Estimating the Black Swan Population: A Null Hypothesis for the Statistics of Extreme Events in Weather and Climate

Philip Sura

Florida State University

EOAS - COAPS

Introduction

- An extreme event is most commonly defined in terms of the **non-Gaussian tail of the data's probability density function (PDF).**
- Understanding extremes is important because climate and weather **risk assessment depends on knowing the tails of PDFs.**
- The general problem of understanding extremes is their scarcity: **We have to extrapolate from the well sampled center of a PDF to the scarcely or unsampled tails.**
- The extrapolation into the uncharted tails of a PDF can divided into **three major categories (not mutually exclusive).**

Three Methods to Study Extremes

- The **statistical** approach

  provides methods to extrapolate from the well sampled center to the unsampled tails of a PDF using mathematical (asymptotic) arguments.

  While based on sound mathematics, it does not provide much insight into the physics of extreme events.

- The **empirical-physical** approach

  uses physical reasoning based on empirical knowledge to provide a basis for the extrapolation into the scarcely sampled tails of the PDF.

  It lacks the mathematical rigor of the statistical method, but provides valuable physical insight into relevant real world problems.

- The **numerical modeling** approach

  aims to estimate the the tails of the PDF by integrating a general circulation model (GCM) for a very long period.

  Here the weakness lies in the largely unknown ability of a model to reproduce the correct statistics of extreme events.

A more General Perspective of Extreme Events in Climate

- The study of extreme climatic events has been **largely empirical:**

  Most investigators used observations or model output to estimate the probabilities of extreme events without actually addressing the detailed dynamical/physical reason for the shape of the PDF.

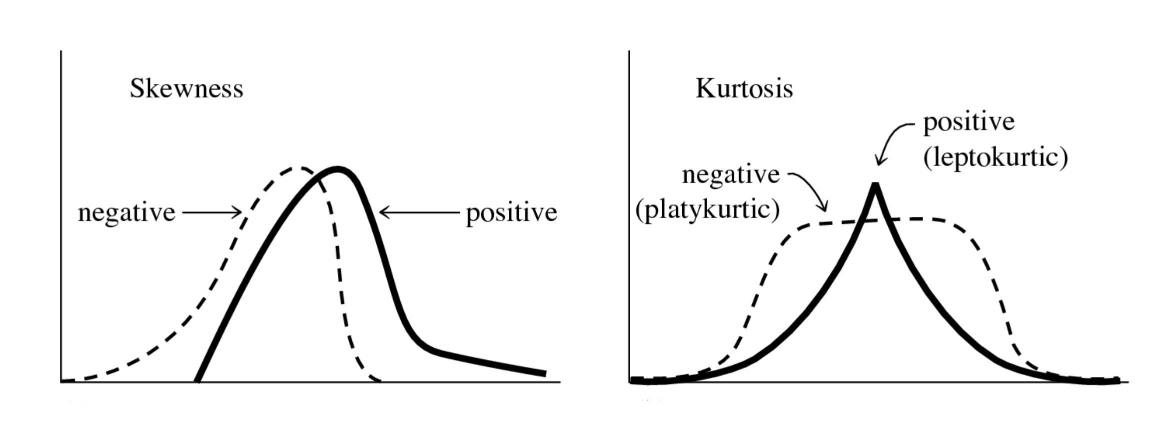
- we try to address the dynamical/physical reason for the shape of the PDF.
- we hope to better understand the dynamics/physics and statistics of extreme events in climate from first dynamical principles.

We provide a null hypothesis derived from first principles !

Measures of Non-Gaussianity: Skewness and Kurtosis

$$skew \equiv \frac{\langle x'^3 \rangle}{\sigma^3}$$

$$kurt \equiv \frac{\langle x'^4 \rangle}{\sigma^4} - 3$$



Skewness and **kurtosis** are non-dimensional measures describing the shape of a probability density function (PDF)

For data drawn from any PDF we have (Pearson 1916):

$$kurt \geq skew^2 - 2$$

Observations

We analyzed the non-Gaussian statistics of many important climate variables:

- Atmospheric Geopotential Height
- Atmospheric Vorticity and Potential Vorticity
- Sea Surface Temperature
- Sea Level Heights

And found the following properties in almost all of them:

Skewness-Kurtosis Link

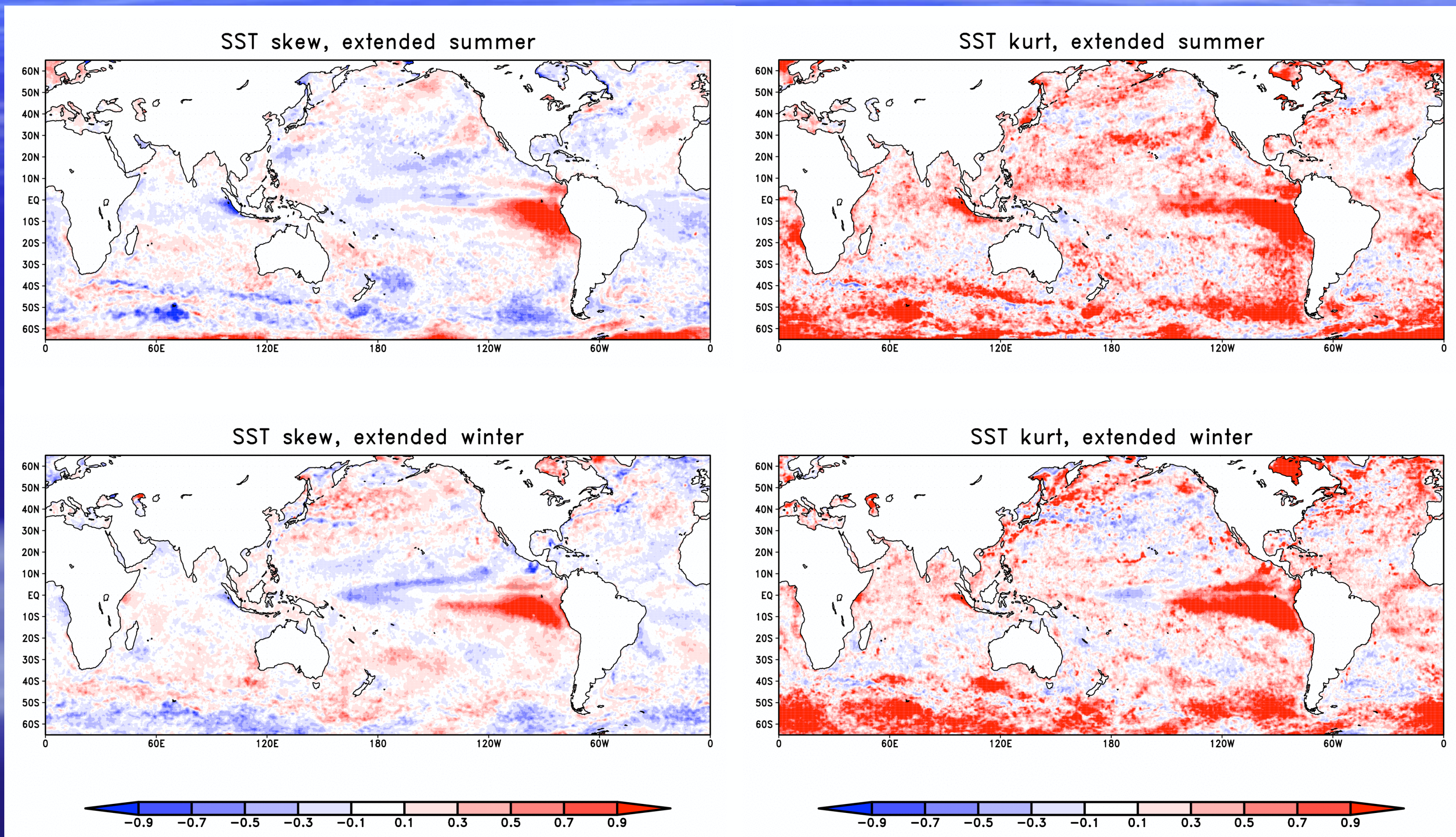
$$kurt \geq \frac{3}{2} skew^2 - r$$

Power-Law Tails

$$p(x) \propto x^{-\alpha}$$

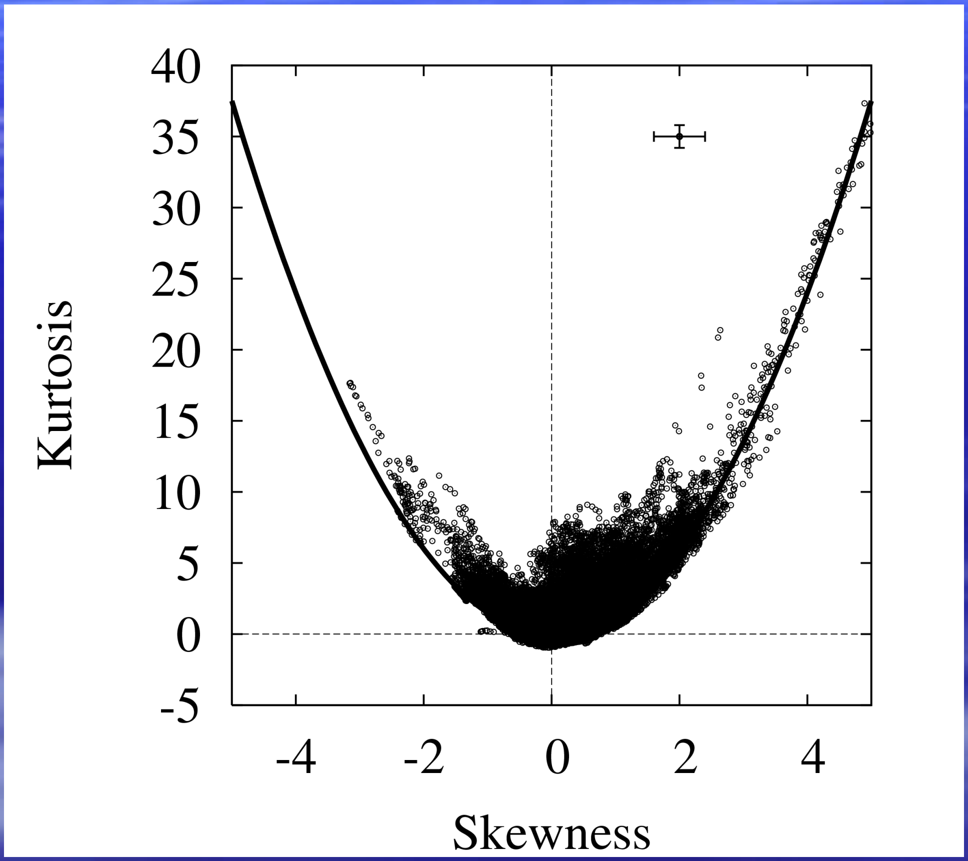
Skewness and Kurtosis - SST Anomalies

Daily AVHRR SSTs blended with in situ data, 1985-2005



Dataset from Reynolds et al. (2006)

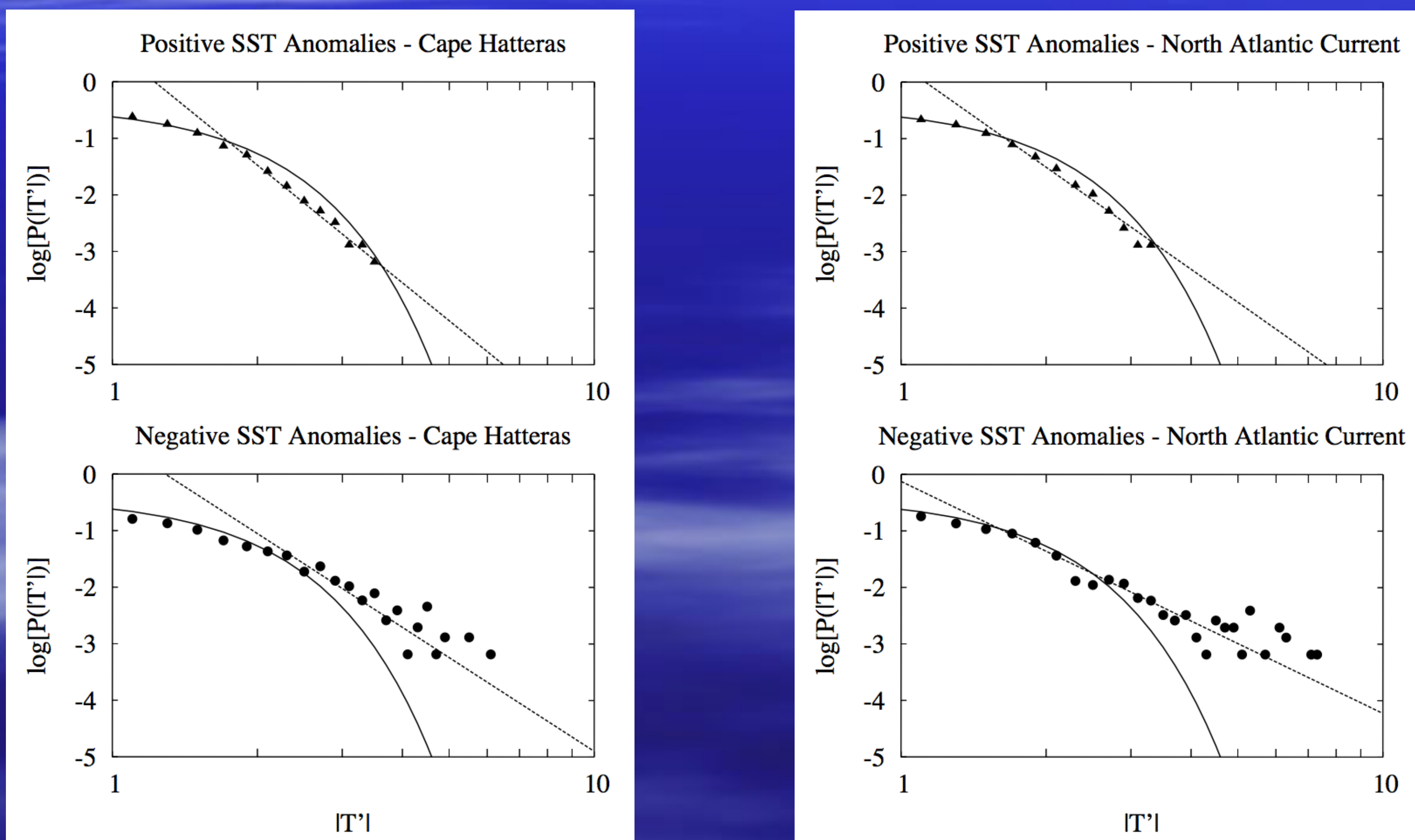
Skewness and Kurtosis - SST Anomalies



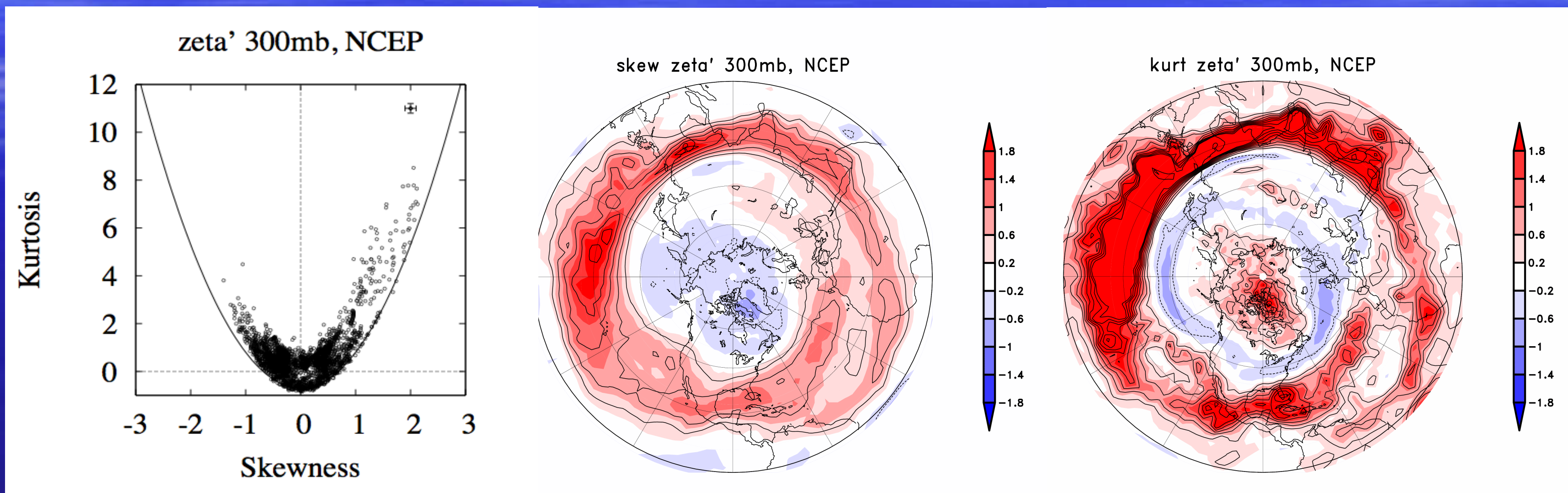
$$kurt \geq \frac{3}{2} skew^2$$

PDFs Follow a Power-Law in the Gulf Stream System

$$p(x) \propto x^{-\alpha}$$



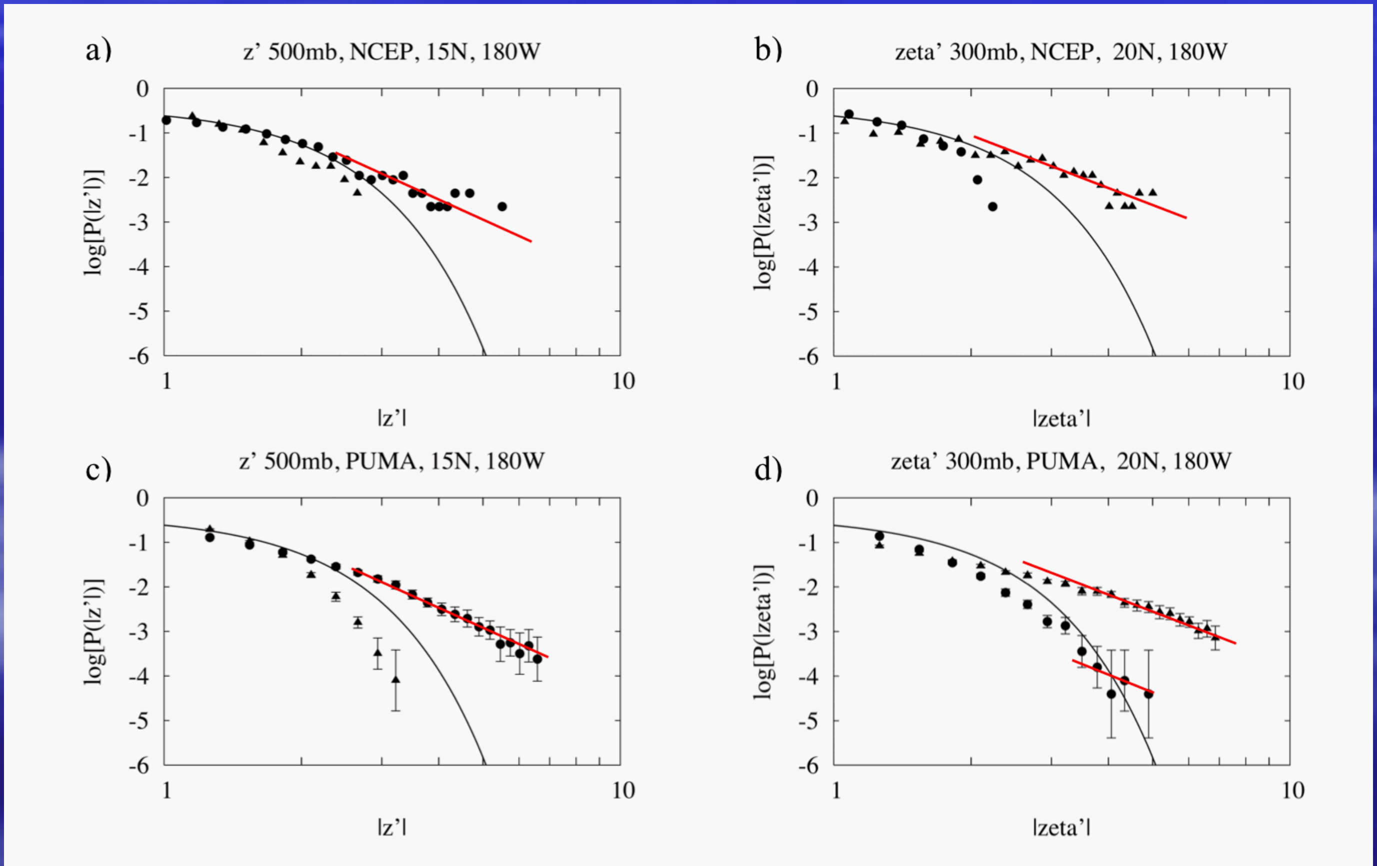
Atmospheric
300 mb Vorticity Anomalies



$$kurt \geq \frac{3}{2} skew^2 - 0.6$$

Observed and Simulated Atmospheric Power-Law PDFs

$$p(x) \propto x^{-\alpha}$$



Math in a Nutshell…

In a quadratically nonlinear system with "slow" and "fast" components *x* and *y*, the anomalous nonlinear tendency for the slow component has terms of the form

$$\frac{dx'}{dt} = \ldots + x'\bar{y} + \bar{x}y' + x'y' - \overline{x'y'}$$

If we parameterize the fast anomaly *y'* as a stochastic white-noise process *η*, we obtain a stochastic differential equation with correlated additive and multiplicative (CAM) noise:

$$\frac{dx'}{dt} = \ldots + x'\bar{y} + (\bar{x} + x')\eta - \overline{x'\eta}$$

In general, the linearized univariate equation can be written:

$$\frac{dx'}{dt} = -\lambda_{eff}x' + (1 - \phi x')F' + R'$$

Properties of our Model

- The (excess) kurtosis *K* is always greater than 1.5 times the square of the skewness *S* minus an off-set *r* :

$$kurt \geq \frac{3}{2} skew^2 - r$$

- The PDF *p(x)* has power-law tails:

$$p(x) \propto x^{-\alpha}$$

- The observed skewness-kurtosis link,
- the observed power-law

are consistent with our model !

Summary and Conclusions

We presented a null hypothesis derived from first principles that can explain the remarkable observed quadratic relationship between kurtosis and skewness, as well as the power-law tails of climate PDFs.

Null Hypothesis for the Statistics of Extreme Events