Towards a more physically based approach to Extreme Value Analysis in the climate system

Prashant Sardeshmukh
Gil Compo
Cecile Penland
Chesley McColl

JOURNAL OF CLIMATE

Need for Caution in Interpreting Extreme Weather Statistics

Prashant D. Sardeshmukh and Gilbert P. Compo
CIRES, University of Colorado, and NOAA/Earth System Research Laboratory, Boulder, Colorado

Cécile Penland
NOAA/Earth System Research Laboratory, Boulder, Colorado

(Manuscript received 2 January 2015, in final form 1 June 2015)
Three Approaches to Extreme Value Analysis

1. Using Extreme Value Distributions (GEV, GP) fitted to extreme values.

2. Using distributions fitted to all values.

3. Using distributions estimated from large ensembles of model simulations.

Our new approach is basically Approach #2, but using a newly introduced general class of non-Gaussian distributions called Stochastically Generated Skewed distributions (SGS distributions).

SGS distributions are associated with linear Markov processes perturbed by asymmetric stochastic noise. Such Markov process models can be used to estimate sampling uncertainties and investigate space-time coherence statistics.

Sardeshmukh, Compo, and Penland, 2015 (Journal of Climate)
Non-Gaussianity has enormous implications for the probabilities of extreme values, and for our ability to estimate their changes using limited samples.

Consider Gaussian vs non-Gaussian PDFs, both $p(0,1)$, and shifted by 1 sigma:

**Gaussian PDFs**

- $P( x \geq 2 ) = 2.3\%$ and increases by a factor of 7
- $P( x \geq 4 ) = 0.003\%$ and increases by a factor of 43

**Non-Gaussian PDFs**

- Skewed and heavy-tailed with
- Skewness $S = 1$
- Kurtosis $K = 5$

- $P( x \geq 2 ) = 3.4\%$ and increases by only a factor of 4
- $P( x \geq 4 ) = 0.34\%$ and increases by only factor of 3
Sharply contrasting behavior of extreme values of non-Gaussian and Gaussian “red noise” linear Markov processes with a 1-day correlation scale, obtained in 10^5-day simulations (equivalent to 10^3 100-day winters).

Gaussian (red) and non-Gaussian (black, S=1, K=5) “SGS” PDFs with same mean and variance.

Non-Gaussian (S=1, K=5) vs Gaussian

Blue curves: Time series of decadal maxima (i.e. the largest daily value in each decade = Ten 100-day winters)

Orange curves: Time series of 99.5th decadal percentile (i.e. the 5th largest daily value in each decade)

Note the much larger extremes, and spurious long term trends of decadal extremes lasting a century or even longer, in the non-Gaussian case even in this statistically stationary world.
The PDFs of daily atmospheric anomalies are distinctively skewed and heavy tailed. This has large implications for assessing, modeling, and predicting extreme anomaly risks.

Left panels:

Skewness

\[ S = \frac{\langle x^3 \rangle}{\sigma^3} \]

of wintertime daily anomalies \( x \) in the 20CRv2 dataset (Compo et al 2011) over the 1872-2011 period

Right panels:

contrast of average histograms for positive and negative anomalies

Sardeshmukh, Compo, and Penland, 2015 (Journal of Climate)
Daily 850 hPa Temperature Probabilities of exceeding ± 2 sigma (DJF 1981-2005)
Both probabilities would be 0.022 if the distributions were Gaussian

\[ P(\ x < -2 \ sigma) \quad P(\ x > 2 \ sigma) \quad \text{Skew} \]

Era-Interim Reanalysis

20CR-v2c Reanalysis

Note the similarity of the exceedance probability and skewness patterns
The distinctively skewed and heavy tailed character of the observed distributions is well captured by SGS distributions (of which Gaussian distributions are special cases).

Stochastically Generated Skewed distributions (SGS distributions)

\[
p(x) = \frac{1}{N} \left[ (Ex + g)^2 + b^2 \right]^{-(1 + \frac{1}{E^2})} \exp \left[ \frac{2g}{E^2 b} \arctan \left( \frac{Ex + g}{b} \right) \right]
\]

with parameters \( E, g, \) and \( b, \) are physically associated with linear Markov processes driven by correlated additive and multiplicative noise (CAM noise), and are particularly relevant to the analysis of weather and climate extremes.

One tail of all such distributions is heavier, and the other lighter, than a Gaussian tail beyond about 1.73 standard deviations, and its heaviness (lightness) is proportional to the Skew \( S. \)
SGS distributions are associated with damped linear Markov processes perturbed by asymmetric stochastic noise

\[ dx = -\left[ \left(1 + \frac{1}{2} E^2 \right) x + \frac{1}{2} E g \right] \lambda dt + \left[ b \eta_1 + (Ex + g) \eta_2 \right] \sqrt{\lambda} dt \]

where \( \lambda > 0 \) is a damping constant, and \( \eta_1 \) and \( \eta_2 \) are Gaussian white noises with zero mean and unit spectral amplitude. The \((Ex + g)\) noise term is referred as Correlated Additive and Multiplicative noise (CAM noise). The asymmetric magnitude of this term for positive and negative \( x \) is the only mechanism for generating skew in this process (Sardeshmukh and Sura, J. Climate 2009)

The mean, variance, and skewness of \( x \) are

\[ \langle x \rangle = 0, \quad \langle x^2 \rangle = \sigma^2 = \frac{g^2 + b^2}{2 - E^2}, \quad S = \frac{\langle x^3 \rangle}{\sigma^3} = \frac{2E}{(1 - E^2)} \frac{g}{\sigma} \]

As \( E \to 0 \), the PDF \( p(x) \) approaches a Gaussian PDF. This becomes obvious from expressing the PDF of the standardized anomalies \( \tilde{x} = x / \sigma \) in this limit as

\[ p(\tilde{x}) = p_0 \exp \left[ -\frac{\tilde{x}^2}{2} + \frac{S}{6} (\tilde{x}^2 - 3) + O(E^2) \right] \]

where \( p_0 \to 1/\sqrt{2\pi} \) as \( E \to 0 \). The first term represents a standard Gaussian. The second term is the leading-order departure from Gaussianity.

This implies that \( p(x) = p(-x) \) at \( x = \sqrt{3} \sigma \), beyond which one tail becomes heavier than the other, and also heavier than a Gaussian.
SGS distributions arise from asymmetric Correlated Additive and Multiplicative (CAM) noise forcing of a linear system

Additive noise forcing alone yields a Gaussian PDF

Uncorrelated Additive + Multiplicative noise yields a symmetric and heavy-tailed non-Gaussian PDF

Correlated Additive + Multiplicative noise yields a skewed and heavy-tailed non-Gaussian “SGS” PDF
A simple rationale for Correlated Additive and Multiplicative (CAM) noise

In a quadratically nonlinear system with “slow” and “fast” components $x$ and $y$, the anomalous nonlinear contribution to $\frac{dx'}{dt}$ may be written as:

\[
(xy)' = \bar{x} y' + x' \bar{y} + x'y' - \bar{x'y'}
= x'y' + (\bar{x} + x')y' - \bar{x'y'}
\]

This is the CAM noise term
And this is its mean “Noise-Induced Drift”

Now if $y'$ is approximated as stochastic noise, the second term becomes a CAM noise forcing term, and the third mean noise-induced drift term can be explicitly represented in terms of the noise amplitude parameters.
Is there a deeper physical basis for the relevance of such

*Stochastically* Generated Skewed distributions in the climate system?

We consider the question of how skewness arises in the simplest energy-preserving dynamical system.

(or more generally, in a system with a quadratic invariant)
Skewed dynamics of a **nonlinear** oscillator

*Sardeshmukh and Penland (Chaos, 2015)*

The skewness arises from the system trajectories lingering near a **line of zero coupling** (which does not exist in a linear system)

\[
\frac{dx}{dt} = -\lambda_x x + \Omega(x, y) y
\]

\[
\frac{dy}{dt} = -\lambda_y y - \Omega(x, y) x + S\eta
\]

\[
\Omega(x, y) = B + Cx + Dy
\]

If \(y\) is stochastically forced, and damped much more strongly than \(x\), then it can be approximated as Gaussian white noise, and the PDF of \(x\) approaches an **SGS** distribution.
The PDF of the slow variable $x$ converges to the SGS distribution if the time scale separation between $x$ and $y$ is large ($\lambda_y / \lambda_x = 300$)

**BLACK** curves: analytical SGS PDF of negative and positive values

**RED** curves: Numerical simulation of the nonlinear system

**BLUE** curve: reference Gaussian

Note that 1. the positive and negative tails cross at $|x'| \sim 1.73$
2. the positive and negative tails have the same power-law slope beyond $|x'| > 10$

*From Sardeshmukh and Penland (Chaos, 2015)*
The SGS versus Gaussian character of mid-tropospheric vertical velocity PDFs has large implications for precipitation PDFs, especially their tails.

To a first approximation, the precip PDF is simply the PDF of (appropriately scaled) upward vertical velocities. This explains why precip PDFs are more skewed in areas of mean descent, and why their tails are often fatter than Gamma PDF tails.
The simple $P$-$w$ relationship explains why the PDF of even seasonal mean precipitation $P$ is more skewed in areas of mean mid-tropospheric descent.

**Color shading:**
- Skewness of seasonal mean $P$

**White areas:**
- Areas of mean descent

**Grey areas:**
- Areas of mean ascent

**Right panels:** Skewness of seasonal mean $P$ versus climatological mean 500 hpa $W$ in five different observational datasets (Sardeshmukh, Compo, Penland 2015)
SGS dynamics also explain the **asymmetry in the persistence of positive and negative anomalies beyond specific amplitude thresholds**, even though the predictable dynamics are linear. This is due to the asymmetry in the probability of initial conditions \( p(x_0) \) and \( p(-x_0) \).

**Number of events that exceed a set of amplitude thresholds for at least T days**

Note that positive anomalies (dark grey) are *less persistent* than negative anomalies (orange), but large positive anomalies (\( > 2\sigma \)) are *more persistent* than large negative anomalies (\( < -2\sigma \)).

This behavior is consistent with that found by Dole and Gordon (1983) in the persistence of observed 500 mb height anomalies in the north Pacific sector (and also in other sectors).
But what is the problem with fitting “universal” extreme value distributions (GEV and/or Generalized Pareto) to extreme values in Extreme Value Analysis?

*No problem*: 1. *IF* the full distribution is unknown  
   (but why bother if we know it to be, say, Gaussian or SGS?)
2. *IF* the “block” sizes are large enough for the GEV limit to be appropriate
   (This can be a concern: see figure below)

The figure shows that the GEV fit to the “true” extreme value distribution of block maxima of Gaussian (CYAN) and SGS (BLACK) processes is poor for block sizes $N < 200$. (note the log scale)

3. *IF* the limited sampling of tail values is not a concern  
   (but it can be a concern: see next slide)
4. *IF* the PDF tails are well approximated as GP distributions  
   (this is probably OK. SGS tails are identical to the GP PDF, and the shape parameter of the GP/GEV PDF is simply related to parameter $E$ of the SGS PDF)
The limited sampling of extreme values ("block maxima") can be a concern in GEV analysis.

The figure shows that for limited (25-winter) records, the PDFs of winter (100-day) maxima of daily values (Extreme Value PDFs) can be more accurately estimated by fitting SGS distributions to all daily values than by fitting GEV distributions to just the maximum values from each winter.

The thick black curves show the "true" extreme value PDFs of the winter maxima in $10^8$ day integrations of Gaussian (TOP) and SGS (BOTTOM) Markov processes with a 1-day correlation time scale.

Note that the SGS approach yields much smaller error bars even in the Gaussian parent case!
The parameters of GEV/GP distributions are also hard to interpret, because they are not related in a simple way to the parameters of the underlying physical process.

For example, for the SGS process:

\[ dx = -\left[ \left( 1 + \frac{1}{2}E^2 \right)x + \frac{1}{2}Eg \right] \lambda dt + [b \eta_1 + (Ex + g)\eta_2] \sqrt{\lambda} dt, \]

With the PDF:

\[ p(x) = \frac{1}{N}[(Ex + g)^2 + b^2]^{-\frac{1+1/E^2}{2}} \exp \left[ \frac{2g}{E^2b} \arctan \left( \frac{Ex + g}{b} \right) \right]. \]

The GEV CDF for large block sizes \( N \) is:

\[ F_G(\hat{x}) = \exp \left\{ - \left[ 1 + \xi \left( \frac{\hat{x} - \hat{\mu}}{\hat{\sigma}} \right) \right]^{-1/\xi} \right\} \]

\[ \xi = \frac{E^2}{2 + E^2}, \quad \hat{\sigma} = \frac{\xi}{E} \left( \frac{C \xi}{E} \right)^{\xi} N^{\xi}, \]

\[ \hat{\mu} = \frac{\hat{\sigma}}{\xi} - \frac{g}{E} = \frac{1}{E} \left( \frac{C \xi}{E} \right)^{\xi} N^{\xi} - \frac{g}{E}. \]

And the GP PDF for large thresholds \( X_T \) is:

\[ p(x \mid x \geq X_T) = \frac{1}{\tilde{\sigma}} \left[ 1 + \xi \left( \frac{x - X_T}{\tilde{\sigma}} \right) \right]^{-\frac{1+1/\xi}{\xi}} \]

\[ \xi = \frac{E^2}{2 + E^2}, \quad \tilde{\sigma} = \left( X_T + \frac{g}{E} \right) \xi. \]
An example: We find no significant differences between SGS distributions fitted to daily NP and NAO indices in the first and second halves of the 1871 to 2011 period.

Results for the first (RED) and second (BLUE) halves using the 20CRv2 dataset

Boxes: Raw histograms
Curves: SGS distributions
Gray: sampling uncertainty

7-yr running means of the wintertime NP and NAO indices in the 20CRv2 and other datasets.

Note that there are large multidecadal variations, but no trend over the full record.

Sardeshmukh, Compo, and Penland, 2015 (Journal of Climate)
Summary: Three Approaches to Extreme Value Analysis

1. Using Extreme Value Distributions (GEV, GP) fitted to extreme values.

2. Using distributions fitted to all values.

3. Using distributions estimated from large ensembles of model simulations.

Our new approach is basically Approach #2, but which goes beyond considering just mean shifts to a better accounting of changes in distribution width and shape as well, using a general class of non-Gaussian distributions called Stochastically Generated Skewed distributions (SGS distributions).

SGS distributions are associated with linear Markov processes perturbed by asymmetric stochastic noise. Such Markov models can be used to estimate sampling uncertainties and investigate space-time coherence statistics.

Sardeshmukh, Compo, and Penland, 2015 (Journal of Climate)