Climatology of Non-Gaussian Atmospheric Statistics

MAXIME PERRON AND PHILIP SURA
Department of Earth, Ocean, and Atmospheric Science, The Florida State University, Tallahassee, Florida

(Manuscript received 8 September 2011, in final form 6 June 2012)

ABSTRACT

A common assumption in the earth sciences is the Gaussianity of data over time. However, several independent studies in the past few decades have shown this assumption to be mostly false. To be able to study non-Gaussian climate statistics, one must first compile a systematic climatology of the higher statistical moments (skewness and kurtosis; the third and fourth central statistical moments, respectively). Sixty-two years of daily data from the NCEP–NCAR Reanalysis I project are analyzed. The skewness and kurtosis of the data are found at each spatial grid point for the entire time domain. Nine atmospheric variables were chosen for their physical and dynamical relevance in the climate system: geopotential height, relative vorticity, quasigeostrophic potential vorticity, zonal wind, meridional wind, horizontal wind speed, vertical velocity in pressure coordinates, air temperature, and specific humidity. For each variable, plots of significant global skewness and kurtosis are shown for December–February and June–August at a specified pressure level. Additionally, the statistical moments are then zonally averaged to show the vertical dependence of the non-Gaussian statistics. This is a more comprehensive look at non-Gaussian atmospheric statistics than has been taken in previous studies on this topic.

1. Introduction

In meteorology, it is often assumed that the probability distribution function (PDF) of many atmospheric variables is Gaussian, or normal, except for midlatitude regions exhibiting potential regime behavior. In a Gaussian PDF, the entire distribution can be described by knowing the mean and variance of the data. Approximately 68%, 95%, and 99% of the data can be found between ±1, 2, and 3 standard deviations, respectively. Indeed, the central limit theorem states that as the sum of independent random variables approaches infinity, the resulting PDF will look more and more Gaussian. Predicting the state of the atmosphere by using a linear stochastic model that includes additive white noise also results in Gaussian distributions (Sura et al. 2005).

However, it has been known for some time that atmospheric variables do not feature Gaussian PDFs as a general rule. It appears that the atmosphere, being a complex nonlinear system with several innate asymmetries (rotation of the earth, uneven insolation, changing water vapor partial pressure with temperature, etc.), does not meet the assumptions of the central limit theorem at all times or locations. Stochastic models can be adapted to give non-Gaussian distributions in three different ways: the first is to make the deterministic term nonlinear, the second is to make the noise term state-dependent, and the third is to make the noise non-Gaussian (Sura et al. 2005). We know that a change in the mean state of the climate can shift the PDF of its different variables (see Houghton 2009; Brönnimann et al. 2008; Easterling et al. 2000), but it is much tougher to determine how the actual shape of the PDF of each variable is changing.

The concept of multiple regimes or bimodality (i.e., non-Gaussianity) in the weather system has been around since modern meteorology began, although it took some time before the first dynamical frameworks were developed. Charney and DeVore (1979), Wiin-Nielsen (1979), and Hart (1979) showed that multiple equilibrium states (blocked and zonal regimes) caused by wave–mean flow interactions could exist in the atmosphere. This led to an effort to find regimes in observed midlatitude flows (e.g., Hansen and Sutera 1986; Mo and Ghil 1988; Molteni et al. 1990; Kimoto and Ghil 1993; Cheng and Wallace 1993; Corti et al. 1999; Smyth et al. 1999; Monahan et al. 2001; Christiansen 2005, and many
Concurrently, several empirical studies on the topic of local non-Gaussianity and higher statistical moments (e.g., White 1980; Trenberth and Mo 1985; Nakamura and Wallace 1991; Holzer 1996; Monahan 2006b; Petoukhov et al. 2008) were published. These papers, which showed significant deviations from Gaussianity for key atmospheric variables, will be compared with this study in future sections. Only recently have we seen theoretical reasons for the non-Gaussian fields shown in these papers. The theory is discussed in papers such as Holzer (1996), Sura and Sardeshmukh (2008), Sardeshmukh and Sura (2009), Sura and Perron (2010), and others. Related to this, other studies in this field have looked at the non-Gaussianity of the first few components of empirical orthogonal functions (EOFs), such as Mo and Ghil (1988), Molteni et al. (1990), Kimoto and Ghil (1993), Corti et al. (1999), Smyth et al. (1999), Monahan et al. (2001), Majda et al. (2003), Berner (2005), Berner and Branstator (2007), Franzke et al. (2005), and Majda et al. (2008).

To determine whether a PDF is Gaussian, we can calculate the higher statistical central moments, skewness and kurtosis. Skewness is a measure of the asymmetry of the PDF. A PDF with a positive (negative) skewness has heavier tails on the positive (negative) anomaly side of the mean. Kurtosis is a measure of how much of the data is found in the tails. Alternatively, one can describe kurtosis as the level of peakedness of the PDF. A PDF with a positive excess kurtosis has a very sharp peak near the mean, which inevitably means that it also has more data in the tails relative to a Gaussian PDF. On the other hand, a PDF with a negative excess kurtosis will have a much flatter profile, with fewer data in the tails and near the mean relative to a Gaussian PDF. These two moments will be explained in a more quantitative way in the next section.

Objectives

To date, no paper has summarized the non-Gaussianity of the atmosphere in a comprehensive and elegant manner. Early studies (White 1980; Trenberth and Mo 1985; Nakamura and Wallace 1991; Holzer 1996) focused only on the geopotential height. White (1980) only looked at the Northern Hemisphere and only had 11 years of midtropospheric data to work with. Trenberth and Mo (1985) did a Southern Hemisphere analysis and had even fewer data. Nakamura and Wallace (1991) filtered the data to remove low-frequency variability, while Holzer (1996) created zonally averaged cross sections of geopotential height. Monahan (2006b) looked at horizontal wind speed for different seasons and for four different datasets. Petoukhov et al. (2008) looked at other variables: temperature, vertical velocity in pressure coordinates, specific humidity, horizontal wind components, and synoptic-scale eddy kinetic energy. Some recent studies (e.g., Sura and Perron 2010) have also looked at variables such as relative vorticity and potential vorticity. However, there is no study that presents the non-Gaussianity of key variables (from one dataset) together in one place.

Thus, the objective of this paper is to present the non-Gaussianity of several key atmospheric variables in a concise and consistent way. The data used cover a 62-yr period between 1948 and 2009. Both horizontal plots and zonally averaged vertical cross sections are shown for each of the nine variables picked. Each of those is split up between the summer and winter season to show the seasonality of the higher statistical moments. Finally, a rigorous analysis is performed to show that our results are statistically significant.

2. Data and methodology

In this section, the rationale for picking a particular dataset and variables to be analyzed will be discussed. Additionally, some background information on non-Gaussianity will be shown.

a. Choice of data and variables

For this study, we needed a dataset of atmospheric variables that spanned several decades. The data chosen were from the National Centers for Environmental Prediction (NCEP)–National Center for Atmospheric Research (NCAR) Reanalysis I Project. Going back to 1948, the data are a cluster of observations from different sources that uses models to fill in the gaps (Kalnay et al. 1996). Daily-averaged data are used in this study. For each grid point, we looked at daily anomalies calculated by subtracting the 62-yr time mean for each day from the full field at each day. After the yearly cycle had been removed from the data, the daily anomalies were plugged into the equations for the higher statistical moments that will be shown later in this section.

The accuracy of reanalysis datasets depends partly on the accuracy and spatiotemporal coverage of observations worldwide. It has been known for some time that data from certain parts of the globe (e.g., the Southern Hemisphere before the 1970s; see Hines et al. 2000) exhibit slight biases. However, in Sura and Perron (2010) it is shown that the general large-scale patterns of higher statistical moments remain stable over the reanalysis period.

The variables examined in the study were chosen for their historical and/or physical importance in the earth–atmosphere system. These are summarized in Table 1.
Vorticity is a fundamental measure of rotation in flow around an axis. In meteorology we most often speak of the vertical vorticity, which is calculated from the two horizontal wind components $u$ and $v$. The relative vorticity $\zeta$ is often added to the Coriolis parameter $f$ to form the absolute vorticity $\eta$. In spherical coordinates, the relative vorticity is given by

$$\zeta = \frac{1}{a \cos \phi} \frac{\partial}{\partial \lambda} \left( \frac{\partial (u \cos \phi)}{\partial \phi} \right),$$

where $a$ is the distance from the center of the earth, $\phi$ is the latitude, and $\lambda$ is the longitude. This variable was recently examined by Sura and Perron (2010).

Additionally, we look at the potential vorticity, which is equal to the absolute vorticity normalized by the “depth” of a barotropic fluid. This variable is conserved following horizontal motion if we ignore forcing such as adiabatic effects and friction. Therefore, it has become valued as a tracer variable in atmospheric dynamics. In this paper, we use the quasigeostrophic potential vorticity (QGPV), which has the advantage of requiring only information about the geopotential height $\Phi$ and temperature $T$ fields:

$$QGPV = f + \frac{1}{f (a \cos \phi)^2} \frac{\partial^2 \Phi}{\partial \lambda^2} + \frac{1}{a^2 \cos \phi} \frac{\partial}{\partial \phi} \left( \frac{\cos \phi \partial \Phi}{f \partial \phi} \right)$$

$$+ f \left( \frac{\partial}{\partial p} \left( \frac{1}{\sigma_p} \frac{\partial \phi}{\partial p} \right) \right),$$

where

$$\sigma_p = -R \left( \frac{p}{P_s} \right)^{R/C_p} \frac{d \theta_{ref}}{dp}$$

is the static stability parameter, $f = 2 \Omega \sin \phi$ is the Coriolis parameter,

$$\theta_{ref} = T \left( \frac{p}{P_s} \right)^{R/C_p}$$

is the potential temperature averaged horizontally, $\Omega$ is the rotation rate of the earth, $R$ is the gas constant for dry air, $C_p$ is the specific heat capacity of dry air, $p$ is the pressure, and $p_s$ is the reference surface pressure (Evans and Black 2003). Previously, Sura and Perron (2010) presented and discussed the higher central moments for this variable. In this study, we will mainly look at the 300-hPa level for vorticity variables, where adiabatic and friction effects are minimized and where the jet stream is found.

3) HORIZONTAL WIND

On a synoptic and global scale, winds exist to create a balance between the main forces in the atmosphere. On the horizontal plane, the most important of these forces are the pressure gradient force and the Coriolis force. Much of the energy transfer in the atmosphere is a result of advection by the wind. Thus, we look at the zonal $u$ and meridional $v$ wind components separately as these variables are the ones included in the different flux terms of meteorological and oceanographic equations. Petoukhov et al. (2008) and Sura and Perron (2010) both examined the higher moments of the individual wind components.

However, when taking a zonal average of the moments, much of the information contained in the meridional wind values gets lost. An inexperienced interpreter could conclude that the global zonally averaged meridional wind distribution is roughly Gaussian. Thus, we also analyze the magnitude of the horizontal wind, given by $|\mathbf{u}| = \sqrt{u^2 + v^2}$. This quantity was simply derived from
the zonal and meridional wind components found in the original data. Monahan (2006b) and Petoukhov et al. (2008) both looked at this variable. For horizontal plots, the 925-hPa level was chosen for all wind variables because low-level wind speed has a large impact on the strength and direction of ocean currents.

4) **VERTICAL VELOCITY IN PRESSURE COORDINATES**

Unlike the horizontal wind components, which are easy to measure and well predicted by current theory, the vertical velocity is a much trickier variable. Vertical motion can happen as a result of sloped terrain, but the most important contributor remains lift caused by convection or frontal features. Thus, vertical motion is often associated with instability and, by definition, ageostrophy. As with geopotential height, we choose 500 hPa, the level of nondivergence, to be our main level of interest. This is because vertical velocity is maximized in nondivergent flow. Petoukhov et al. (2008) has skewness and kurtosis plots for this variable.

5) **TEMPERATURE**

The air temperature is a thermodynamic variable that has been measured for centuries. At the most basic level, it is the air temperature that drives the global atmospheric circulation. Air warmed up by the sun in the tropics rises far into the upper troposphere, forming one branch of the Hadley cell. The transport of heat through advection into the higher latitudes is an important part of cyclogenesis in the midlatitudes. However, we must not forget that about half of the energy transport to the poles is done by the oceans, which have a higher heat capacity but advect more slowly than the atmosphere. Here, we choose 925 hPa for our analyses because most of the solar radiation is used to warm the surface. Again, Petoukhov et al. (2008) examined the non-Gaussianity of temperature in more detail.

6) **SPECIFIC HUMIDITY**

Specific humidity is an interesting variable because very little research has been done on its non-Gaussianity. One issue with humidity is the fact that the ocean/land distribution as well as the temperature field plays a major role in its distribution in the lower troposphere. Again, we pick 925 hPa for this variable because of its proximity to the ocean surface. The NCEP–NCAR Reanalysis I Project does not record humidity data above 300 hPa (likely owing to the scarcity of moisture near the troposphere), so we cannot analyze the range between 300 and 100 hPa for this variable. Petoukhov et al. (2008) looked at this variable.

**b. Methodology: Non-Gaussian statistics**

1) **HIGHER STATISTICAL MOMENTS**

In statistics, distributions can be described quantitatively by the central moments, or moments about the mean. The central moments are calculated using the anomalies from a mean state of a time series. Therefore, the first central moment, essentially the mean of the anomalies, is defined to be always zero. The second central moment, variance, is a measure of the spread of the data on either side of the mean—typically, we use the square root of the variance, the standard deviation, instead. The two values we are interested in, however, are the third and fourth central moments: skewness \( \gamma_1 \) and kurtosis \( \gamma_2 \):

\[
\gamma_1 = \frac{E[(X - \mu)^3]}{E[(X - \mu)^2]^{3/2}} \quad \text{and} \quad \gamma_2 = \frac{E[(X - \mu)^4]}{E[(X - \mu)^2]^2},
\]

where \( E[\ ] \) denotes the expected value, \( X \) is the variable, and \( \mu \) is the mean of the variable, that is, \( E[X] \) (see, e.g., Lomax 2007).

Skewness is a measure of the asymmetry of the distribution. For a Gaussian distribution, this number is zero, since it is fully symmetric. A positive number denotes a heavier tail in the positive anomalies, and a negative number denotes a heavier tail in the negative anomalies. Kurtosis is a measure of the peakedness of the distribution. A Gaussian distribution has a kurtosis of three. Anything below 3 means that the distribution is flatter and anything above 3 means that the distribution is more peaked. As a corollary of this definition, a high kurtosis means that more data points are located in the tails and near the mean relative to a Gaussian distribution, while a low kurtosis means the opposite. For ease of analysis, we will plot the excess kurtosis instead of the kurtosis. The excess kurtosis is simply the kurtosis minus 3; this means that a Gaussian distribution has a skewness and an excess kurtosis of zero. Significant nonzero values for the third and fourth central moments would represent a non-Gaussian distribution.

In this study, we calculate the skewness and kurtosis of the nine variables of interest for the entire 62-yr period covered by the daily data. Since we are also interested in seeing the seasonal variations of the non-Gaussianity, we calculate the moments for the December–February (DJF) or June–August (JJA) subsets. Furthermore, we have also averaged these values zonally, so that we can show the data as vertical cross sections with latitude on the x axis. It is important to remind the reader that anomalies, not the field itself, are required in the calculation of central
statistical moments. In this study, we use anomalies from the 62-yr mean with the seasonal cycle removed.

2) Statistical significance

One way to calculate the statistical significance of our results is to use the standard error for skewness and kurtosis:

$$\varepsilon_{\gamma_i} = \sqrt{6/N_i} \quad \text{and} \quad (5)$$

$$\varepsilon_{\gamma_2} = \sqrt{24/N_i} = 2\varepsilon_{\gamma_1}, \quad (6)$$

where $N_i$ is the number of independent data points in our time series at each location (Brooks and Carruthers 1953). Note that (5) and (6) are derived assuming a Gaussian distribution and are, therefore, approximately valid for weakly non-Gaussian data. The exact standard errors of skewness and kurtosis depend on their underlying population.

Thus, one major obstacle in using these formulas is the fact that we do not know a priori the shape of the distribution, and therefore the assumptions going into the standard error formulas are not fully met. Ideally, one would use resampling or subsampling methods, which work under weaker assumptions than most other statistical analysis techniques, to find the errors of the statistical moments (Gluhovsky and Agee 2009; Gluhovsky 2011). However, given the large sample size of the data, it is more practical to use a slightly less computational intensive technique, assuming that this technique can adequately reflect the statistical parameters of the original data. Thus, we choose a Monte Carlo approach to estimate the confidence intervals of skewness and kurtosis from a Gaussian red-noise process fitted to the data.

Using this method, we get a 95% skewness confidence interval that is smaller than $\pm0.12$ for eight of the nine variables (QGPV being the only outlier) outside of the tropics. Thus, for consistency, we chose $\pm0.12$ as the nonshaded contour interval for each global plot. In most areas these errors are even smaller, so this value is an underestimate. For more details on the results of the Monte Carlo approach, see the appendix.

Having established the standard errors from a Monte Carlo approach, let us see what a back-of-the-envelope calculation using (5) and (6) yields for a midlatitude locale. Take 91 days in a season and 62 yr in the dataset for a total of 5642 days. Assume that it takes roughly one week for a midlatitude Rossby wave to travel a full wavelength over a particular area. Then we get $5642/7 = 806$ independent observations $N_i$. Plugging this value into the equation for skewness above gives an error of 0.086. Therefore, we would be 95% confident in the skewness being found within $\pm2$ standard errors ($\pm2 \times 0.086$) of the mean. That is, we established that (5) and (6) are good approximations even for non-Gaussian data. To reiterate, the nonshaded contour interval in the plots was not based on the standard error formula, but rather on the Monte Carlo method.

3. Climatology of non-Gaussian statistics

In this section, we will describe and analyze the non-Gaussian statistics of the nine variables mentioned earlier. A full-page figure composed of eight panels (described in detail below) represents each atmospheric variable for a total of 72 panels. Each figure is set up in the same manner; the top four panels (a, b, c, and d) show the horizontal variation of non-Gaussian statistics at a specific pressure level (the reasoning behind picking each level is explained in the previous section). In these panels, the x-axis is the longitude and the y-axis is the latitude. The bottom four panels (e, f, g, and h) show the vertical profile of the zonally averaged non-Gaussian statistics. In these panels, the x-axis is the latitude and the y-axis is the logarithm of the pressure. The left panels (a, c, e, and g) show the skewness of that particular variable while the right panels (b, d, f, and h) show the kurtosis. Finally, the first and third line of the panels (a, b, c, and f) feature the seasonal subset of the data taken during the boreal winter months (DJF), while the second and fourth line of panels (c, d, g, and h) feature the seasonal subset of the data taken during the boreal summer months (JJA).

a. Geopotential height

We start with an analysis of Fig. 1, which shows the higher statistical moments of the geopotential height. At the 500-hPa level, we notice that the significant non-Gaussian values are arranged in bands that travel (for the most part) all the way around the globe zonally. There exists much more variation in these moments as one travels north and south. The DJF skewness, for example, features negative values centered on 30° in both hemispheres. Positive skewness tends to be found in the tropics and in polar regions. However, during the JJA period, the positive skewness almost entirely disappears except for certain areas near the Antarctic continent. These patterns are consistent with the findings of White (1980), Trenberth and Mo (1985), Nakamura and Wallace (1991), and Holzer (1996). In these papers, the areas featuring non-Gaussianity were linked to climatological features of the synoptic atmosphere such as the mean jet location and troughs and ridges. Part of the non-Gaussianity is attributed to blocking in the atmosphere, in which the large-scale flow pattern becomes stuck in an untypical configuration for a few to several weeks. Holzer (1996)
also attributes some of the non-Gaussianity of the geopotential height field at midlatitudes to the so-called “rectification process.” Essentially, the advective nonlinearity of the system causes a rectification of the near-symmetric velocity components, which distorts the PDF of geopotential heights.

The kurtosis of the geopotential height is also mainly a function of latitude, although the patterns become

---

**Fig. 1.** Geopotential height: (a) 500-hPa DJF skewness, (b) 500-hPa DJF kurtosis, (c) 500-hPa JJA skewness, (d) 500-hPa JJA kurtosis, (e) zonally averaged DJF skewness, (f) zonally averaged DJF kurtosis, (g) zonally averaged JJA skewness, and (h) zonally averaged JJA kurtosis.
distorted over the large landmasses of the Northern Hemisphere. This time, around 30°, we see positive kurtosis bounded by negative values to the north and south. The negative values of the kurtosis become much more prominent in the JJA period, covering large swaths of territory over the Arctic, tropics, and Southern Hemisphere midlatitudes. Like the skewness patterns commented on earlier, the kurtosis also features slight seasonal variability, especially in the polar regions.

Looking at the zonally averaged vertical cross sections, we notice that in the lower troposphere (below 700 hPa) the positive skewness regions as a whole are not significant. We do see the large negative-valued bands between roughly 15° and 45° that extend from the surface of the earth to the tropopause. The tropics are nearly neutral no matter the time of the year, while the polar regions are only significantly positive in regions mostly above the tropopause. The strongest of these positive values are centered at the 250-hPa level south of 60°S in the austral summer. The cross sections are analogous to the ones found in Holzer (1996).

The kurtosis pattern during DJF is composed of two negative bands that span most of the depth of the troposphere around 50° in both hemisphere along with a smaller positive band in the Southern Hemisphere subtropics that goes up to 300 hPa. In the JJA period, the positive signal disappears and more negative areas show up, especially in the tropics and over the North Pole.

b. Relative vorticity

At 300 hPa, the skewness of the relative vorticity \( \zeta \) (Fig. 2) depends mainly on latitude. In the Northern Hemisphere, this value is positive equatorward of the mean climatological storm track and negative poleward of it. In the Southern Hemisphere, this pattern is reversed: positive skewness poleward of the storm track and negative skewness equatorward of it. As the DJF season turns into JJA, the pattern is almost the same. The main differences are that the bands of skewness have moved a few degrees north and that the magnitude of the skewness has decreased a small amount. This is likely due to the boreal winter having more baroclinicity than the austral winter (the Northern Hemisphere has twice as much land as the Southern Hemisphere, and land typically has larger meridional temperature gradients than water), which increases the strength and amplitude of the jet stream, resulting in higher vorticity values aloft.

The kurtosis is, on average, positive globally with neutral pockets scattered around the tropics. However, around the mean storm tracks, the kurtosis is strongly negative. One can see that during DJF there is a well-defined (and quite narrow) band of negative kurtosis going around the globe just north of Antarctica. In the Northern Hemisphere, this band is broader, again owing to the larger amplitude of Rossby waves in the boreal winter.

The zonally averaged cross section of \( \zeta \) shows that the equatorward bands of skewness seen in the horizontal plots extend from the surface to the tropopause. The poleward bands only show up above a certain level (roughly 600 hPa). All bands have their maximum value in the upper troposphere around 200 hPa. The kurtosis is positive everywhere except for a region of negative values in the upper troposphere at midlatitudes (the one associated with the storm track in the horizontal plots).

c. Quasigeostrophic potential vorticity

The skewness field of the 300-hPa QGPV, as seen in Fig. 3, is quite similar to the one for \( \zeta \) except that, at the 300-hPa level, the skewness is neutral poleward of the storm track instead of being significantly positive or negative.

Poleward of the storm tracks, we can see that the kurtosis of the QGPV is almost uniformly negative. Between the storm tracks, in the tropical region, both positive and negative kurtosis is found. Overall, the tropical kurtosis changes to a more negative value around JJA. This is especially true in North Africa and southern Asia.

Taking a zonal average of these moments reveals that the hemispheric character of the 300-hPa pattern extends throughout the entire column of air. Positive skewness regions in the Northern Hemisphere correspond to negative skewness regions in the Southern Hemisphere and vice versa. Below 300 hPa, the non-Gaussianity does not depend much on latitude, while above 300 hPa the tropopause seems represented by a region of opposite skewness as the rest of the non-Gaussian regions of that hemisphere.

The kurtosis patterns are simpler to describe. There exists a region of positive kurtosis in the tropics at virtually all levels of the troposphere, and this region extends to the midlatitudes closer to the earth’s surface. Poleward and upward of this positive kurtosis region, the sign switches to negative. By virtue of the kurtosis being an even-powered statistical moment, the sign of the original field (i.e., the QGPV field) does not matter and both hemispheres are close to mirror images of each other.

The local QGPV significance testing (see the appendix and, in particular, Fig. A2, described in detail below) is the only one that failed out of the nine variables. However, it is still possible that the patterns are real, given that the testing is done on local time series and the patterns are consistent spatially.

d. Zonal wind

When compared with the previous variables, both horizontal wind components are more Gaussian in nature. The zonal component (Fig. 4) has positive skewness around the equator and negative skewness (to
In the boreal summer, we see a decrease in the magnitude of the skewness in the Northern Hemisphere. The zonal wind features several mesoscale pockets of alternating skewness around the globe, possibly demonstrating that local topography has a large effect on wind distributions in the lower troposphere. Overall, the wind skewness is stronger over the oceans. The kurtosis of the zonal wind is decidedly harder to analyze.
In the Hadley cell (from 30°S to 30°N in the troposphere), the zonal wind skewness is generally positive except in the vicinity of the subtropical jets, where the skewness turns neutral and eventually negative in the center of the jet. However, the zonally averaged signal is quite weak globally. Zonal wind kurtosis is mostly insignificant in the troposphere except for narrow regions of negative values in the subtropical latitudes.
The skewness of the meridional wind (Fig. 5) at 925 hPa is dependent mostly on the hemisphere. In the Northern Hemisphere, values are neutral or positive except for local pockets of negative values. In the Southern Hemisphere, this trend is reversed. The small pockets of opposite skewness are found almost
exclusively over land. The kurtosis, on the other hand, is positive in the tropics and over Antarctica and negative elsewhere. Such a pattern suggests that the mean transport of energy toward the poles could be a source of non-Gaussianity for the meridional wind distribution worldwide.

Similar to the zonally averaged zonal wind, the zonally averaged meridional wind skewness is quite insignificant.
Averaging the positive and negative areas from the horizontal plot for each latitude gives skewness values near zero almost everywhere. Surprisingly, the kurtosis does show meaningful patterns. For example, in the tropics the kurtosis of the meridional wind is positive. Additionally, the Southern Hemisphere features positive kurtosis over Antarctica and negative kurtosis in the lower troposphere at midlatitudes.

**f. Horizontal wind speed**

Figure 6 shows the horizontal wind speed’s higher statistical moments. At the 925-hPa level, we notice that the skewness is positive throughout most of the spatial domain. There is a band of neutral skewness over 50°S and also a mix of negative and neutral skewness over the tropics. The main difference between the JJA and DJF seasons is a large decrease in the skewness values near the Indian Ocean, no doubt related to the onset of the monsoon season. The reason for the spatial dominance of the positive values is simple: the wind speed at any given point cannot be negative, and from an energetics point of view the wind usually stays close to zero; that is, extreme events related to wind speed are almost always on the right side of the axis, leading to positive skewness. The kurtosis of the wind speed, however, is mostly negative. The main exceptions to this are zonal bands near the poles and in isolated pockets in the tropics and subtropics. Monahan (2006a) found the statistical moments for the sea surface wind speed for the four seasons for four different datasets, one of which was the NCEP–NCAR reanalysis. Monahan’s plots look very similar to Fig. 6, with small differences accounted for by the fact that he looked at the sea surface wind and we looked at 925 hPa.

Averaged zonally, we clearly see that the skewness is neutral in places and positive in the rest. The kurtosis is neutral or positive over the poles and the tropics (especially in DJF) and negative over the midlatitudes. Monahan (2006b) raises an interesting point about the predicted PDF of a wind speed dataset. Essentially, the horizontal wind speed is the magnitude of the horizontal wind velocity vector, and this magnitude is the square root of the squares of the individual components summed together. It can be shown that the PDF of the wind speed can be approximated by the Weibull distribution, which assumes that the PDF of the individual wind components are roughly Gaussian. However, even the individual wind components feature some non-Gaussianity, so that may be part of the reason why observed low-level wind speed distributions are not fully “Weibullian” in nature.

**g. Vertical wind speed in pressure coordinates**

The quantity $\omega$ (Fig. 7), representing the vertical wind speed in pressure coordinates, has distinct non-Gaussian patterns restricted mostly by the instability of the atmosphere. Looking at the skewness panels, we see that almost all of the atmosphere at 500 hPa has a neutral or negative skewness for this variable. Only parts of Antarctica and smaller pockets where trade wind inversions are common have significant positive skewness. These patterns agree with those found in Petoukhov et al. (2008). The main explanation for this is that it is easier for a parcel of air to rise quickly than it is for it to sink quickly. Given that most of the warming caused by solar radiation occurs at the surface of the earth and the lowest layer of the atmosphere, parcels have a tendency to form thermal bubbles that are unstable relative to their environment. On the other hand, it is rarer that a parcel of air in the mid to upper troposphere will cool suddenly and drop toward the surface. Thus, in most places we see negative skewness, which means more upward motion extreme events than downward motion extreme events. In other words, upward motion is typically a smaller-scale event (both spatially and temporally) while downward motion is a larger-scale event.

The kurtosis of the vertical velocity is decidedly positive almost everywhere except for thin neutral bands around the equator and the midlatitude storm tracks most active in that season (northern in DJF; southern in JJA). This means that vertical motion usually belongs to one of two modes: the common stable mode in which air might be ascending or descending very slowly, and the rarer unstable mode in which air suddenly gets caught in a convective cell and launched upward at high velocities, resulting in a PDF with negative kurtosis.

**h. Air temperature**

The air temperature (Fig. 8) skewness at 925 hPa shows interesting patterns that are very dependent on the local geography and the locations of the climatological highs and lows. For example, in the Northern Hemisphere we observe positive skewness to the west and negative skewness to the east of climatological highs (such as the Icelandic and Aleutian lows). Conversely, we observe negative skewness to the west and positive skewness to the east of climatological highs (such as the Bermuda and Siberian highs). These vary with the strength of the storm tracks such that the skewness gradient is largest in the winter season for each hemisphere. In general, the temperature distribution above continents tends to have negative skewness. Over the oceans, the main difference between each season is the sign of the skewness in polar waters. While the air above the Southern Ocean is mainly positive in the DJF period, it switches to negative in JJA; and the roughly neutral air above the icy Arctic waters during DJF become overwhelmingly positive in the summertime.
On the other hand, the patterns of kurtosis in both DJF and JJA periods are quite noisy. Overall, there is a trend of negative kurtosis at mid- to high latitudes stretching around the globe concentrated over the ocean basins, but there are dozens of pockets with alternating signs found throughout the atmosphere, especially in the tropics. However, these are not randomly distributed; the values of kurtosis are quite high, showing that the significance of these values is not.

**FIG. 6.** Horizontal wind speed: (a) 925-hPa DJF skewness, (b) 925-hPa DJF kurtosis, (c) 925-hPa JJA skewness, (d) 925-hPa JJA kurtosis, (e) zonally averaged DJF skewness, (f) zonally averaged DJF kurtosis, (g) zonally averaged JJA skewness, and (h) zonally averaged JJA kurtosis.

On the other hand, the patterns of kurtosis in both DJF and JJA periods are quite noisy. Overall, there is a trend of negative kurtosis at mid- to high latitudes stretching around the globe concentrated over the ocean basins, but there are dozens of pockets with alternating signs found throughout the atmosphere, especially in the tropics. However, these are not randomly distributed; the values of kurtosis are quite high, showing that the significance of these values is not.
questioned, and there is a correlation of positive skewness with positive kurtosis found over the cold ocean currents. These values for air temperature skewness and kurtosis were calculated by Petoukhov et al. (2008), but differ slightly because of the different reanalysis data used. Since the non-Gaussianity of the temperature depends much on the land/ocean distribution, the skewness and

![Vertical Velocity in Pressure Coordinates](image)

**Fig. 7.** Vertical velocity in pressure coordinates (a) 500-hPa DJF skewness, (b) 500-hPa DJF kurtosis, (c) 500-hPa JJA skewness, (d) 500-hPa JJA kurtosis, (e) zonally averaged DJF skewness, (f) zonally averaged DJF kurtosis, (g) zonally averaged JJA skewness, and (h) zonally averaged JJA kurtosis.
Kurtosis values do not show interesting patterns when averaged zonally. However, we see persistent negatively skewed regions at the midlatitudes between 700 and 400 hPa that are close to the jet stream.

Specific humidity

Finally, the specific humidity (Fig. 9) is also non-Gaussian. In deserts and polar regions, skewness and
kurtosis are quite positive. We can explain this by saying that the mean specific humidity in those areas is usually close to zero, and because it cannot be negative, the extreme events must necessarily lie on the right side of the distribution. Thus, typically dry regions that get a few days of moisture advection feature positive skewness. The tropical oceans, on the other hand, have negative skewness values. The main difference between the

FIG. 9. Specific humidity (a) 925-hPa DJF skewness, (b) 925-hPa DJF kurtosis, (c) 925-hPa JJA skewness, (d) 925-hPa JJA kurtosis, (e) zonally averaged DJF skewness, (f) zonally averaged DJF kurtosis, (g) zonally averaged JJA skewness, and (h) zonally averaged JJA kurtosis. Note the regions of missing data over Antarctica.
DJF and JJA plots is the decrease in positive skewness over the high latitudes in the Northern Hemisphere in the summer.

The kurtosis field looks quite similar to the skewness field, except that in the kurtosis field the gradients are typically sharper since the mid- to high latitudes now feature negative values. Again, Petoukhov et al. (2008) shows similar patterns but in the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) dataset instead.

Averaged zonally, the specific humidity’s skewness is naturally quite positive above 800 hPa in the extra-tropics, where the air is usually dry. Near the surface and in the whole tropical troposphere, the values are closer to neutral and sometimes negative. Again, the kurtosis fields look quite similar to the skewness fields, even after taking the zonal average.

### 4. Summary and conclusions

In this paper we have presented a consistent (i.e., from one dataset) climatology of non-Gaussian atmospheric statistics for daily-sampled key atmospheric variables. While the first two moments, mean and variance, are the most common statistical quantities to include in a climatology, it is crucial to recognize that many important phenomena, such as extreme events, are described by higher-order statistics. That is, the variability of a non-Gaussian variable is not sufficiently described by its mean and variance; higher-order moments are needed to describe the characteristics of non-Gaussian phenomena. For example, for weather risk management applications, the detailed non-Gaussian statistics are required to make accurate statements about the probability of extreme events. In addition, the knowledge of higher-moment statistics is necessary to validate the ability of numerical models to reproduce extreme events. Until now, this has not been done systematically, and the non-Gaussian climatology presented here might serve as a one-stop benchmark for future model validations.

From the results presented here, it is evident that Gaussianity is actually quite rare in the atmosphere. In fact, for daily observations, non-Gaussian statistics are the norm and not an exception. Only if we calculate averages in space and/or time does the central limit theorem kick in and the statistics become more Gaussian. We can summarize the basic global skewness and kurtosis patterns of the analyzed key variables as follows:

- Geopotential height is the most widely studied variable in the literature, it features bands of negative skewness at midlatitudes and positive skewness around the equator and near the poles. These patterns are consistent vertically for the most part. Current hypothesized causes for the non-Gaussianity of this variable speak of blocking flow and the rectification caused by the nonlinearity of the advective processes.
- Relative vorticity’s higher statistical moments have their maximum values in the upper troposphere near the jet streams. Skewness is mainly a function of latitude, with four major bands of alternation values separated by the equator and the mean climatological storm tracks. Kurtosis is mainly positive around the globe except for narrow negative-valued bands around these storm tracks.
- For QGPV, although non-Gaussianity is visible at all locations globally, the skewness of the QGPV is generally confined to tropical or midlatitude regions. The sign of the skewness is almost entirely a function of the hemisphere: positive in the NH and negative in the SH. One possible explanation is the fact that the rotation of the earth, embodied in the Coriolis parameter that is found in the QGPV equation, is an asymmetric process.
- For zonal wind, in general, we see positive skewness near the tropics and negative skewness poleward of 30° latitude. These moments are stronger over the oceans. Zonally averaged, the moments almost disappear.
- For meridional wind, the skewness is mostly positive in the Northern Hemisphere and negative in the Southern Hemisphere, except for localized pockets of opposite-signed skewness over land. The kurtosis is positive in the tropics and Antarctica and positive elsewhere, but again the low-level meridional wind is affected by several smaller-scaled processes. Zonally averaged, only the kurtosis patterns have any significance.
- The skewness of horizontal wind speed is mainly positive year-round, but there are small pockets of negative skewness found in the tropics. Here we have to relate this distribution to the Weibull distribution; the sum of two independent quasi-Gaussian variables cannot result in a Gaussian distribution.
- For vertical velocity in pressure coordinates, we see almost the opposite of the horizontal wind speed: negative skewness is the rule of thumb for vertical velocity. We can assume that the large-scale atmosphere is stable and that the few unstable days during the year when upward vertical motion takes over are what causes the positive skewness throughout the world.
- Air temperature is a common variable with a complicated non-Gaussian spatial distribution. At first glance, there seems to be a correlation between the strength and location of the climatological highs and lows and the spatial distribution of the positive and negative skewness worldwide.
Specific humidity features, to a first approximation, negative skewness in the tropics and positive elsewhere. Specific humidity is typically low at high latitudes, but the occasional weather system brings moisture poleward and gives this region positive skewness.

Going beyond those broad skewness and kurtosis aspects, the patterns become quite complex with many interesting small-scale features. It seems that the small-scale features are (dependent on the actual variable) induced by land–sea contrasts, local sea surface temperatures, and topography.

At this point, it is important to note and realize that we only have a very rudimentary physical understanding of why we see the global patterns presented here. In fact, we do not have a fully developed theory of non-Gaussian statistics in the atmosphere. Only recently, several investigators (Sura and Sardeshmukh 2008; Sardeshmukh and Sura 2009; Sura and Perron 2010; Sura and Gille 2010; Sura 2011) have tried to fill this gap by analyzing local non-Gaussian oceanic and atmospheric variability in a stochastic–dynamical framework. Their theory attributes extreme anomalies to stochastically forced linear dynamics, where the strength of the stochastic forcing depends on the flow itself (multiplicative noise). Because stochastic theory makes clear and testable predictions about non-Gaussian variability, the multiplicative noise hypothesis can be verified by analyzing the detailed non-Gaussian statistics of oceanic and atmospheric variability. However, the theory is, at this point, not capable of explaining the large-scale skewness and kurtosis patterns. The theory only predicts the general form of the stochastic equation, but not the values of the parameters going into it. Thus, we still have a long way to go before we understand the statistics presented here.

While we established that the overall patterns of skewness and kurtosis do not change significantly over the reanalysis period (Sura and Perron 2010), we are interested in longer-range changes, both past (using the twentieth-century reanalysis data) and future (using Intergovernmental Panel on Climate Change). Besides our future plans, we also hope that the comprehensive non-Gaussian climatology of key atmospheric variables presented here will inspire students and researchers to help in answering the plethora of outstanding questions and challenges.

Acknowledgments. We thank the anonymous reviewers for their helpful comments that led to significant improvements of the manuscript. This project was funded by the National Science Foundation through Award ATM-0903579, “Assessing Atmospheric Extreme Events in a Stochastic Framework.”

APPENDIX

Monte Carlo Approach to Statistical Significance

Because the PDFs of our data are not necessarily close to Gaussian at each grid point, we cannot use the standard error defined in (5) and (6) to create confidence intervals around our calculated values of skewness and kurtosis. Instead, we use a Monte Carlo approach to test whether the observed higher-moment values can be recreated in (by construction) strictly Gaussian red-noise surrogate data fitted at every grid point (e.g., von Storch and Zwiers 1999; Wilks 2006). Thus, we can calculate local confidence intervals of skewness and kurtosis based on a Gaussian red-noise null hypothesis,
where the decorrelation time scale and variance are based on the original data.

Let us start with the Langevin equation of our null hypothesis:

\[
\frac{dx}{dt} = -\lambda x + \eta(t),
\]

(A1)

where \(\lambda\) is a positive constant representing the damping of the variable \(x\) and \(\eta(t)\) is Gaussian white noise with amplitude \(\sigma\). Of course, \(x\) will also be Gaussian by construction. Discretizing (A1) in time gives the first-order autoregressive process [AR(1)]

\[
x_{t+\Delta t} = (1 - \lambda \Delta t)x_t + \eta_{t+\Delta t},
\]

(A2)

where \(\eta_t\) is now discrete (with time increment \(\Delta t\)) Gaussian white noise with standard deviation \(\sigma_\eta\). In this form, it can be shown that the variance of \(x\) is

**FIG. A2.** Absolute values of the skewness 95% confidence intervals for DJF (a) 500-hPa geopotential height, (b) 300-hPa relative vorticity, (c) 925-hPa zonal wind, (d) 925-hPa meridional wind, (e) 925-hPa horizontal wind speed, (f) 500-hPa vertical velocity in pressure coordinates, (g) 925-hPa air temperature, and (h) 925-hPa specific humidity. These values were calculated from the Monte Carlo process described in the appendix.
These papers focus on low-frequency variability (e.g., described by the wave amplitude index) and, therefore, the statistical tests have to preserve the entire spectral structure.

Figures A2 and A3 show the absolute values of the confidence intervals for all nine variables of interest. For most variables, the 95% confidence interval of the skewness is smaller than the 0.12 shading interval found in the main skewness and kurtosis plots. Only geopotential height and air temperature feature a band of higher error values in the tropics, but at the mid and high latitudes the patterns are fully significant. The only exception to this is the error of the QGPV, which is only small in Northern Hemisphere storm tracks. Nevertheless, we argue that the skewness and kurtosis patterns found in Fig. 3 are significant because the patterns are consistent spatially. The Monte Carlo simulation can only determine significance of the QGPV locally in the time dimension. Therefore, we do not shade values between ±0.12 in the presented skewness and kurtosis plots, giving us approximately 95% confidence in skewness and 68% confidence in kurtosis.

REFERENCES


