CLIMATE EXTREMES AND GLOBAL WARMING: A STATISTICIAN'S PERSPECTIVE

Richard L. Smith
Department of Statistics and Operations Research
University of North Carolina, Chapel Hill
rls@email.unc.edu

Statistics and Climate Change: Is There a Consensus?
136th Annual Meeting
American Public Health Association
San Diego, October 27 2008

Author's Discosure Statement

Portions of this work have been funded by NOAA, NSF and EPA.

I would like to acknowledge the contributions of Thomas Knutson (NOAA), Evangelos Evangelou (UNC), Gabi Hegerl (Edinburgh) and Amy Grady (NISS), as well as the author team of Climate Change Science Program Synthesis and Assessment Product 3.3 ("Weather and Extremes in a Changing Climate"). However all opinions stated herein are mine.

I have no financial conflicts of interest.

BACKGROUND

OVERVIEW OF THIS TALK

IPCC (2007) has claimed there is *unequivocal* evidence of global warming induced by human activities. This has been endorsed by (among many others) the American Statistical Association.

However, there remains much disagreement over whether other forms of observed climate change are anthropogenic in origin (or are real trends at all).

I focus here on two problems for which I think there is still room for doubt:

- Precipitation extremes
- Increasing intensity or frequency of Atlantic tropical cyclones

In both cases, my objective is to argue that statisticians have an important role to play in resolving key scientific issues.

PART 1:
PRECIPITATION EXTREMES

Many papers in the climatological literature have documented increases in the frequency of extreme precipitation events. A representative paper is Groisman *et al.* (2005, *Journal of Climate*). However, with a few notable exceptions, this literature has not made use of *extreme value theory*, which is a well-established branch of statistics.

My objective here is to explore what we might learn by applying advanced methods of extreme value theory to this problem.

DATA SOURCES

- NCDC Rain Gauge Data (Groisman 2000)
  - Daily precipitation from 5873 stations
  - Select 1970–1999 as period of study
  - 90% data coverage provision — 4939 stations meet that

- NCAR-CCSM climate model runs
  - 20×41 grid cells of side 1.4^{o}
  - 1970–1999 and 2070–2099 (A1B scenario)

- PRISM data
  - 1405×621 grid, side 4km
  - Elevations
  - Mean annual precipitation 1970–1997

EXTREME VALUES METHODOLOGY

The essential idea is to fit a probability model to the exceedances over a high threshold at each of ≈5000 data sites, and then to combine data across sites using spatial statistics.

The model at each site is based on the *generalized extreme value distribution*, interpreted as an approximate tail probability in the right hand tail of the distribution.

$$\Pr\{Y \geq y\} \approx \delta_t \left(1 + \xi \frac{y-\mu}{\psi}\right)_+^{-1/\xi} \text{ for large } y,$$

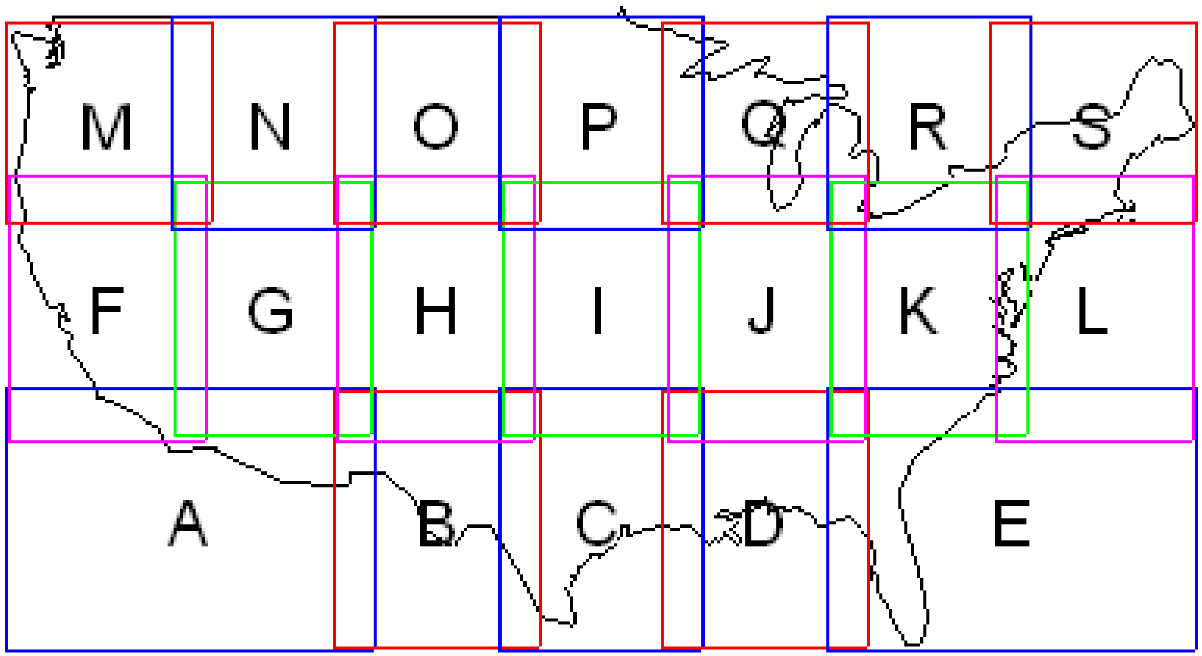
Here $x_+ = \max(x, 0)$, δ_t is a time increment (here 1 day based on a time unit of 1 year) and the parameters μ, ψ, ξ represent the location, scale and shape of the distribution. In particular, when $\xi > 0$ the marginal distributions have a Pareto (power-law) tail with power $-1/\xi$.

TEMPORAL AND SPATIAL DEPENDENCE

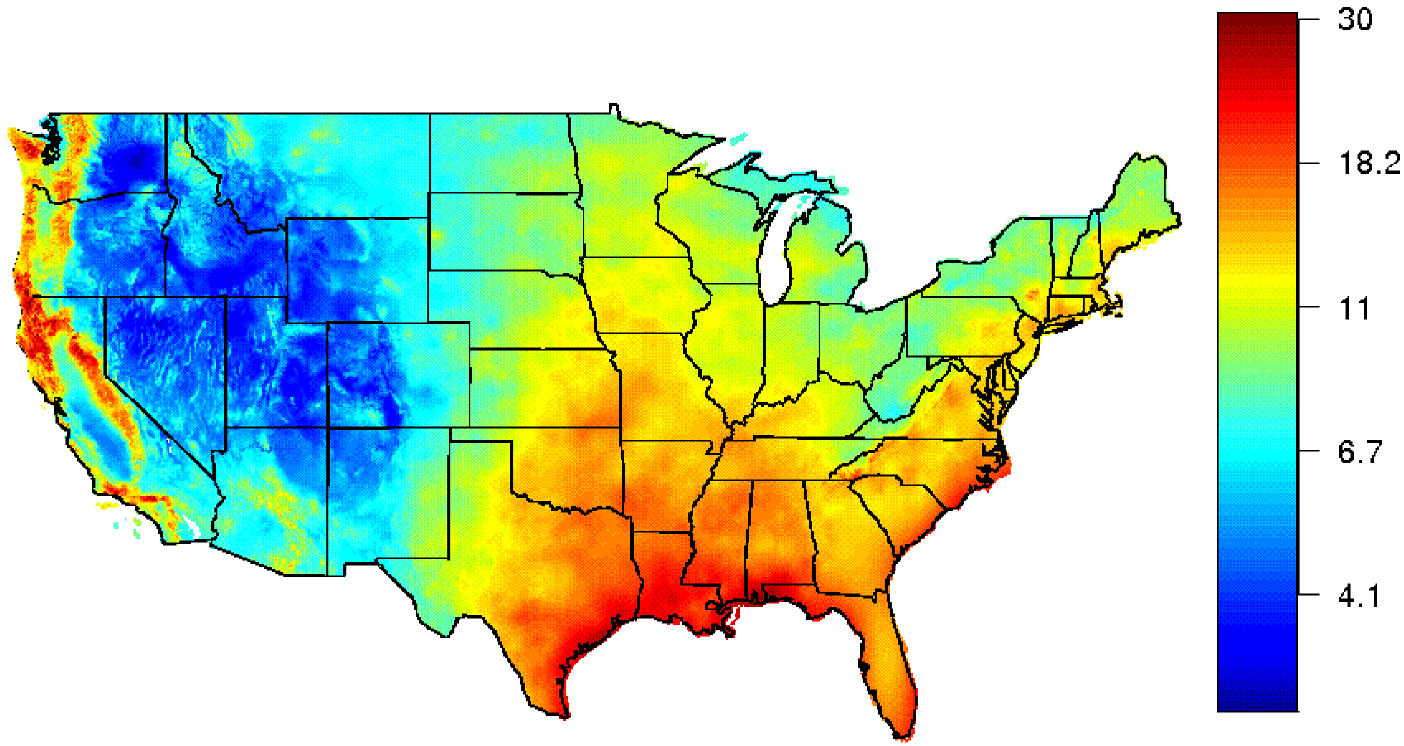
Here, we make two extensions of the basic methodology.

First, the parameters μ, ψ, ξ are allowed to be time-dependent through covariates. This allows a very flexible approach to seasonality, and we can also introduce linear trend terms to examine changes in the extreme value distribution over the time period of the study.

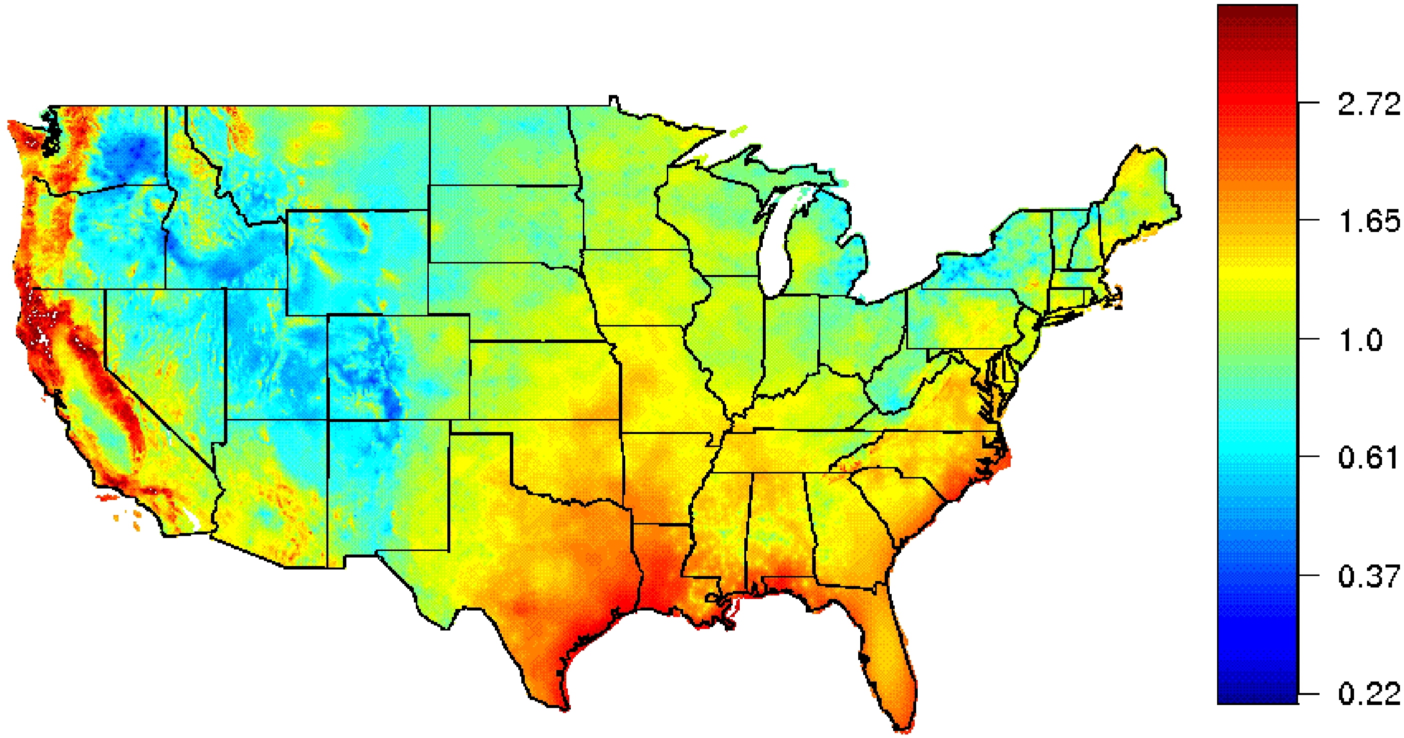
The second extension is *spatial smoothing*: after estimating the 25-year return value at each site, we smooth the results across sites by a technique similar to kriging. We allow for spatial nonstationarity by dividing the US into 19 overlapping boxes, and interpolating across the boundaries.



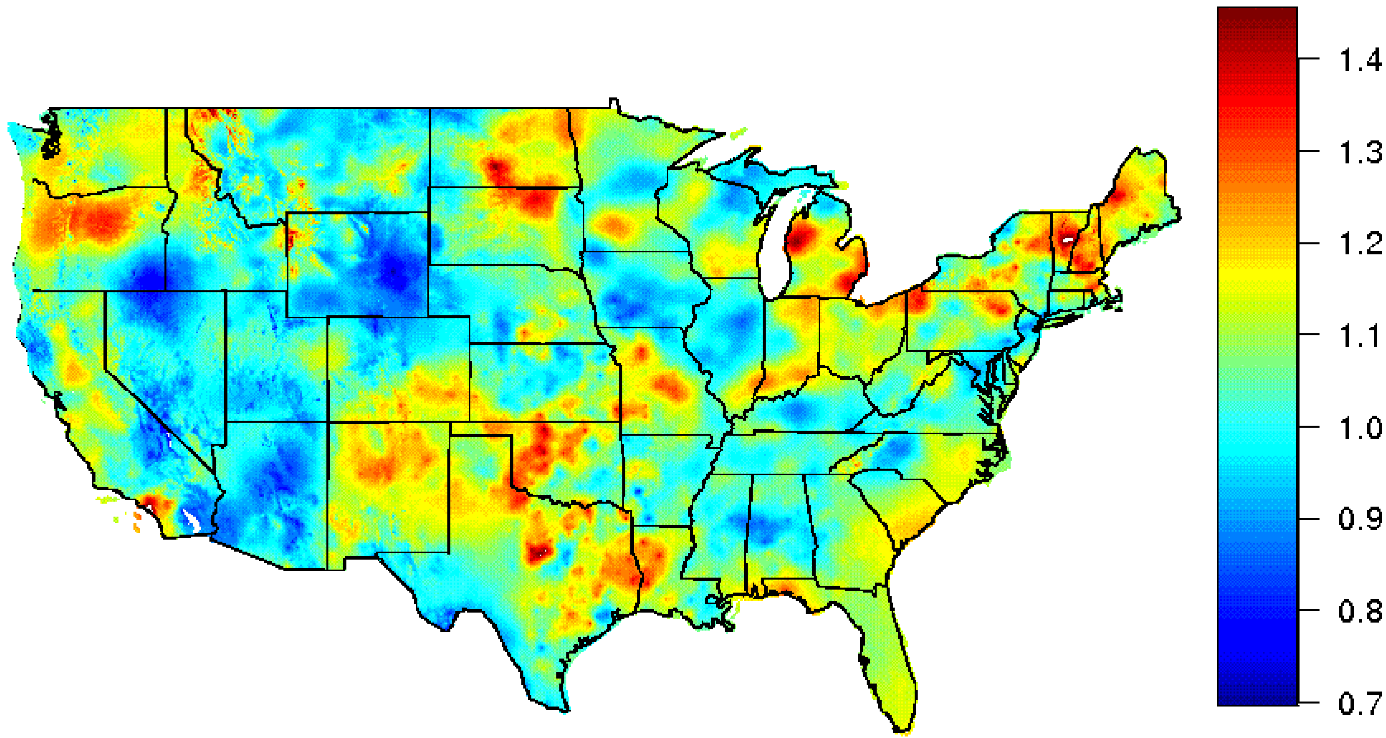
Continental USA divided into 19 regions



Map of 25-year return values (cm.) for the years 1970–1999



Root mean square prediction errors for map of 25-year return values for 1970–1999



Ratios of return values in 1999 to those in 1970, using a statistical model that assumes a linear trend in the GEV model parameters

	Change	RMSPE		Change	RMSPE
A	−0.01	.03	K	0.08***	.01
B	0.07**	.03	L	0.07***	.02
C	0.11***	.01	M	0.07***	.02
D	0.05***	.01	N	0.02	.03
E	0.13***	.02	O	0.01	.02
F	0.00	.02	P	0.07***	.01
G	−0.01	.02	Q	0.07***	.01
H	0.08***	.01	R	0.15***	.02
I	0.07***	.01	S	0.14***	.02
J	0.05***	.01			

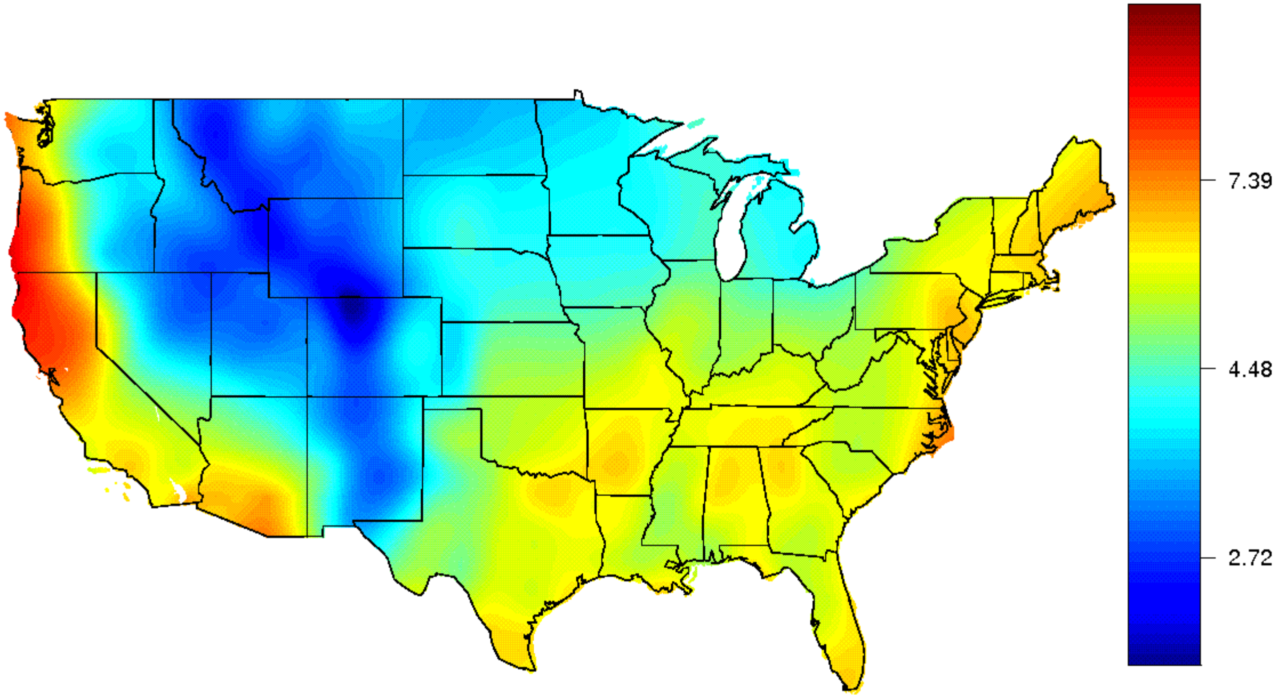
For each grid box, we show the mean change in log 25-year return value (1970 to 1999) and the corresponding standard error (RMSPE)

Stars indicate significance at 5%*, 1%**, 0.1%***.

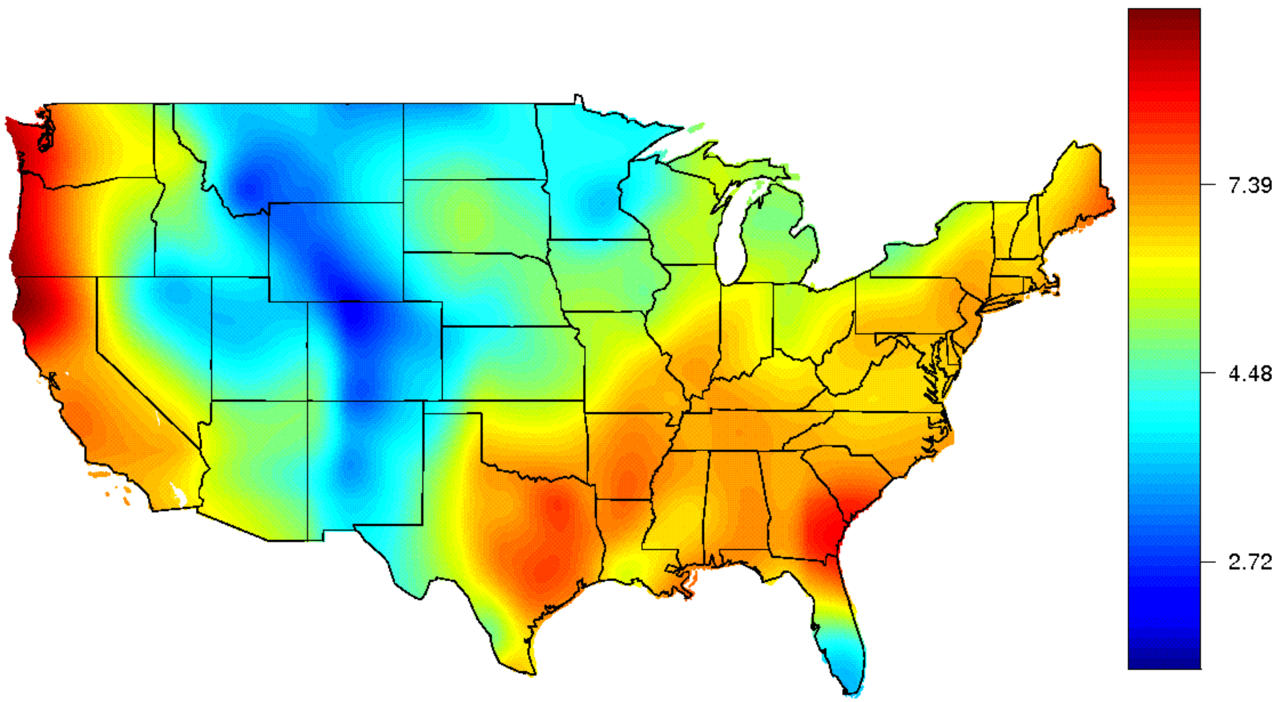
14 of 19 regions are statistically significant increasing: the remaining five are all in western states

We can use the same statistical methods to project future changes by using data from climate models.

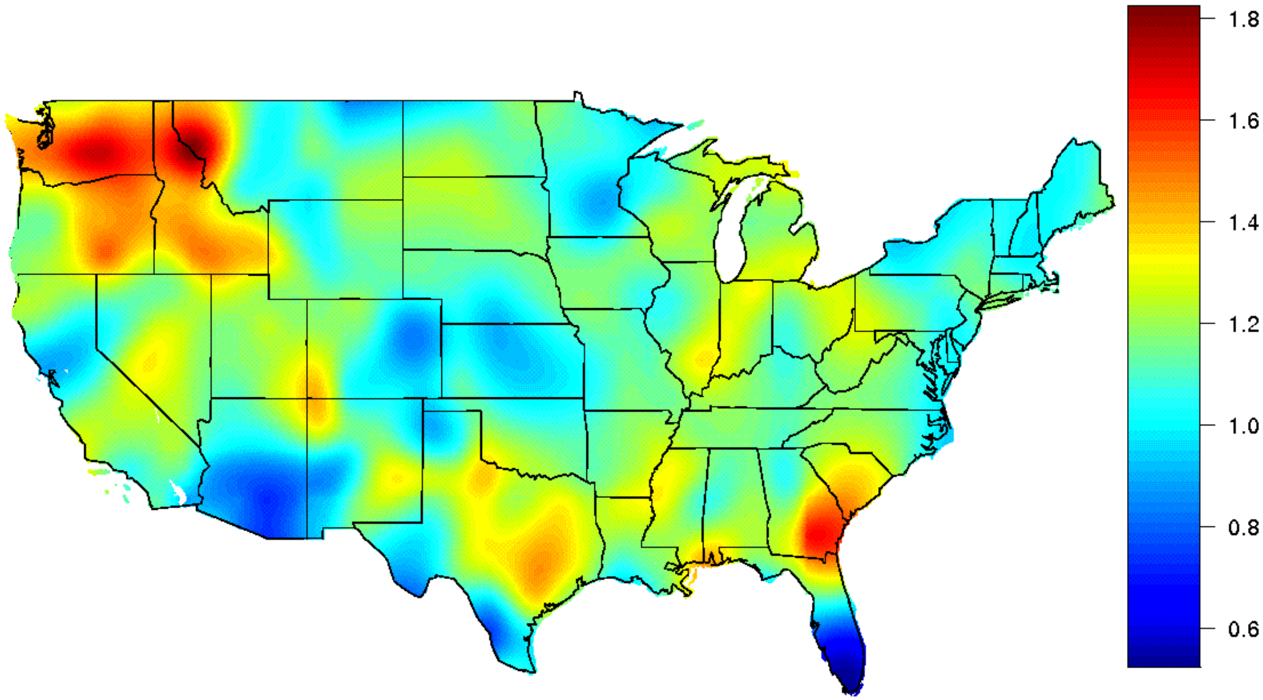
Here we use data from CCSM, the climate model run at NCAR.



Return value map for CCSM data (cm.): 1970–1999



Return value map for CCSM data (cm.): 2070–2099



Estimated ratios of 25-year return values for 2070–2099 to those of 1970–1999, based on CCSM data, A1B scenario

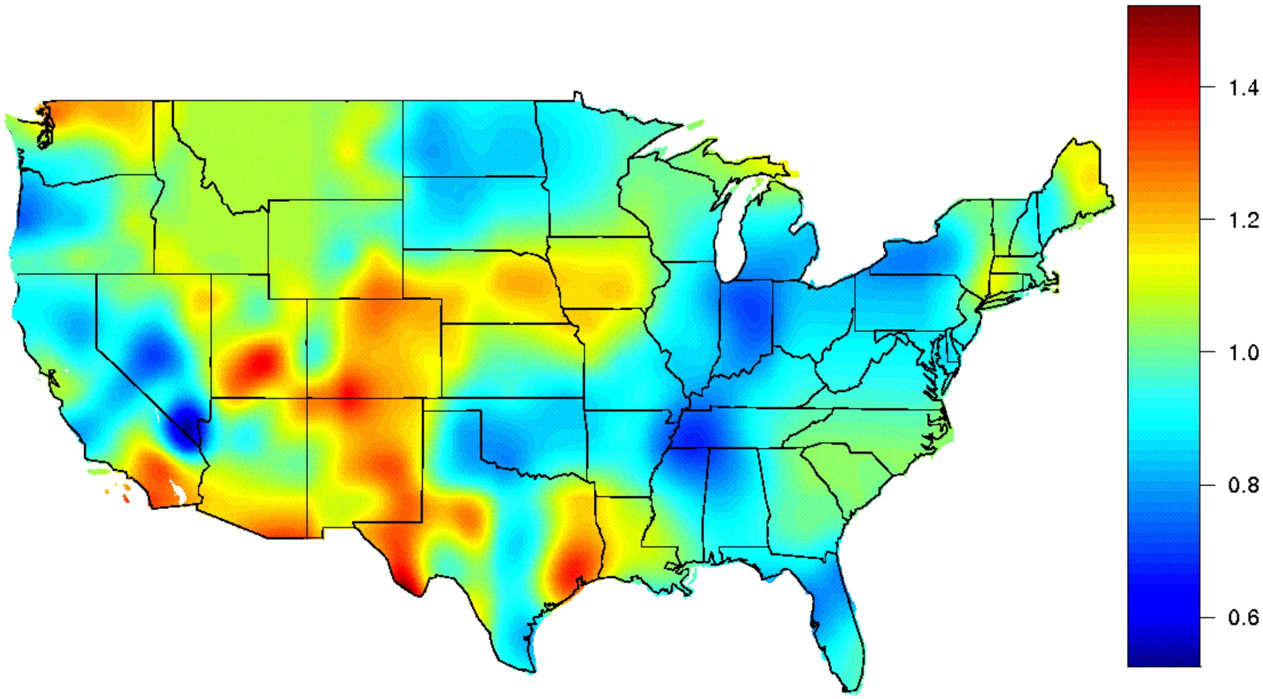
The climate model data show clear evidence of an increase in 25-year return values over the next 100 years, as much as doubling in some places.

So what's wrong with this picture?

Two issues concern me.

1. Although the overall increase in observed precipitation extremes is similar to that stated by other authors, the spatial pattern is completely different. There are various possible explanations, including different methods of spatial aggregation and different treatments of seasonal effects.

2. Even when the *same* methods are applied to CCSM data over 1970–1999, the results are different.



Extreme value model with trend: ratio of 25-year return value in 1999 to 25-year return value in 1970, based on CCSM data

MY CONCLUSIONS FROM THIS

1. More sophisticated statistical methods really can produce different answers. It's not just a case of added statistical complexity for the sake of it.

2. This kind of analysis does show up the limitations of climate models in projecting precipitation extremes.

PART 2: HURRICANES AND GLOBAL WARMING

Numerous recent papers have documented an increase in tropical cyclones and Atlantic hurricanes over the past 40 years, in particular in the intensity of the most extreme hurricanes. These increases have been tied to increases in sea surface temperature (SST), and this has led many hurricane researchers to conclude there is a direct causal link to greenhouse gas-induced global warming. However, there are a number of reasons why a simple trend analysis may not be sufficient to resolve the question.

nature
Vol 436|4 August 2005|doi:10.1038/nature03906

LETTERS

Increasing destructiveness of tropical cyclones over the past 30 years

Kerry Emanuel[1]

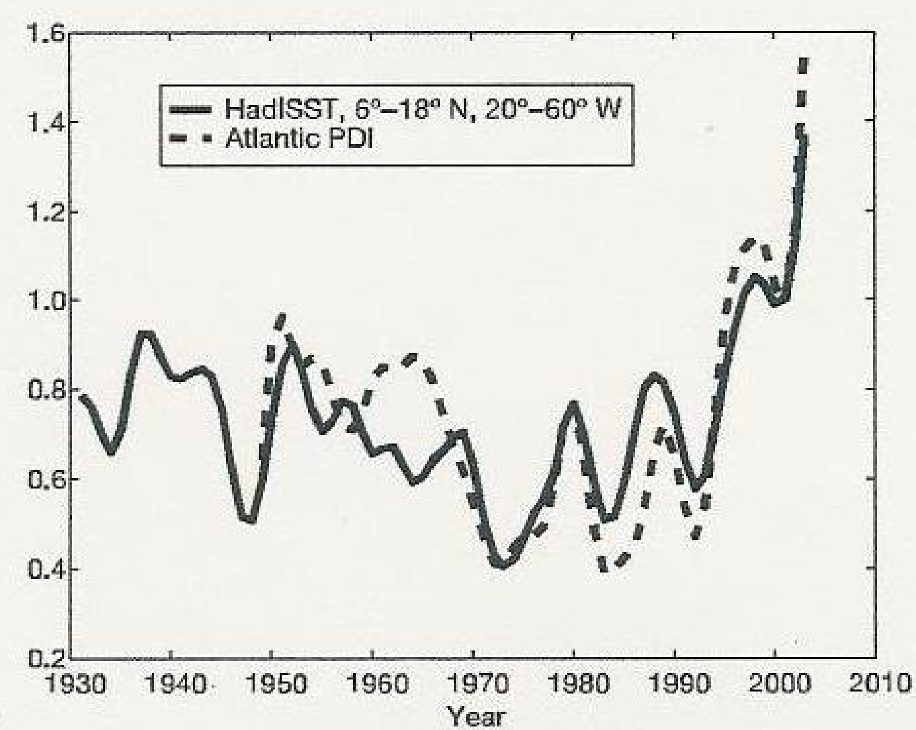


Figure 1 | A measure of the total power dissipated annually by tropical cyclones in the North Atlantic (the power dissipation index, PDI) compared to September sea surface temperature (SST). The PDI has been multiplied by 2.1×10^{-12} and the SST, obtained from the Hadley Centre Sea Ice and SST data set (HadISST)[22], is averaged over a box bounded in latitude by 6° N and 18° N, and in longitude by 20° W and 60° W. Both quantities have been smoothed twice using equation (3), and a constant offset has been added to the temperature data for ease of comparison. Note that total Atlantic hurricane power dissipation has more than doubled in the past 30 yr.

Changes in Tropical Cyclone Number, Duration, and Intensity in a Warming Environment

P. J. Webster,[1] G. J. Holland,[2] J. A. Curry,[1] H.-R. Chang[1]

We examined the number of tropical cyclones and cyclone days as well as tropical cyclone intensity over the past 35 years, in an environment of increasing sea surface temperature. A large increase was seen in the number and proportion of hurricanes reaching categories 4 and 5. The largest increase occurred in the North Pacific, Indian, and Southwest Pacific Oceans, and the smallest percentage increase occurred in the North Atlantic Ocean. These increases have taken place while the number of cyclones and cyclone days has decreased in all basins except the North Atlantic during the past decade.

16 SEPTEMBER 2005 VOL 309 SCIENCE www.sciencemag.org

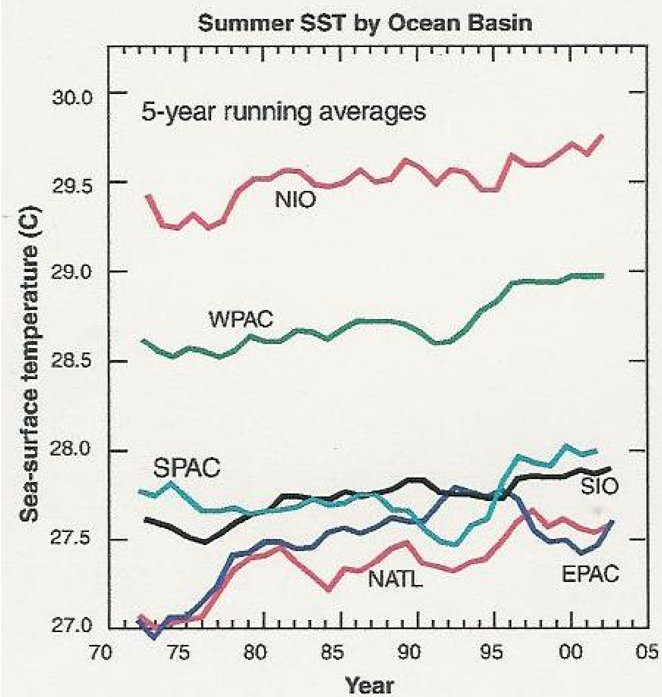


Fig. 1. Running 5-year mean of SST during the respective hurricane seasons for the principal ocean basins in which hurricanes occur: the North Atlantic Ocean (NATL: 90° to 20°E, 5° to 25°N, June-October), the Western Pacific Ocean (WPAC: 120° to 180°E, 5° to 20°N, May-December), the East Pacific Ocean (EPAC: 90° to 120°W, 5° to 20°N, June-October), the Southwest Pacific Ocean (SPAC: 155° to 180°E, 5° to 20°S, December-April), the North Indian Ocean (NIO: 55° to 90°E, 5° to 20°N, April-May and September-November), and the South Indian Ocean (SIO: 50° to 115°E, 5° to 20°S, November-April).

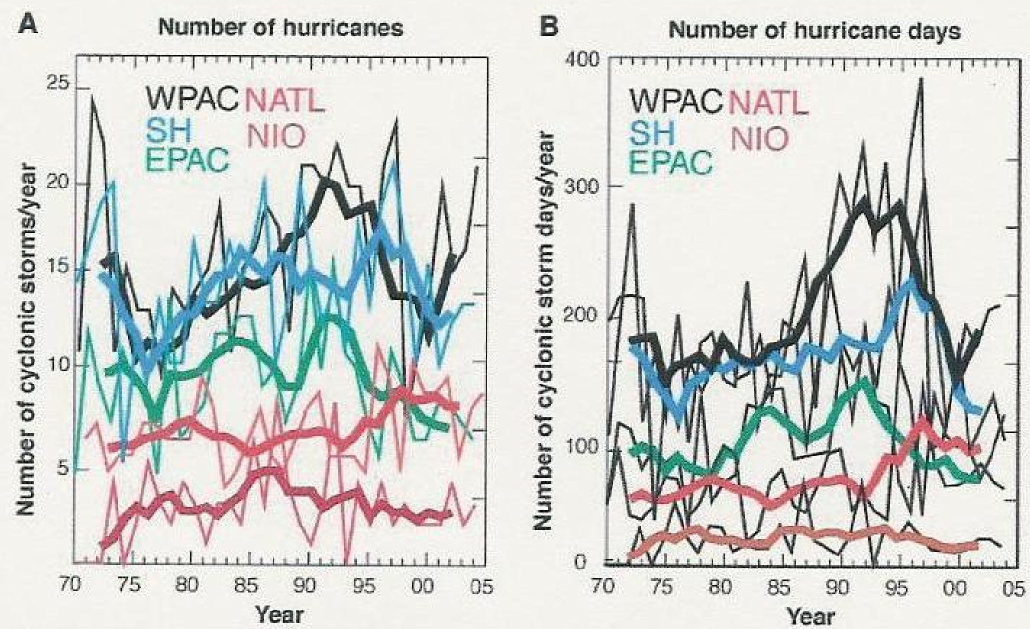
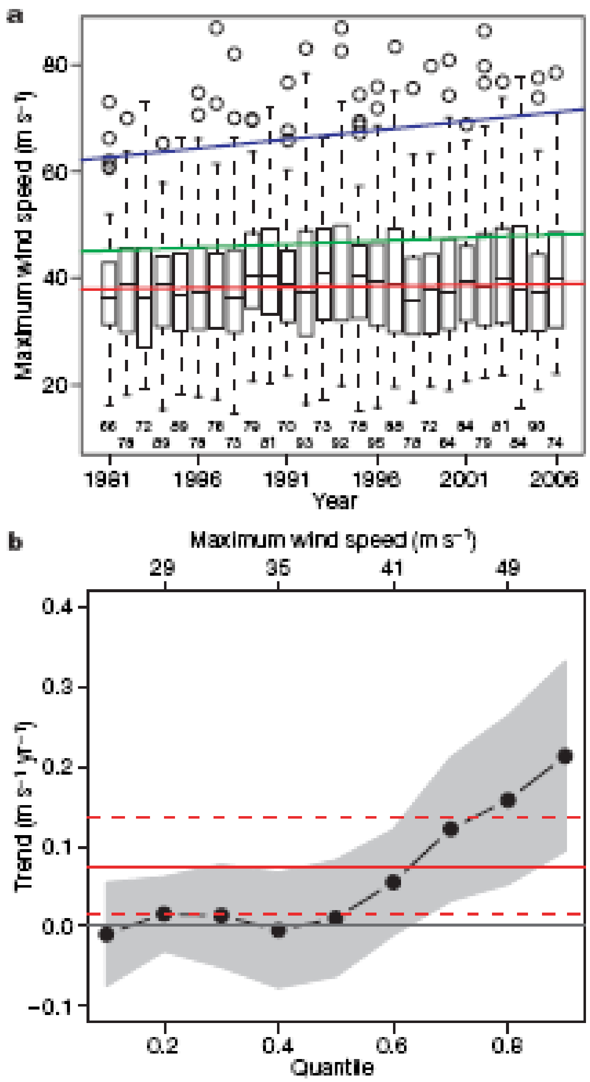


Fig. 3. Regional time series for 1970–2004 for the NATL, WPAC, EPAC, NIO, and Southern Hemisphere (SIO plus SPAC) for (**A**) total number of hurricanes and (**B**) total number of hurricane days. Thin lines indicate the year-by-year statistics. Heavy lines show the 5-year running averages.

The increasing intensity of the strongest tropical cyclones

James B. Elsner[1], James P. Kossin[2] & Thomas H. Jagger[1]



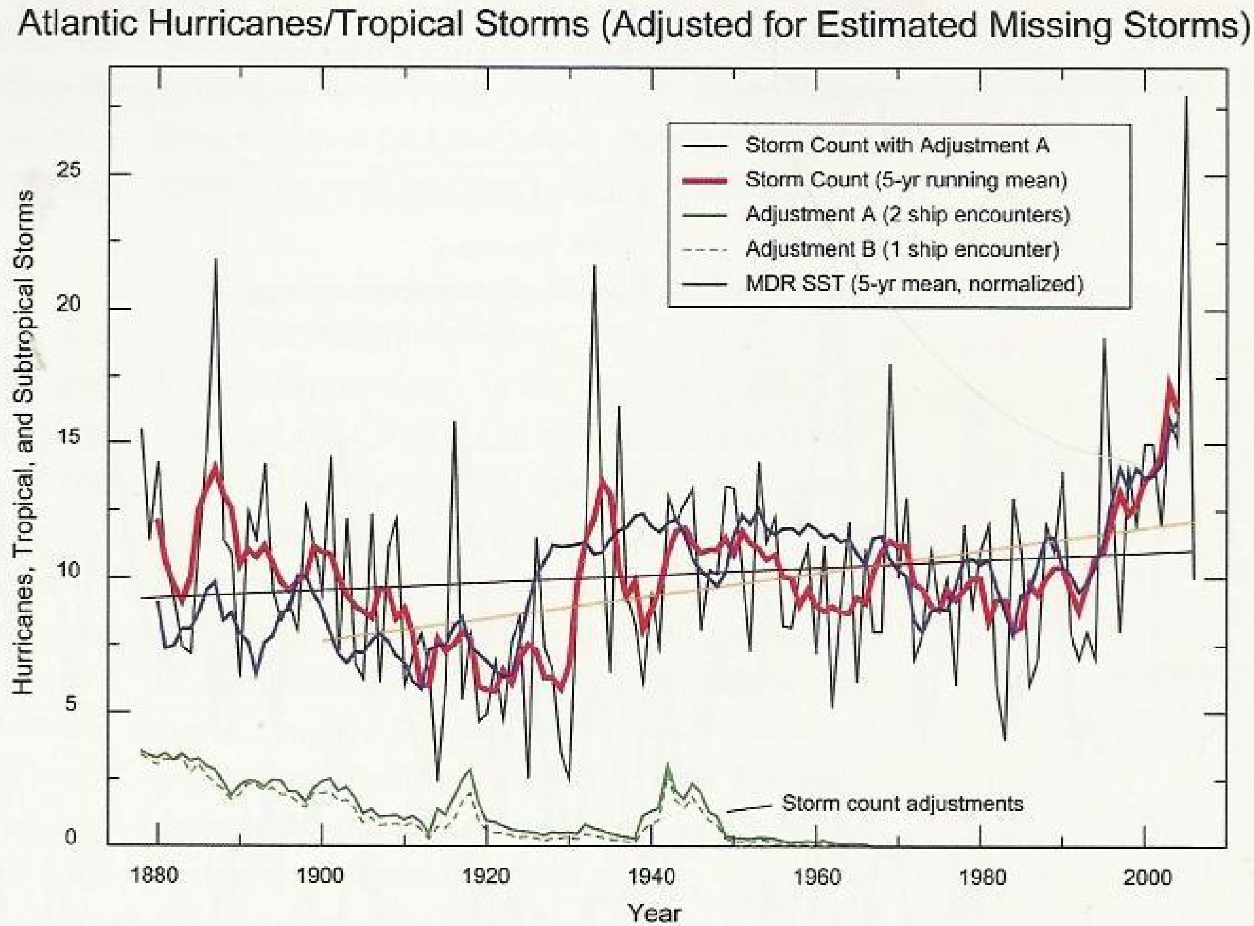


Figure 2.16 Atlantic hurricanes and tropical storms for 1878-2006, using the adjustment method for missing storms described in the text. Black curve is the adjusted annual storm count, red is the 5-year running mean, and solid blue curve is a normalized (same mean and variance) 5-year running mean sea surface temperature index for the Main Development Region of the tropical Atlantic (HadISST, 80-20°W, 10-20°N; Aug.-Oct.). Green curve shows the adjustment that has been added for missing storms to obtain the black curve, assuming two simulated ship-storm "encounters" are required for a modern-day storm to be "detected" by historical ship traffic for a given year. Straight lines are least squares trend lines for the adjusted storm counts. (Adapted from Vecchi and Knutson, 2007).

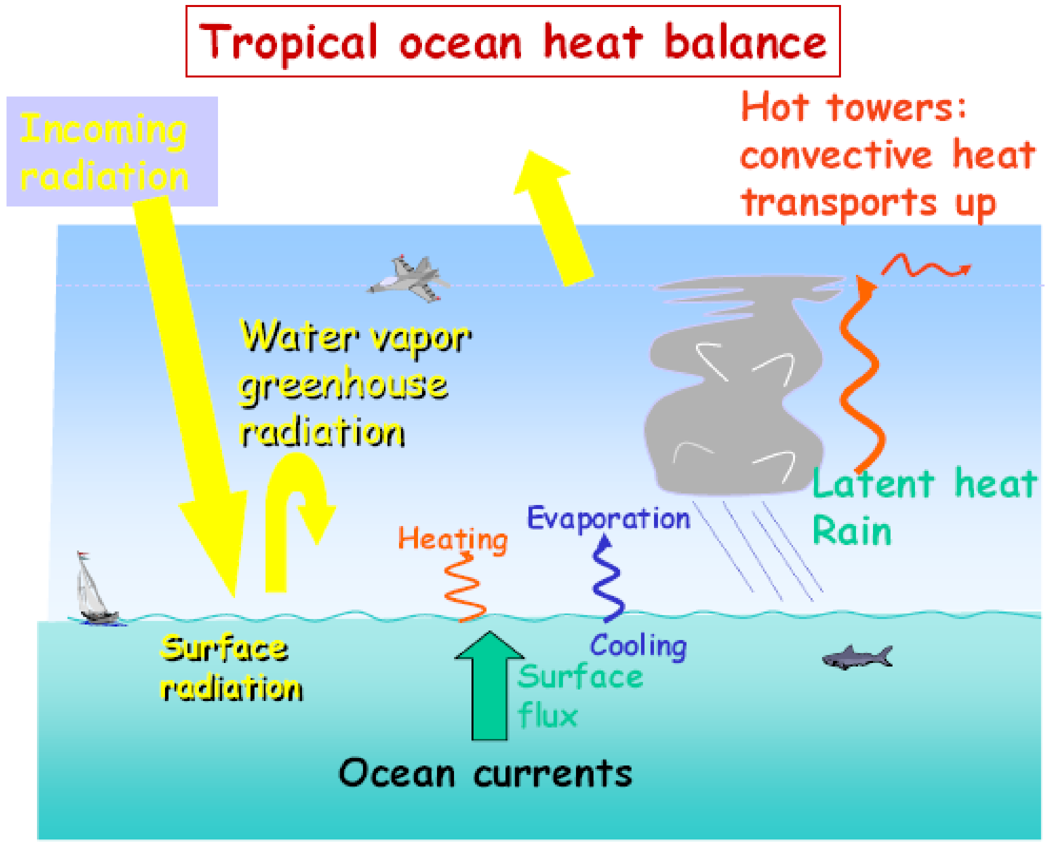
Some Basic Facts About Hurricanes

(From a presentation by Kevin Trenberth)

- SSTs $> 26^{o}$C (80^{o}F)
- High water vapor
- Weak wind shear
- Weak static stability
- Pre-existing disturbance

Also:

Large variability from year to year
El Ninõ means more activity in Pacific, less in Atlantic
Large interdecadal variability in Atlantic (AMO)



Source: Presentation by Kevin Trenberth

- Hurricanes play a key role in climate, but are not in models and are not parameterized

- Competition between thunderstorms and convection, but these are not resolved and are treated as sub-grid-scale phenomena

- Climate models have premature onset of convection

- Result: Existing models are likely to underpredict hurricanes

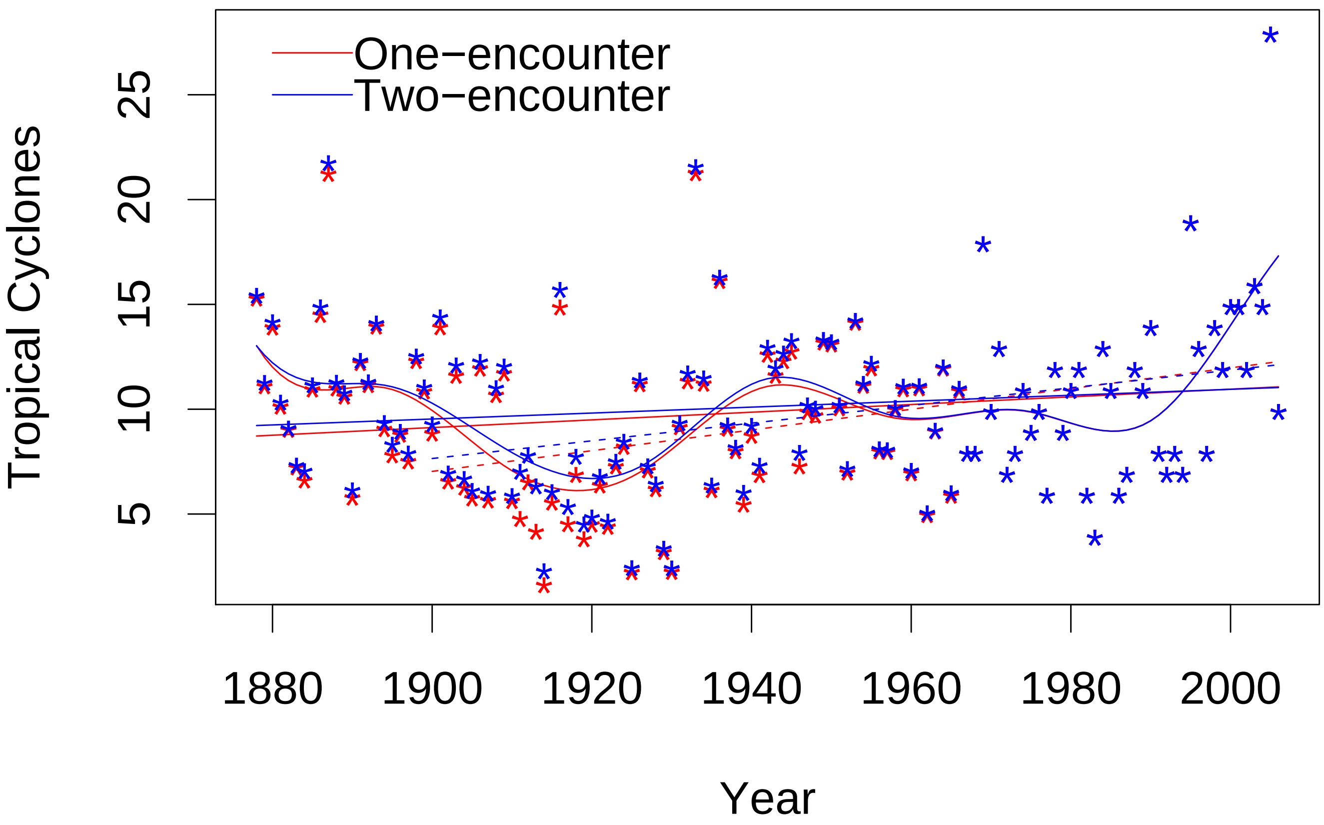
This analysis compares two reconstructions of the TC series (Vecchi and Knutson)

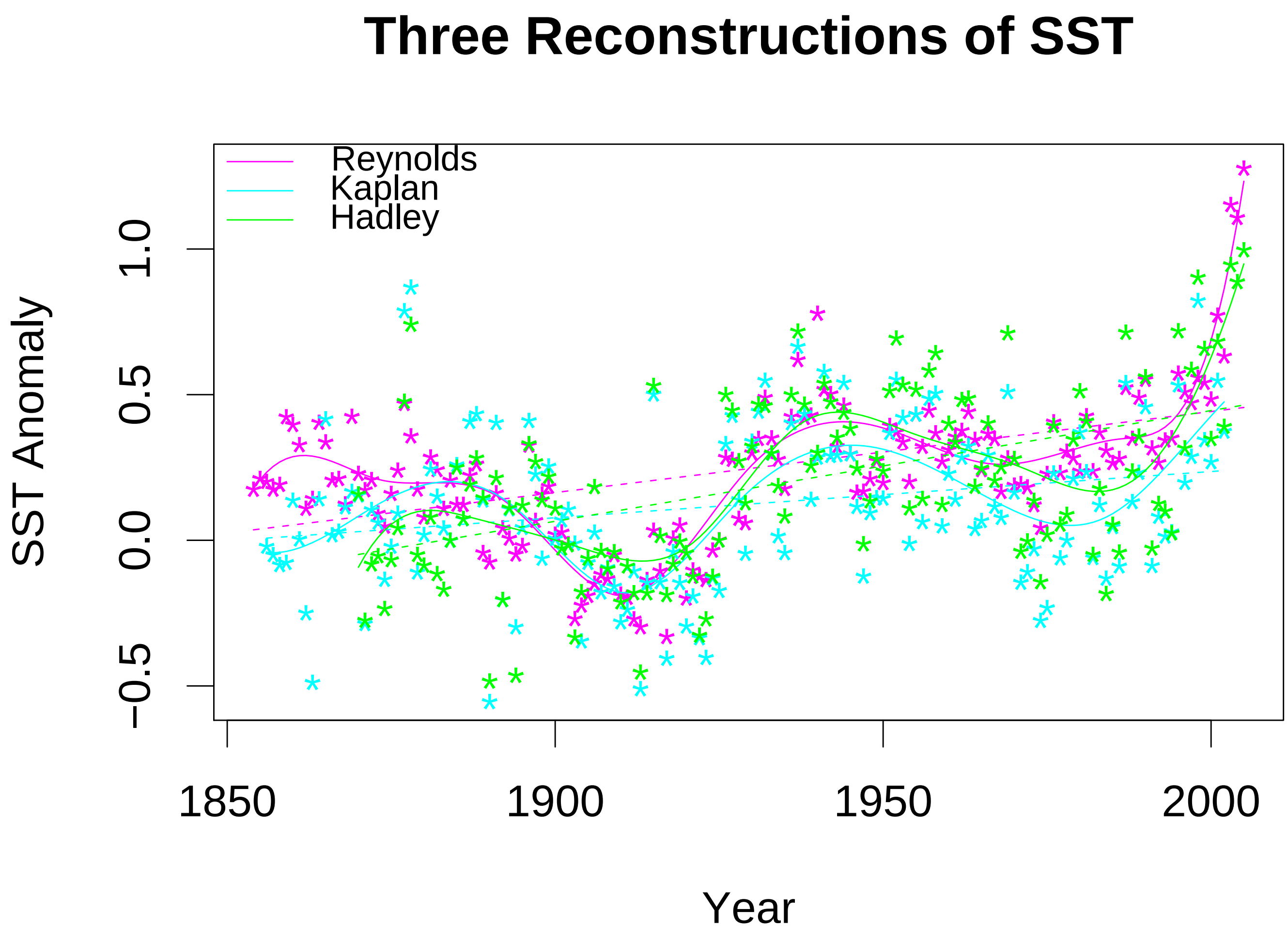
- One-encounter: Assumes a single encounter between a modern storm and a historical ship track is sufficient to count the storm as one that would have been observed historically

- Two-encounter: Similar, but requires two ship $\times$ storm encounters

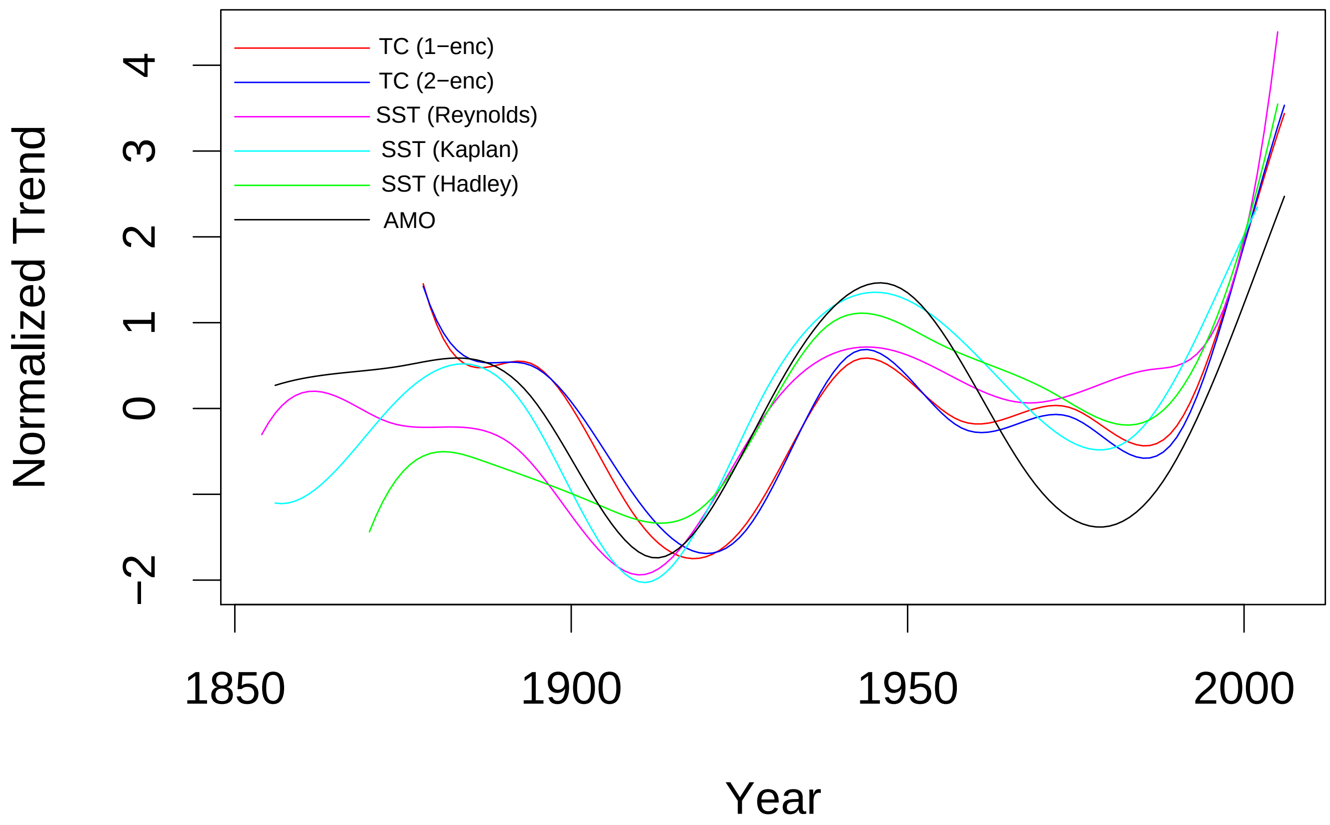
The "two-encounter" model applies a stronger correction to the historical record

Tropical Cyclones by Year
(Data by Vecchi and Knutson)



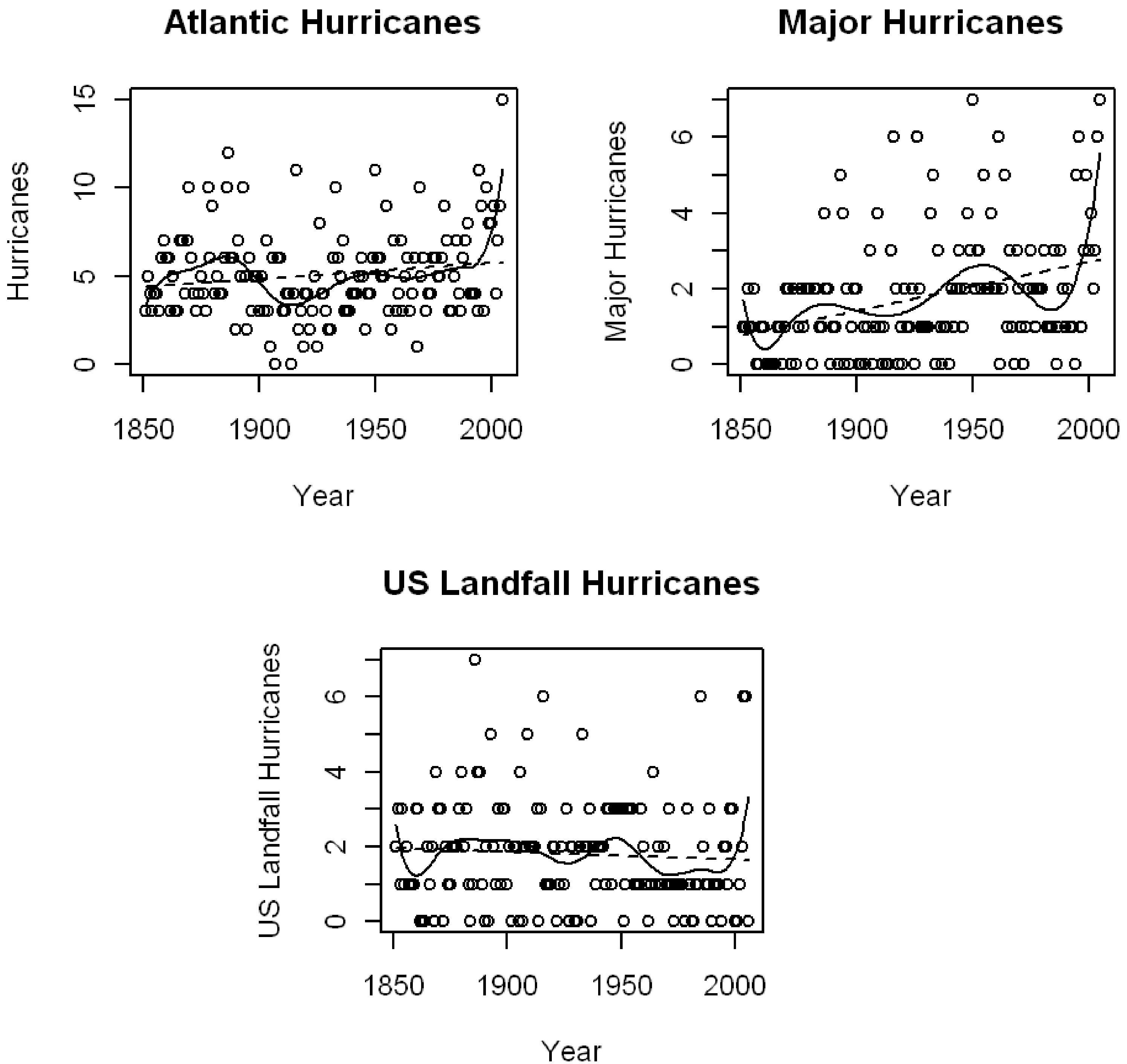

Three Reconstructions of SST
Reynolds
Kaplan
Hadley
SST Anomaly
−0.5
0.0
0.5
1.0
1850
1900
1950
2000
Year

Trends of TC Counts and SSTs



Trends in hurricane counts are similar to those in TCs but the data are much sparser

US landfalling hurricanes show no trend or a slight decrease but this can be explained as a sampling effect



Linear and Nonlinear Trends in Three Series

STATISTICAL ANALYSIS

- The basic idea is to decompose the trend in tropical cyclones into a long-term (possibly anthropogenic) trend and a component due to SSTs (which includes the AMO)

- Similar analyses with temperature, precipitation etc. have been done many times using a method called *detection and attribution analysis*, which uses climate models to generate the trend

- However in the absence of a reliable climate model signal for tropical cyclones, researchers have used a linear trend as an approximation.

- As a technical point, we use square roots of tropical cyclone countsa to approximate a Gaussian time series. A key issue is selection of the order of the time series.

Trends Fitted to Tropical Cyclones (No SST Component)

Period	ARMA	Trend	SE	Trend/SE	p-value
1878–2006	(0,0)	0.018	0.0094	1.94	0.055
1878–2006	(9,2)	0.022	0.022	0.97	0.33
1900–2006	(0,0)	0.049	0.012	4.11	8×10^{-5}
1900–2006	(9,4)	0.050	0.020	2.54	0.011

BIVARIATE TIME SERIES MODEL

Define

x_t: SST in year t

y_t: square root TC in year t

Also let T_t denote some modeled trend (initially linear, but later we consider alternatives)

Model:

$$
\begin{aligned}
x_t &= \alpha_0 + \alpha_1 T_t + u_t \ (u_t \text{ ARMA}) \\
\hat{u}_t &= x_t - \hat{\alpha}_0 - \hat{\alpha}_1 T_t, \ \text{(residuals)} \\
y_t &= \beta_0 + \beta_1 T_t + \beta_2 \hat{u}_t + \beta_3 \hat{u}_{t-1} + \beta_4 \hat{u}_{t-2} + v_t \ (v_t \text{ ARMA})
\end{aligned}
$$

Results for T_t linear

Use TC1 dataset ("1-encounter"), Hadley SST
ARMA(1,1) for u_t
ARMA(7,2) for v_t

Focus on coefficient of trend in y_t (β_1 parameter), various start years ending in 2005

Start Year	$\hat{\beta}_1$	SE	t ratio	p value
1880	0.33	0.26	1.29	0.20
1890	0.54	0.24	2.30	0.02
1900	0.75	0.09	8.07	0.00
1910	0.97	0.15	6.48	0.00
1920	0.81	0.13	6.38	0.00
1930	0.65	0.09	7.46	0.00
1940	1.01	0.56	1.80	0.07
1950	1.29	0.45	2.86	0.00
1960	2.00	0.72	2.77	0.01
1970	1.85	0.97	1.91	0.06
1980	0.46	0.65	0.71	0.47

Conclusions from this analysis

The results improve on fitting linear trends without involving SST, and several of the results are highly statistically significant.

However not all the results are statistically significant — in particular, those beginning in 1880 and 1940 still show the effect of AMO.

Further work is still needed!

Overall Conclusions For This Session

- Although IPCC appears to have made up its mind about the anthropogenic causes of global warming, there remain many other climate change questions for which there is not currently a consensus.

- I believe statistical methods and statisticians will have an important role to play in resolving those issues!