider the regression $x_i = \rho_{\mathrm{ij}} x_j + \varepsilon_{\mathrm{ij}}$ of x_i on x_j, where $|\rho_{ij}| < 1$ and $\varepsilon_{\mathrm{ij}}$ is uncorrelated with x_j. Then $< x_i^{n-1} x_j > \; = \; < (\rho_{\mathrm{ij}} x_j + \varepsilon_{\mathrm{ij}})^{n-1} x_j > \; \approx \; \rho_{ij}^{n-1} < x_j^{\,n} >$, which approaches zero for large n. The second term in (26) is thus small for large n. By a similar argument, the contributions to the fourth term from cross-correlations between x_i and $x_{j\neq i}$ are also small for large n. This enables the third and fourth terms to be approximated and combined as follows:

$$\sum_{\mathrm{m}} \left\{ G_{im} < x_i^{n-1} \eta_m > + \sum_{\mathrm{j}} E_{ijm} < x_i^{n-1} x_j \eta_m > \right\}$$

$$= \quad < x_i^{n-1} \sum_m \left\{ G_{im} + \sum_j E_{ijm} x_j \right\} \eta_m > \quad \approx \quad < x_i^{n-1} \sum_m \{ G_{im} + E_{iim} x_i \} \eta_m > \quad .$$

Then, dropping the subscript *i* for clarity, (26) may be approximated for each component x_i as

$$0 \quad \simeq \quad A < x^n > \quad + \quad < x^{n-1} \sum_m \left(G_m + E_m x \right) \eta_m > \quad - \quad < x^{n-1} > \frac{1}{2} \sum_m G_m E_m$$

for large *n*. And finally, without loss of generality, one may combine all the uncorrelated additive noise terms into a single term $b\xi_1$ and all the CAM noise terms into another term $(Ex+g)\xi_2$, where ξ_1 and ξ_2 are independent Gaussian white noises, as follows:

$$\begin{aligned} 0 \quad &\simeq A < x^n > \quad + \quad < x^{n-1} \sum_m \left(G_m + E_m x \right) \eta_m > \quad - \quad < x^{n-1} > \frac{1}{2} \sum_m G_m E_m \\ &= A < x^n > \quad + \quad < x^{n-1} \sum_{m'} G_{m'} \eta_{m'} > \quad + \quad < x^{n-1} \sum_{m''} \left(G_{m''} + E_{m''} x \right) \eta_{m''} > - \quad < x^{n-1} > \frac{1}{2} \sum_{m''} G_{m''} E_{m''} \\ &= A < x^n > \quad + \quad < x^{n-1} b \xi_1 > \quad + \quad < x^{n-1} \left(g + Ex \right) \xi_2 > - \quad < x^{n-1} > \frac{1}{2} Eg \end{aligned}$$

but this is *identical* to the equation for $< x^n >$ in the 1-d system (5). The increasing importance of self-correlation terms in the higher-order moment equations is thus the basic reason for the relevance of the 1-d model (5) in the dynamics of the higher-order moments even in multivariate systems.

The fact that the higher-order moment equations are more diagonally dominant also helps one understand why the *K-S* parabolas in Figures 2 and 5 are shifted slightly downward

relative to the prediction (20) of the 1-d theory. Let us say that for each component x_i of a multivariate system, the dynamical equation is (5) with an additional error term

$$\frac{dx}{dt} = Ax + b\,\eta_1 + (Ex+g)\,\eta_2 - \frac{1}{2}Eg + error\,, \tag{27}$$

where we have again dropped the subscript i for convenience. The moment equations are then $<x> = 0$ and

$$< x^2 > = \sigma^2 = -\frac{g^2+b^2}{2M+E^2} + \varepsilon^{(2)}\,\sigma^2 \quad \text{for the first two moments, and}$$

$$< x^n > = -\frac{\left(\frac{n-1}{2}\right)}{\left[M+\left(\frac{n-1}{2}\right)E^2\right]}\left[2Eg < x^{n-1} > + (g^2+b^2) < x^{n-2} >\right] + \varepsilon^{(n)}\,\sigma^n \tag{28}$$

for $n>2$. The quantities $\varepsilon^{(n)}$ represent the error made in $< x^n > / \sigma^n$ by ignoring the non-local dynamics. From this the local K-S relationship between K and S may be derived as

$$\begin{aligned} K = {} & \frac{3}{2}\left[\frac{1+\alpha}{1+(3/2)\alpha}\right]S^2 + 3\left[\frac{1+(1/2)\alpha}{1+(3/2)\alpha} - 1\right] \\ & - \frac{3}{2}\left[\frac{1+\alpha}{1+(3/2)\alpha}\right]S\,\varepsilon^{(3)} - 3\left[\frac{1+(1/2)\alpha}{1+(3/2)\alpha}\right]\varepsilon^{(2)} + \varepsilon^{(4)} \end{aligned} \tag{29}$$

in which, as in (20), the ratios involving α are all greater than unity. We may therefore rewrite this as

$$K > \frac{3}{2}\,S^2 + r\,, \tag{30}$$

where

$$\begin{aligned} r = {} & 3\left[\frac{1+(1/2)\alpha}{1+(3/2)\alpha} - 1\right] - 3\left[\frac{1+(1/2)\alpha}{1+(3/2)\alpha}\right]\varepsilon^{(2)} \\ & - \frac{3}{2}\left[\frac{1+\alpha}{1+(3/2)\alpha}\right]S\,\varepsilon^{(3)} + \varepsilon^{(4)} \end{aligned} \tag{31}$$

Since diagonal dominance is stronger for the higher moments, we expect that $|\varepsilon^{(4)}| < |\varepsilon^{(3)}| < |\varepsilon^{(2)}|$, and therefore the first two terms on the right hand side to dominate in (31). As mentioned previously, the first term is always positive. The second term is negative if $\varepsilon^{(2)}$ is positive. One can provide a theoretical argument as well as empirical evidence that $\varepsilon^{(2)}$ is positive. The former relies on the fact the linear operator A in (1) is in almost all geophysical contexts a "non-normal" operator that does not commute with its transpose. This non-normality of A (which refers to the non-orthogonality of the eigenfunctions of A and should not be confused with non-Gaussianity) leads to a greater variance of x than for a "normal" A with the same eigenvalues (Ioannou 1995). In most cases, this non-normality is associated with the ability of anomalies to draw energy from a background state, of which there is pervasive evidence and which is indeed one of the cornerstones of dynamical meteorology and oceanography. Most recently, Newman and Sardeshmukh (2008) have confirmed that the variance budget of observed extratropical weekly circulation anomalies is dominated by a local balance between stochastic forcing and local damping, and that non-local dynamical effects increase the anomaly variance at all locations. Such enhancements of variance cause the K-S parabola to have a slight negative bias, as evident in Figs 2 and 5. Note, however, that the $1.5\ S^2$ dependence of K on S is still preserved, consistent with our simple analysis here.

7. Discussion and Concluding Remarks

In this paper we demonstrated that certain types of non-Gaussian statistics are consistent with linear stochastically forced (LSF) dynamics with correlated additive and

multiplicative (CAM) noise forcing. In particular, we emphasized that skewed PDFs can be reconciled with such LSF models. We also showed that some remarkable relationships found between the third, fourth, and fifth moments, and also power-law tails, in both observations and in a long dry adiabatic GCM simulation are consistent with the simplest 1-d LSF model with CAM noise. We attributed the 1-d model's success to a principle of increasing "diagonal dominance" in the higher-order moment equations of multivariate systems, associated with the increasing importance of the self-correlation terms in those equations.

It should be emphasized that not all types of non-Gaussian behavior observed in the climate system may be reconcilable with LSF dynamics with CAM noise. The 1-d model predicts, for instance, a unique PDF maximum and therefore cannot account for the multiple PDF maxima sometimes claimed to exist in observations and climate model simulations (e.g. Hansen and Sutera 1986, Corti et al 1999, Monahan et al 2000, 2001). It is likely that multi-dimensional LSF models also cannot account for multiple PDF maxima, although we did not actually show this. A clear demonstration of more than one PDF maximum has been hindered in previous studies by the sampling uncertainties associated with limited observational records and relatively short climate model integrations, and also by methodological limitations (e.g. Stephenson et al 2004, Christiansen 2005). Very long integrations with state-of-the-art coupled climate models could resolve the issue, but the fact that one has to work so hard to show this may be a sign that any multi-modality as may be exist is weak and arguably not of great practical consequence. It is also noteworthy that Berner and Branstator (2007) found only

unimodal PDFs in the longest integration performed to date, of 14 million days, with an (admittedly low-resolution) atmospheric GCM.

It should also be recognized that the LSF approximation of coarse-grained anomaly dynamics is ultimately only an approximation, and like all approximations is not equally accurate in all situations. In climatic contexts, its applicability is mostly limited to departures from the annual cycle, and does not extend to the annual cycle itself (Huang and Sardeshmukh 2002). Some role for deterministic nonlinear dynamics has also been argued, for instance, in tropical SST variability (e.g. Penland and Sardeshmukh 1995, Monahan and Dai 2004, An and Jin 2004), extratropical atmospheric circulation variability (e.g. Kravtsov et al 2005, Kondrashov et al 2006, Newman and Sardeshmukh 2008), and sea surface wind variability (e.g., Monahan 2004). Nonetheless, the LSF approximation is a powerful approximation for diagnostic and prediction purposes, whose utility has been demonstrated in numerous studies, and also in this paper.

Perhaps the most important result from our analysis is that LSF models with CAM noise can explain skewed statistics. They also make falsifiable predictions, as in (20) and (21), of the specific manner in which the kurtosis and fifth moments are related to skew. From the evidence presented here, these predictions appear to be borne out both in observations and in a dry adiabatic GCM simulation. We are currently unaware of any simple nonlinear model with the same ability to explain such relationships among higher-order moments. The existence of power-law PDF tails is another prediction of the simple linear 1-d model (5), that also appears to be borne out in reality and in our GCM simulation. While it is true that several types of nonlinear models can also account for power-law

tails (see Newman 2005 for a review), our 1-d model can additionally account for the differing magnitudes of the positive and negative tails as in lower right panel of Figure 7. It is not clear to what extent nonlinear models can do this.

Finally, our analysis raises the exciting possibility of using the SGS distribution (22) (or its extended forms (A1) and (A2)) to estimate and predict the probabilities of extreme anomalies. Given the relevance of Diagonal Dominance, we believe that this would provide a simple, dynamically justifiable, and arguably more accurate way to estimate the tails of anomaly PDFs than direct estimations from short observational records or GCM integrations. The key point is not only that one can approximate the distribution of many climate variables as an SGS distribution, but also that its parameters can be accurately estimated from relatively short observational records or GCM integrations by fitting (5) to the observed, simulated, or predicted time series of those variables. This is a topic of current research.

Acknowledgments

Discussions with our colleagues at the Climate Diagnostics Center, especially C. Penland and G.P. Compo, are gratefully acknowledged. This research was partly supported by funding from NOAA's Climate Program Office.

Appendix

Stationary probability density p(x) of the stochastically perturbed process (24)

The solution *p(x)* of (25) is different if E^2G^2 is larger or smaller than β^2.

If $E^2G^2 > \beta^2$, let $\gamma^2 = E^2G^2 - \beta^2$. The solution of (25) may then be written as

$$p(x) \;=\; \frac{1}{\mathcal{N}}\left[E^2x^2 + 2\beta x + G^2\right]^{\frac{1}{\alpha}-1} \exp\left[\;\frac{2}{\gamma}\left(f_{ext} - \frac{\beta}{\alpha}\right)\;\arctan\left(\frac{E^2x+\beta}{\gamma}\right)\right] \qquad \text{(A1)}$$

If $E^2G^2 < \beta^2$, let $\gamma^2 = \beta^2 - E^2G^2$. Then the solution of (25) is

$$p(x) \;=\; \frac{1}{\mathcal{N}}\left[E^2x^2 + 2\beta x + G^2\right]^{\frac{1}{\alpha}-1} \left|\frac{E^2x+\beta-\gamma}{E^2x+\beta+\gamma}\right|^{\frac{1}{\gamma}\left(f_{ext}-\frac{\beta}{\alpha}\right)} \qquad \text{(A2)}$$

Note that (A1) reduces to (22) for $f_{ext} = 0$ and 2-d noise with $(E_1, g_1, c_1) = (0, b, 0)$ and $(E_2, g_2, c_2) = (E, g, 0)$ in (24). Also, (A2) reduces to the Gamma probability density function $p \sim x^{1/2-1}\exp[-x/(b/-M)]$ for $f_{ext} = \beta/2$ and $G \equiv 0$ and $E \equiv 0$ in (24).

References

An, S.I., and F. F. Jin, 2004: Nonlinearity and asymmetry of ENSO. *J. Climate*, **17**, 2399-2412

Barsugli, J.J., and P.D. Sardeshmukh, 2002: Global atmospheric sensitivity to tropical SST anomalies throughout the Indo-Pacific basin. *J. Climate*, **15**, 3427-3442.

Barsugli, J.J., Shin, S.I., and P.D. Sardeshmukh, 2006: Sensitivity of global warming to the pattern of tropical ocean warming. *Climate Dynamics*, **27**, 483-492.

Berner, J., and G. Branstator, 2007: Linear and Nonlinear signatures in the planetary wave dynamics of an AGCM probability density functions. *J. Atmos. Sci*., 64, 117-136.

Borges, M. D., and P. D. Sardeshmukh, 1995: Barotropic Rossby wave dynamics of zonally varying upper level flows during northern winter. *J. Atmos. Sci*., **52**, 3779-3796.

Branstator, G. W., 2002: Circumglobal teleconnections, the jet stream waveguide, and the North Atlantic oscillation. *J. Climate*, **15**, 1893-1910.

Christiansen, B., 2005: The shortcomings of nonlinear principal component analysis in identifying circulation regimes. *J. Climate*, **18** , 4814-4823.

Corti, S., F. Molteni, and T. N. Palmer, 1999: Signature of recent climate change in frequencies of natural atmospheric circulation regimes. *Nature,* **29**, 799-802.

DelSole, T., 2004: Stochastic models of quasigeostrophic turbulence. *Surv. Geophys.*, **25**, 107–149.

Farrell, B. F., and P. J. Ioannou, 1995: Stochastic dynamics of the midlatitude atmospheric jet. *J. Atmos. Sci.,* **52**, 1642–1656

Franzke, C., Majda, A.J., and E. Vanden-Eijnden, 2005: Low-order stochastic mode reduction for a realistic barotropic model climate. *J. Atmos. Sci*., **62**, 1722-1745.

Gardiner, C. W., 2004: *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences,* Third Edition. Springer-Verlag, 415 pp.

Hall, N.M.J., and P. D. Sardeshmukh, 1998: Is the time mean Northern Hemisphere flow baroclinically unstable? *J. Atmos. Sci.* 55, 41-56.

Hannachi, A., Stephenson, D.B., and K. R. Sperber, 2003: Probability-based methods for quantifying nonlinearity in the ENSO. *Climate Dynamics*, **20**, 241-256.

Hansen, A.R., and A. Sutera 1986: On the probability density distribution of planetary-scale atmospheric wave amplitude. *J. Atmos. Sci*., **43**, 3250-3265.

Hasselmann, K., 1976: Stochastic climate models. Part I. Theory. *Tellus*, **28**, 474--485.

Hoskins, B.J., and T. Ambrizzi, 1993: Rossby-wave propagation on a realistic longitudinally varying flow. *J. Atmos. Sci*., **50**, 1661-1671.

Huang, H.P., and P.D. Sardeshmukh, 2000: Another look at the annual and semiannual cycles of atmospheric angular momentum. *J. Climate*, **13**, 3221-3238.

Ioannou, P.J., 1995: Nonnormality increases variance. *J. Atmos. Sci*., **52**, 1155-1158.

van Kampen, N., 1981: *Stochastic Processes in Physics and Chemistry*. Amsterdam.

Khas'minskii, R. Z., 1966: A limit theorem for the solutions of differential-equations with random right hand sides. *Theory Probab. Appl.,* **11**, 390–406.

Kimoto, M., and M. Ghil, 1993: Multiple flow regimes in the Northern Hemisphere winter. Part I: Methodology and hemispheric regimes. *J. Atmos. Sci.,* **50**, 2625-2643.

Knutson, T.R., Delworth, T.L., Dixon, K.W., Held, I.M., Lu, J., Ramaswamy, V., Schwarzkopf, M.D., Stenchikov, G., and R.J. Stouffer, 2006 : Assessment of Twentieth-century regional surface temperature trends using the GFDL CM2 Coupled Models. *J. Climate* ,**19**, 1624-1651

Kondrashov, D., Kravtsov, S., and M. Ghil, 2006: Empirical mode reduction in a model of extratropical low-frequency variability. *J. Atmos. Sci*., **63**, 1859-1877.

Kravtsov, S., Kondrashov, D., and M. Ghil, 2005: Multilevel regression modeling of nonlinear processes: Derivation and applications to climatic variability. *J. Climate*, **18**, 4404-4424.

Majda, A.J., Timofeyev, I., and E. Vanden-Eijnden, 2003: Systematic strategies for stochastic mode reduction in climate. *J. Atmos. Sci*., 60, 1705-1722.

Monahan, A.H., 2004: A simple model for the skewness of global sea surface winds. *J. Atmos. Sci*., **61**, 2037-2049.

Monahan, A.H., Fyfe, J.C,, and G.M. Flato, 2000: A regime view of northern hemisphere atmospheric variability and change under global warming. *GRL*., **27**, 1139-1142.

Monahan, A.H., Pandolfo, L., and J.C. Fyfe, 2001: The preferred structure of variability of the Northern Hemisphere atmospheric circulation. *GRL*., **28**, 1019-1022.

Monahan, A.H., and A.G. Dai, 2004: Spatial and temporal structure of ENSO nonlinearity. *J. Climate*, 17, 3026-3036 .

Müller, D., 1987: Bispectra of sea-surface temperature anomalies. *J. Phys. Oceanogr.*, **17**, 26-36.

Newman, M., Sardeshmukh, P.D., Winkler, C.R., and J.S. Whitaker, 2003: A study of subseasonal predictability. *Mon. Wea. Rev*., **131**, 1715-1732.

Newman, M., and P.D. Sardeshmukh, 2008: Tropical and stratospheric influences on extratropical short-term climate variability. *J. Climate*. doi: 10.1175/2008JCLI2118.1.

Newman, M.E.J.,2005: Power laws, Pareto distributions and Zipf's law. *Contemporary Phys*ics **46**, 323-351.

Papanicolaou, G. and W. Kohler, 1974: Asymptotic theory of mixing stochastic ordinary differential equations. *Commun. Pure Appl. Math*., **27**, 641-668.

Penland, C., 1989: Random forcing and forecasting using principal oscillation pattern-analysis. *Mon. Wea. Rev*., **117**, 2165-2185.

Penland, C., 1996: A stochastic model of Indo-Pacific sea surface temperature anomalies. *Physica D*, 98, 534-558.

Penland, C., and M. Ghil, 1993: Forecasting northern-hemisphere 700-mb geopotential height anomalies using empirical normal-modes. *Mon. Wea. Rev*., **121**, 2355-2372.

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*.

Penland, C., and P. D. Sardeshmukh, 1995: The optimal growth of tropical sea surface temperature anomalies. *J. Climate.*, **8***,* 1999-2024.

Saha, S., Nadiga, S., Thiaw, C., Wang, J., Wang,W., Zhang, Q., van den Dool, H.M., Pan, H. -L., Moorthi, S., Behringer, D., Stokes, Pena, D., Lord, S., White, G., Ebisuzaki, W., Peng, P., and P. Xie, 2006: The NCEP Climate Forecast System. *J. Climate*, 19, 3483-3517.

Sardeshmukh, P.D., and P. Sura, 2007: Multiscale impacts of variable heating in climate. *J. Climate*, **20**, 5677-5695.

Schneider, E.K., Bengtsson, L., and Z. -Z. Hu, 2003 : Forcing of northern hemisphere climate trends. *J. Atmos. Sci*. 60, 1504-1521.

Stephenson, D.B., Hannachi, A., and A. O'Neill, 2004: On the existence of multiple climate regimes. *Quart. J. Royal Met. Soc*., **130**, 583-605.

Sura, P., M. Newman, C. Penland, and P. D. Sardeshmukh, 2005: Multiplicative noise and

non-Gaussianity: A paradigm for atmospheric regimes? *J. Atmos. Sci.*, **62**, 1391–1409.

Sura, P., and P. D. Sardeshmukh, 2008: A global view of non-Gaussian SST variability. *J. Phys. Oceanogr.*, in press.

Whitaker, J.S., and P. D. Sardeshmukh, 1998 A linear theory of extratropical synoptic eddy statistics. *J. Atmos. Sci,* **55***, 237-258.*

White, G.H., 1980 : Skewness, kurtosis and extreme values of northern hemisphere geopotential heights. *Mon. Wea. Rev*., 108, 1446-1455.

Winkler, C. R., M. Newman, and P. D. Sardeshmukh, 2001: A linear model of wintertime low-frequency variability. Part I: Formulation and forecast skill. *J. Climate*, **14**, 4474-4494.

Figure Captions

Fig. 1. Observed Skewness S and excess Kurtosis K of daily 300 mb vorticity variations during the northern winters of 1970-99, estimated using the NCEP/NCAR reanalysis dataset. The fields are both colored and contoured for clarity. The contours are drawn at intervals of 0.4, starting at 0.2.

Fig. 2. The S and K values from Figure 1 displayed in the form of a scatter plot. The solid curve is a parabola $K = 1.5\ S^2 - 0.6$. The estimated local 95% confidence intervals are indicated in the upper right corner of the figure.

Fig. 3. *Left panel*: Departure from Gaussianity of the joint PDF of the principal component time series associated with the two dominant EOFs of weekly-averaged 750 mb streamfunction anomalies during the northern winters (DJF) of 1950-2002. *Right panel* : Departure from Gaussianity of the joint PDF of the time series of the real and imaginary parts of the dominant barotropic perturbation eigenmode of the northern hemispheric circulation obtained when it is steadily forced and stochastically damped. Adapted from Sura et al (2005). See text for more explanation.

Fig. 4. As in Figure 1, but obtained from a long 1200-winter simulation of a dry adiabatic GCM with prescribed constant forcing, as described in the text.

Fig. 5. As in Figure 2 but from the 1200-winter GCM simulation.

Fig. 6. Scatter plots of the fifth moments versus skewness S of the daily 300 mb vorticity (left) and 500 mb geopotential height (right) variations in the 1200-winter GCM simulation. The straight lines and curves facilitate comparison with the prediction of the linear 1-d theory (Eq. 21). The straight lines are $10S$; the curves are $10S + S^3$.

Fig. 7. Observed (left) and GCM simulated (right) PDFs of standardized daily wintertime 500 mb geopotential height (upper panels) and 300 mb vorticity (lower panels) anomalies at the locations of largest skew in the north Pacific. The results are shown on a log-log scale, with the probabilities of the negative anomalies (circles) flipped over to the positive side for better comparison with those of positive anomalies (triangles). The curve in all panels is a reference Gaussian. Results are not shown for standardized anomaly magnitudes of less than unity. The straight lines are simple linear fits to the PDF tails.

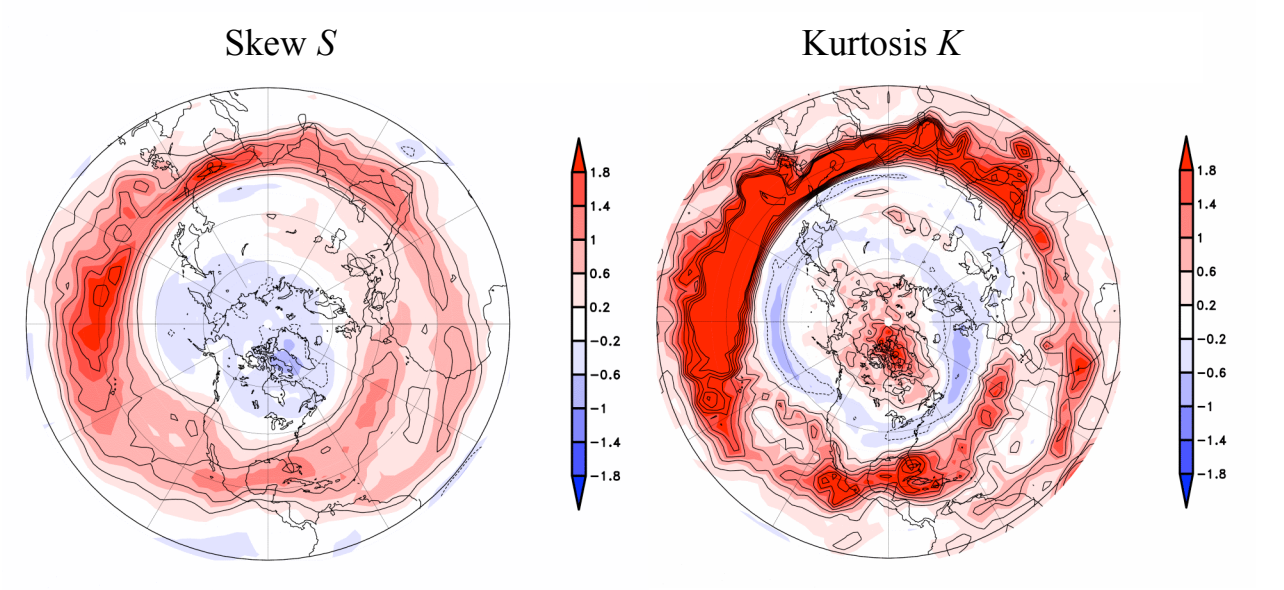

Fig. 1. Observed Skewness *S* and excess Kurtosis *K* of daily 300 mb vorticity variations during the northern winters of 1970-99, estimated using the NCEP/NCAR reanalysis dataset. The fields are both colored and contoured for clarity. The contours are drawn at intervals of 0.4, starting at 0.2.

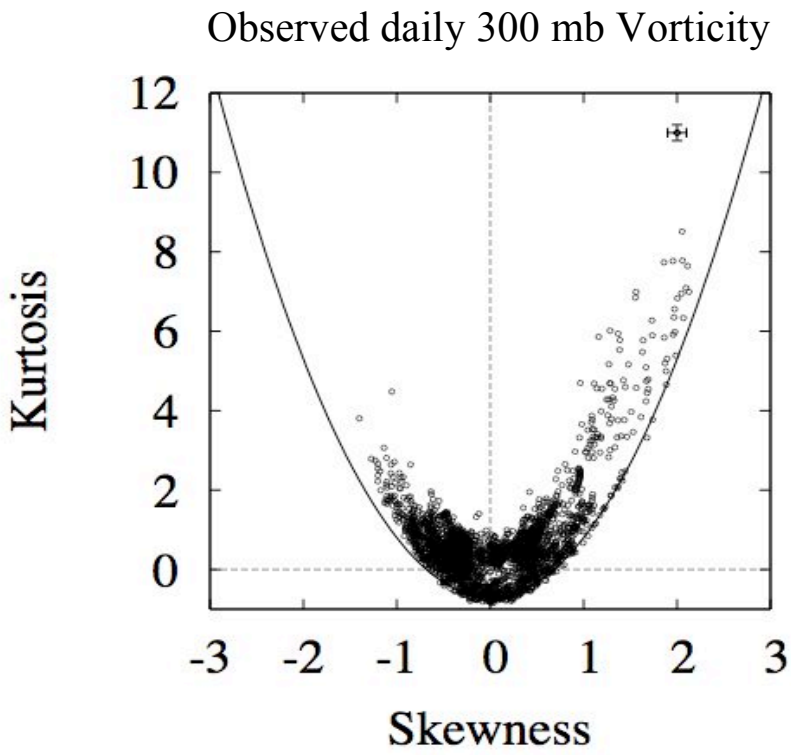

Fig. 2. The S and K values from Figure 1 displayed in the form of a scatter plot. The solid curve is a parabola $K = 1.5\ S^2 - 0.6$. The estimated local 95% confidence intervals are indicated in the upper right corner of the figure.

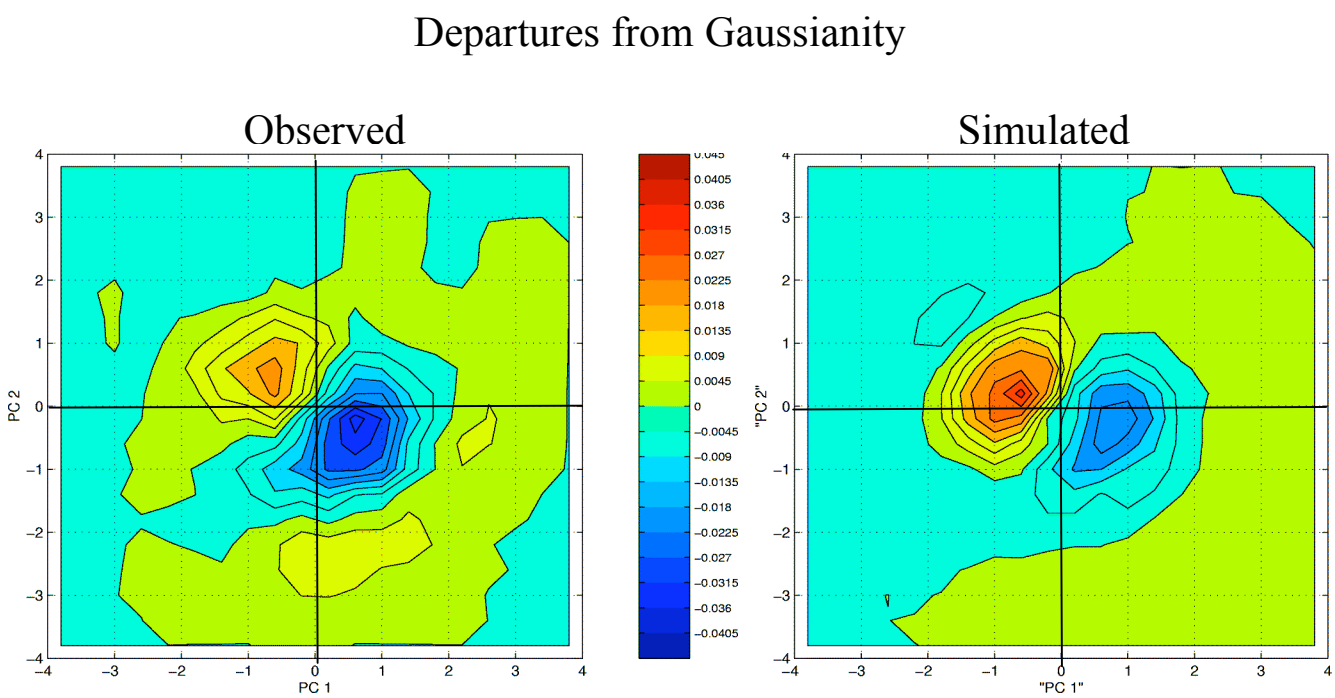

Fig 3. *Left panel*: Departure from Gaussianity of the joint PDF of the principal component time series associated with the two dominant EOFs of weekly-averaged 750 mb streamfunction anomalies during the northern winters (DJF) of 1950-2002.

Right panel : Departure from Gaussianity of the joint PDF of the time series of the real and imaginary parts of the dominant barotropic perturbation eigenmode of the northern hemispheric circulation obtained when it is steadily forced and stochastically damped. Adapted from Sura et al (2005). See text for more explanation.

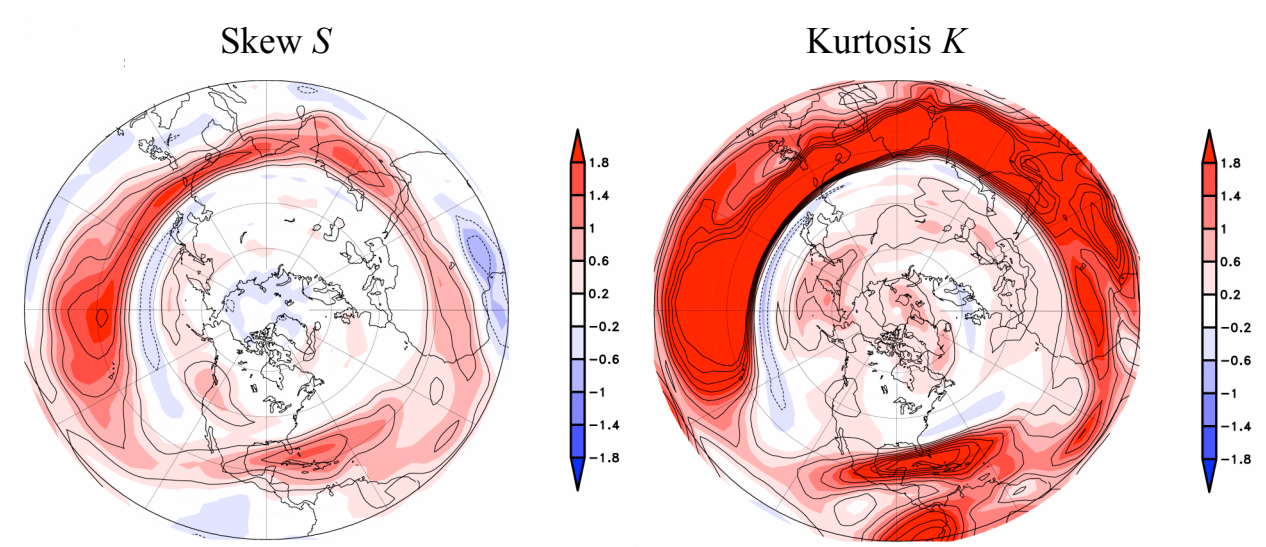

Fig. 4. As in Figure 1, but obtained from a long 1200-winter simulation of a dry adiabatic GCM with prescribed constant forcing, as described in the text.

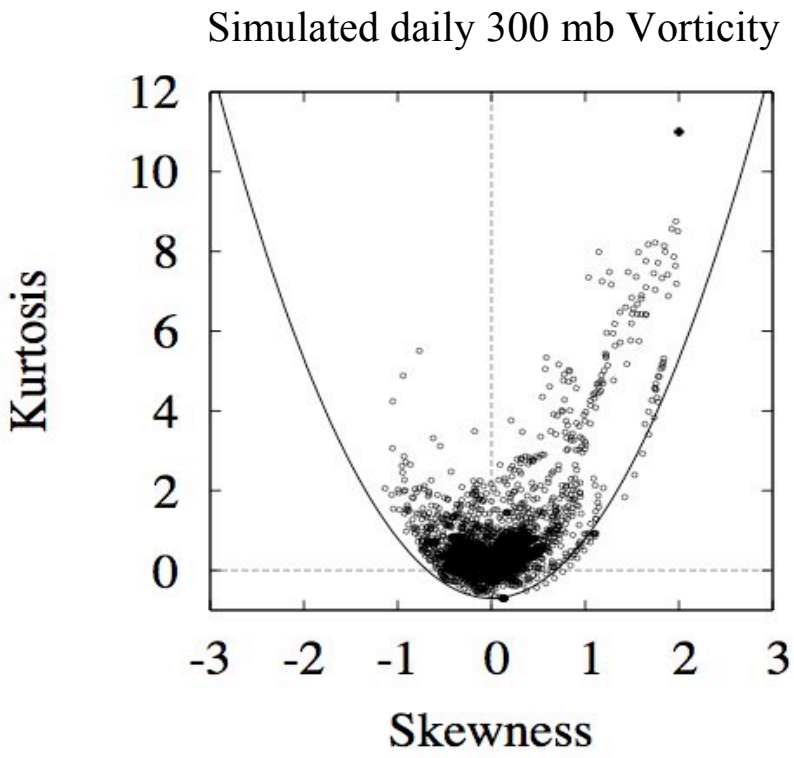

Figure 5. As in Figure 2 but from the 1200-winter GCM simulation.

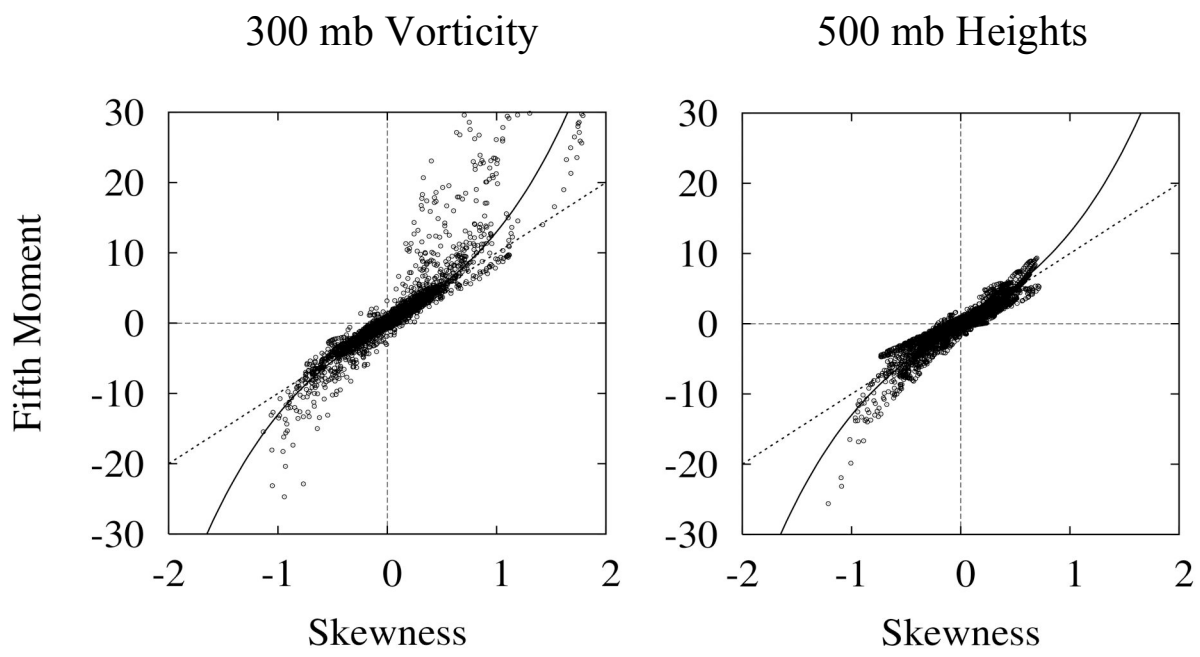

Fig. 6. Scatter plots of the fifth moments versus skewness S of the daily 300 mb vorticity (left) and 500 mb geopotential height (right) variations in the 1200-winter GCM simulation. The straight lines and curves facilitate comparison with the prediction of the linear 1-d theory (Eq. 21). The straight lines are $10S$; the curves are $10S + S^3$.

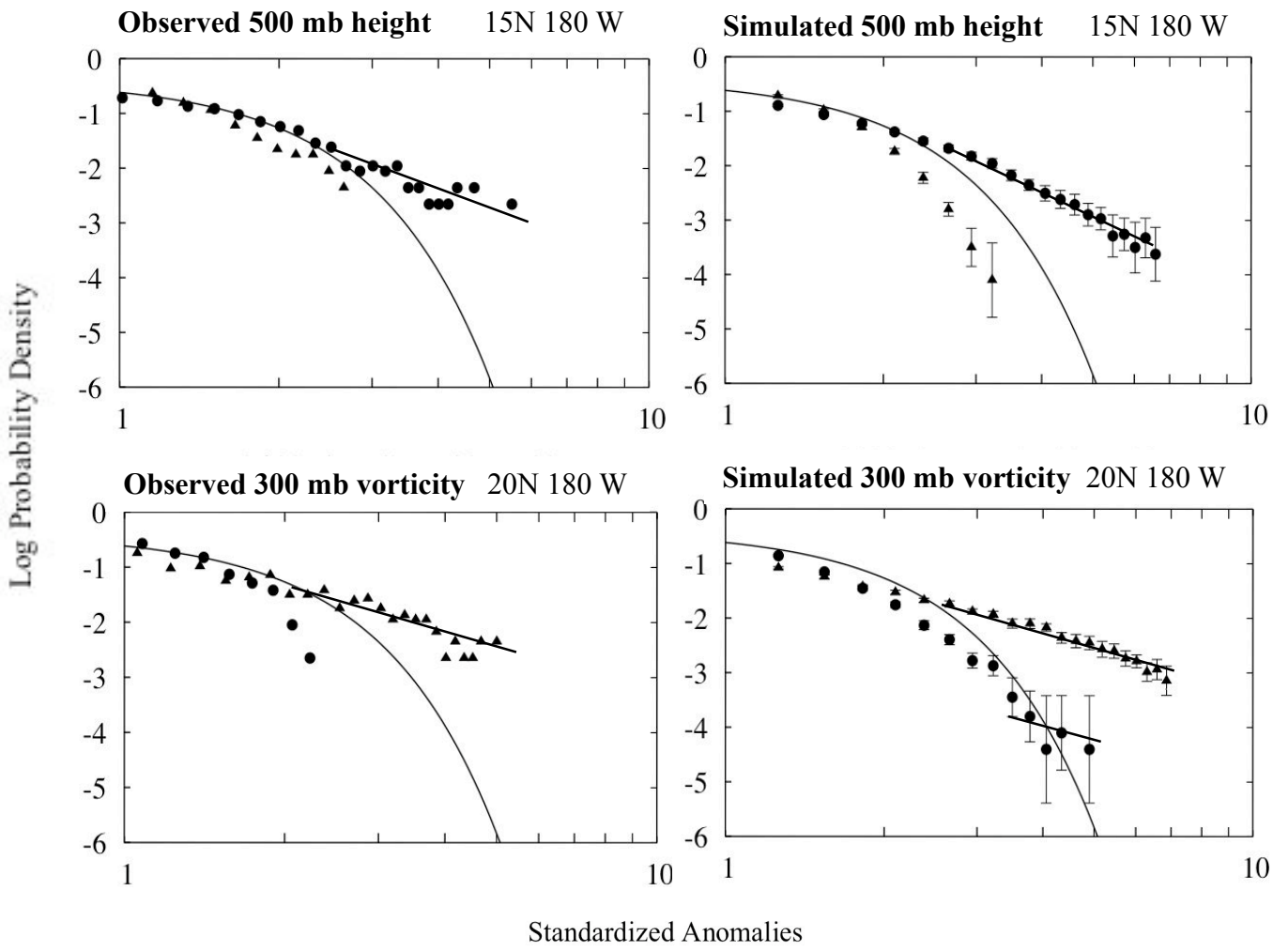

Fig. 7. Observed (left) and GCM simulated (right) PDFs of standardized daily wintertime 500 mb geopotential height (upper panels) and 300 mb vorticity (lower panels) anomalies at the locations of largest skew in the north Pacific. The results are shown on a log-log scale, with the probabilities of the negative anomalies (circles) flipped over to the positive side for better comparison with those of positive anomalies (triangles). The curve in all panels is a reference Gaussian. Results are not shown for standardized anomaly magnitudes of less than unity. The straight lines are simple linear fits to the PDF tails.