THEORY AND METHODOLOGY FOR CLIMATE EXTREMES

Richard L. Smith
University of North Carolina, USA, and

Statistical and Applied Mathematical Sciences Institute

Frontiers in the Detection and Attribution of Climate Change
Banff International Research Station
May 29, 2012

I. Univariate Extremes

II. Multivariate Extremes

III. Spatial Extremes

IV. D&A for Extreme Events

I. Univariate Extremes

- Generalized Extreme Value distribution (GEV) for block maxima:

$$F(x) \;=\; \exp\left\{-\left(1+\xi\frac{x-\mu}{\sigma}\right)_+^{-1/\xi}\right\}$$

- Generalized Pareto distribution (GPD) for threshold exceedances:

$$F(x) \;=\; 1-\left(1+\xi\frac{x-u}{\tau}\right)_+^{-1/\xi}, \quad x \geq u,$$

- GEV for threshold exceedances (point process approach): Apply GEV censored at $x \geq u$ for threshold exceedance data

Application to climate change:

Apply GEV with time-changing parameters, e.g.

$$\begin{aligned}\mu_t &= \beta_0 + \beta_1 t, \\ \sigma_t &= \sigma, \\ \xi_t &= \xi.\end{aligned}$$

This is a simple model. Some more complicated ones are

- More general forms for μ_t, e.g. B-splines or natural splines to represent a nonlinear trend

- μ_t (and maybe also σ_t, ξ_t) derived from a climate model to reflect more general trends and future projections

- If this approach is applied separately to models with anthropogenic and natural forcings, we are already on the way to a D&A for extremes

II. Multivariate Extremes

Extend to $d > 1$ dimensions, simplest case $d = 2$, e.g.

- temperature and precipitation
- temperature at d different locations
- temperature on d consecutive days (heatwaves)

Multivariate generalization of GEV: if X_{in} is value of variable i at time point n, look for

$$\lim_{n \to \infty} \Pr \left\{ \frac{\max(X_{i1}, ..., X_{in}) - b_{in}}{a_{in}} \leq x_i, \;\; i = 1, ..., d \right\} \;\; = \;\; G(x_1, ..., x_d).$$

If G exists, it is called a *multivariate extreme value distribution* (MEVD) — also known as *max-stable*

Threshold versions exist (Coles and Tawn 1991, Joe, Smith & Weissman 1992, and a gazillion papers since)

G may be specified parametrically or nonparametrically

- Nonparametric estimation derives from spectral characterizations of G (Pickands, Deheuvels, De Haan-Resnick, etc.)

- Many parametric forms but none universally accepted (unlike, say, the role of the multivariate normal distribution in multivariate analysis)

- But maybe this isn't even the right formulation!

Ledford and Tawn (1996, 1997):

- Consider $d = 2$, standardize margins so that $\Pr\{X_i \le x\} = e^{-1/x}$, $i = 1, 2$. So $\Pr\{X_i > x\} \sim x^{-1}$ as $x \to \infty$.

- Under weak conditions one can show that $\Pr\{\max(X_1, X_2) > x\} \sim x^{-1/\eta}$ as $x \to \infty$, where $\eta \in (0, 1]$.

- Case $\eta = \frac{1}{2}$ corresponds to (exact) independence, $\eta = 1$ to asymptotic dependence.

- All MEVDs have $\eta = 1$!

- But many cases are known for $\eta \in (0, 1)$ which we call *asymptotic independence*. For instance, the bivariate normal has $\eta = \frac{1+\rho}{2}$.

The bottom line: we need something more general than MEVDs.

How does one develop statistics for these more general models?

- Ledford and Tawn (1997) proposed a model for $d = 2$ with asymptotic independence
- Alternative version due to Ramos and Ledford (2009)
- Resnick (2002), Maulik and Resnick (2005) developed a more general mathematical framework called *hidden regular variation*, which includes the Ledford-Tawn theory but also shows how it can be generalized to $d > 2$
- Statistical approaches for $d > 2$ so far very limited (ongoing PhD work of G. Weller, CSU)
- There is yet another approach (Heffernan and Tawn, 2004) but this is relatively little utilized so far
- Nobody understands multivariate extremes! (anon.)

The structure variable approach

Maybe we shouldn't worry about multiple extremes. Instead, identify a *single* variable that best characterizes the impact of interest, and do univariate extreme value analysis on that. Coles and Tawn (1994) called this the *structure variable approach*

Potential reasons for preferring multivariate analysis:

- Better fit of the extreme value model (maybe)
- Opportunity to use longer records of single variables than may be available for joint variables
- Better understanding of what is going on in terms of climate change
- Ability to study several impacts simultaneously

III. Spatial Extremes

- Interest in the distribution of extremes at many spatial locations — say, $X(s)$ is block maximum at a location $s \in \mathcal{S}$ for some set of spatial locations $\mathcal{S}$.

- Concept of a *max-stable process*: for any $s_1, ..., s_d$, joint distribution of $X(s_1), ..., X(s_d)$ is max-stable

- Characterizations due to De Haan (1984), Giné, Hahn and Vatan (1990). First statistical exploration in an unpublished paper of Smith (1990)

- Schlather (2002) developed a whole new class of max-stable processes, further extended by Kabluchko *et al.* (2009) (*Brown-Resnick processes*)

- Statistics still difficult: we can calculate closed-form densities for $d = 2$, but in most cases not for any $d > 2$. However the *method of composite likelihood* (Padoan, Ribatet and Sisson, 2010) offers a possible way of estimating these models without full joint densities

- Current research includes threshold methods (Huser and Davison, Jeon and Smith), Bayesian equivalents of the MCLE (Ribatet, Cooley and Davison; Shaby), nonstationary processes (e.g. Blanchet and Davison, 2011) and many other things

- But one question is completely open: how to develop spatial process equivalents of the asymptotically independent theory of multivariate extremes

IV. D&A for Extreme Events

- Observe some extreme weather event
- Run a large number of climate models under anthropogenic forcings; measure weather variable corresponding to the observed extreme event
- Repeat but under either natural forcings or using control model runs
- Estimate P_1: probability of extreme event under anthropogenic scenario and P_0: probability of extreme event under natural or control scenario
- The *fraction of attributable risk* is

$$FAR = 1 - \frac{P_0}{P_1}$$

or just consider the *risk ratio* $\frac{P_1}{P_0}$.

Approaches

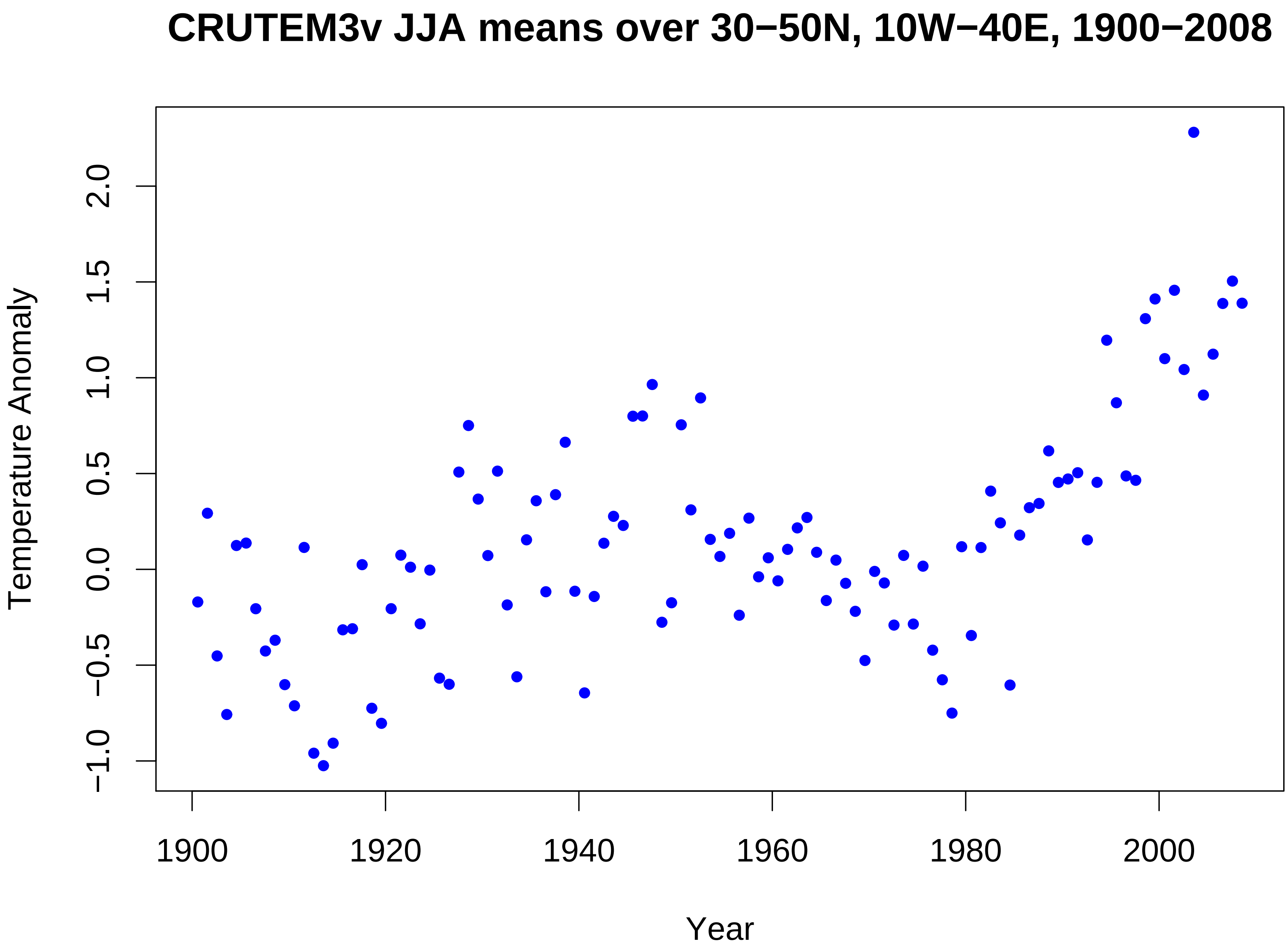
- Stott, Stone and Allen (2004): first major development of methodology, illustrated on 2003 European heatwave

- Statistical approaches based on normal distributions have been proposed (e.g. Schär *et al.* 2004, Jaeger *et al.* 2008) but studying extreme value problems generally requires statistical methods designed for that purpose

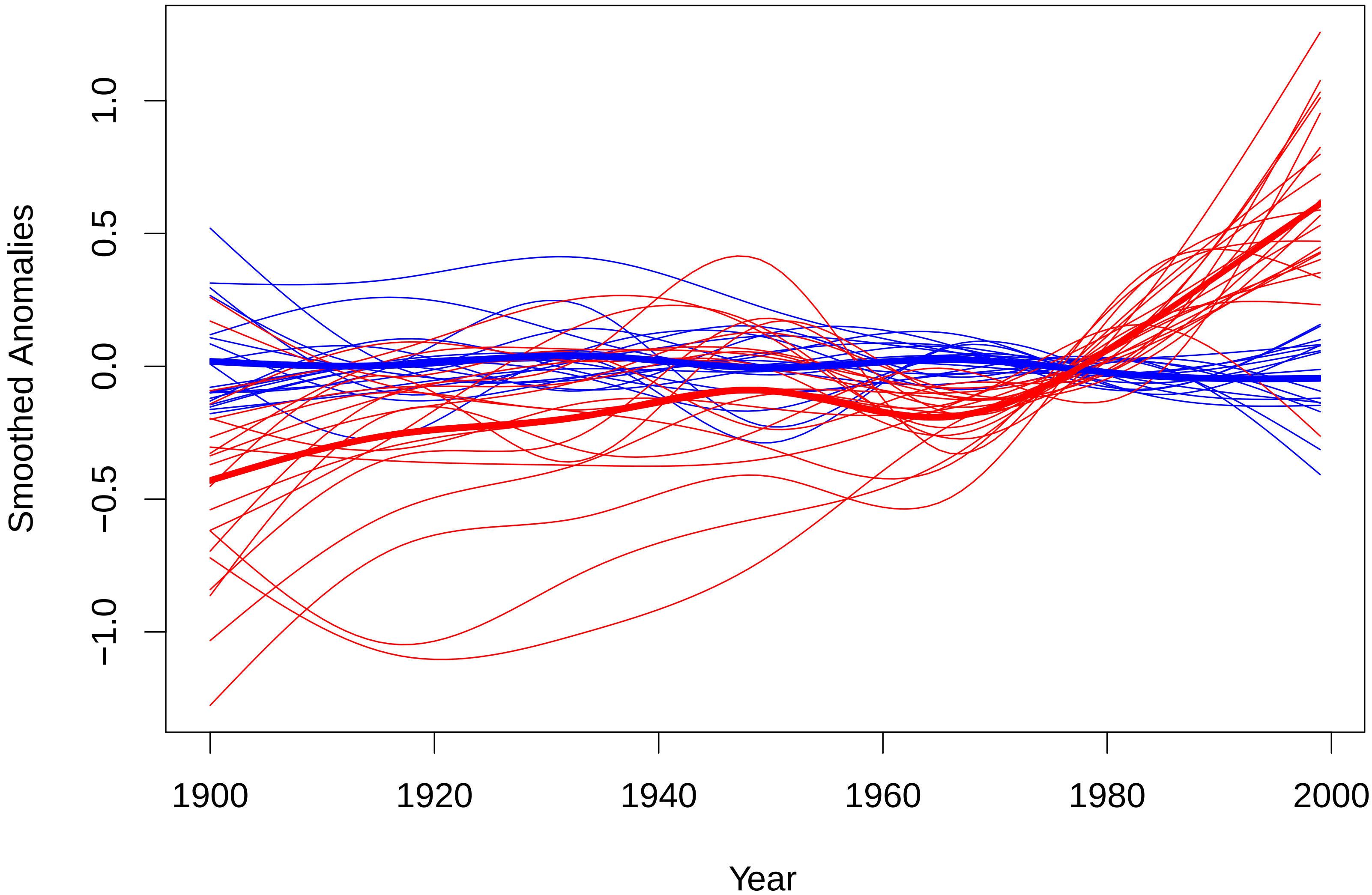
- Another method developed by NOAA group (Dole, Hoerling, Perlwitz, etc.), including paper by Dole *et al.* (2011) disputing anthropogenic role in 2010 Russian heatwave

- A very data-intensive approach due to Pall *et al.* (2011)

- Here I follow an approach of Smith and Wehner which is based on trying to generalize the method of Stott, Stone and Allen



Observed anomalies for JJA mean temperatures for 1900–2008



Smoothed annual temperature means based on 95 control model runs (blue) and 57 twentieth century model runs (red).

Analysis of Observational Data

- Fit GEV model to exceedances over a high threshold, 1900–2000
- Set $\mu_t = \sum_{k=1}^{K} \beta_k x_{kt}$ where the $\{x_{kt}\}$ are a set of natural spline basis coefficients, σ_t and ξ_t constant
- Use the fitted model to compute the probability of exceeding 2.3K in a single year at the end of the series
- Critical questions include the selection of a threshold (characterized as a fixed quantile of the observations) and of K, which determines the smoothness of the fitted trend. The Akaike Information Criterion (AIC) is one tool (among many) that could assist in the selection of K, but it doesn't help us choose the threshold

	85% Threshold	80% Threshold	75% Threshold
DF	AIC	AIC	AIC
1	75.6	107.8	110.4
2	71.3	105.2	98.2
3	43.6	74.2	68.6
4	43.9	74.1	71.6
5	37.1	74.0	71.2
6	43.3	82.9	70.3
7	40.6	83.3	71.3
8	42.8	85.4	71.2
9	39.6	87.3	68.1
10	46.3	81.2	76.5

Table 1. Fit to observational data: AIC values for DF=1,...,10 and three threshold levels

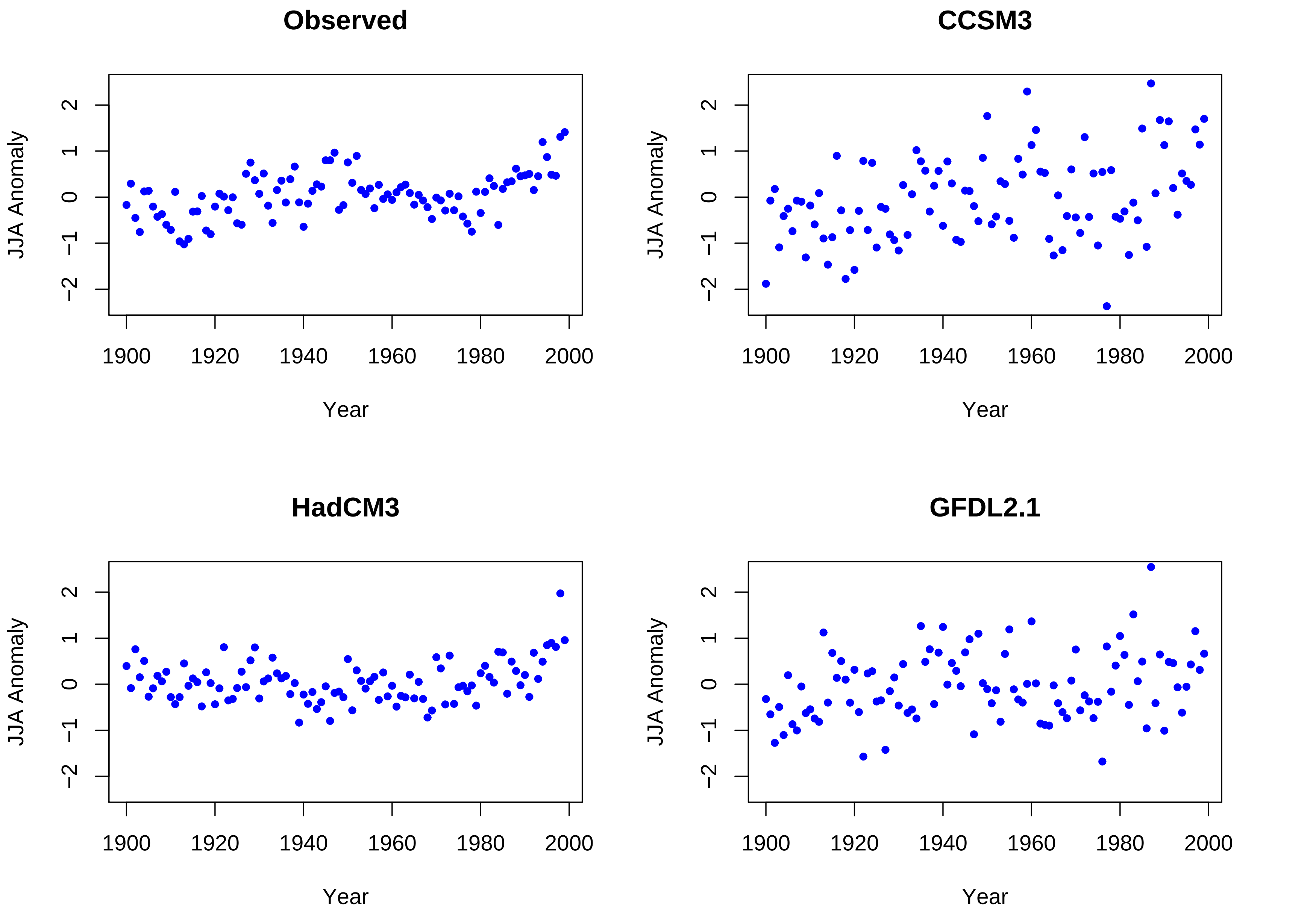
DF	Threshold (%)	MLE	Posterior Mean	Posterior Quantiles		
				.025	.5	.975
4	80%	0	.011	0	.007	.042
4	75%	.001	.013	.0003	.008	.048
4	85%	.0007	.020	.001	.015	.068
6	80%	.00002	.027	.002	.019	.090
6	75%	.0004	.025	.002	.019	.080
6	85%	.0011	.032	.002	.024	.110
8	80%	.0001	.035	.004	.028	.104
8	75%	.0008	.028	.003	.022	.082
8	85%	.0017	.052	.005	.040	.167

Table 2. Summaries of estimates for the probability of exceedance of 2.3K, for different DF and threshold. Shown are the maximum likelihood estimate (MLE), the posterior mean, and the .05, .5 and .975 quantiles of the posterior distribution. These are estimated by a Metropolis sampler with 100,000 iterations, with a flat prior distribution.

How can we extend this to a D&A analysis?

The obvious approach is to run two extreme value analyses, one using the anthropogenic forcings climate models and the other using the natural forcings climate models, then compare the results.

The problem is that the climate models are not all consistent with each other or with the data. I call this the *scale mismatch problem*.



The scale mismatch problem: Observed series and three 20C model runs, up to 2000

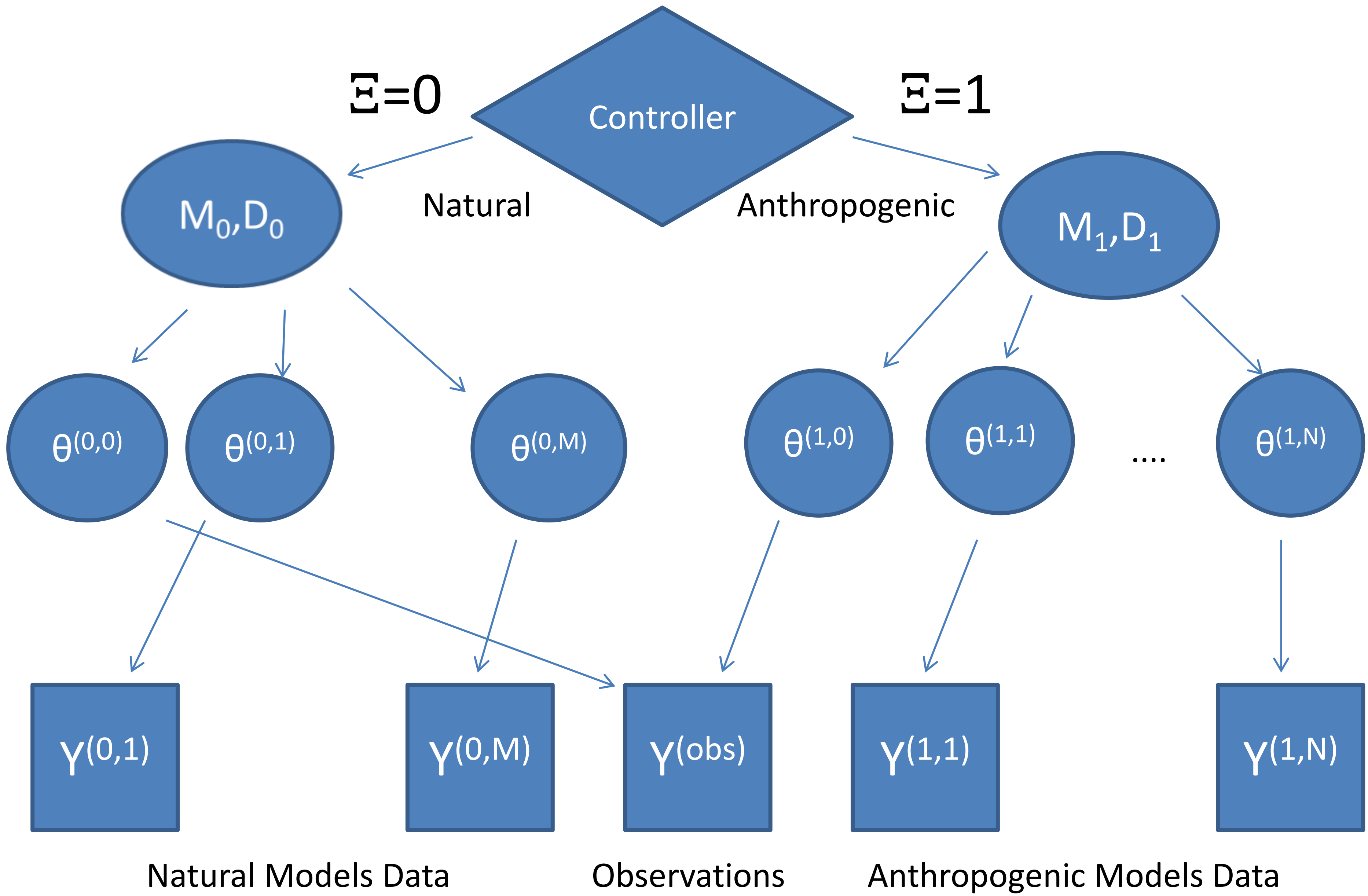
Our approach:

- Fit univariate GEV distributions with trends represented by natural splines to both observational data and to each of the climate models under both natural and anthropogenic scenarios

- However, instead of assuming that the observational and climate model data are all generated by the same statistical distributions, we create two *hierarchical models* (one corresponding to natural forcing models, the other to anthropogenic forcing)

- Model fit by Bayesian statistics — uses MCMC, computationally intensive for a statistical estimation method but not by comparison with climate models!

Hierarchical Model

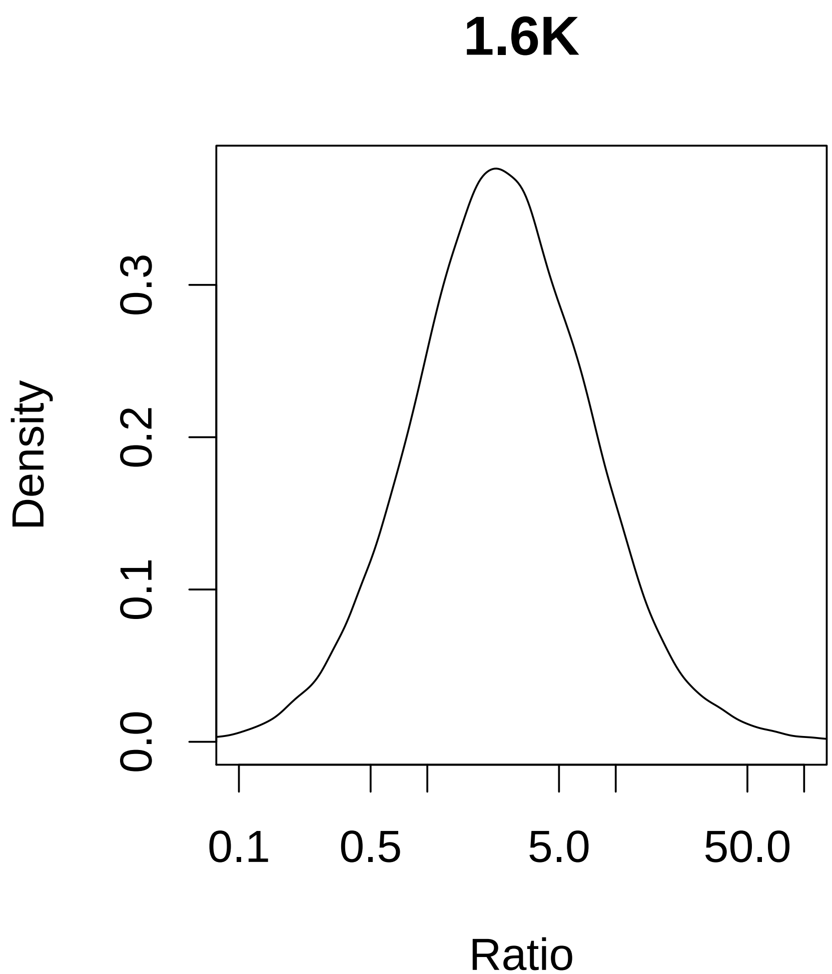


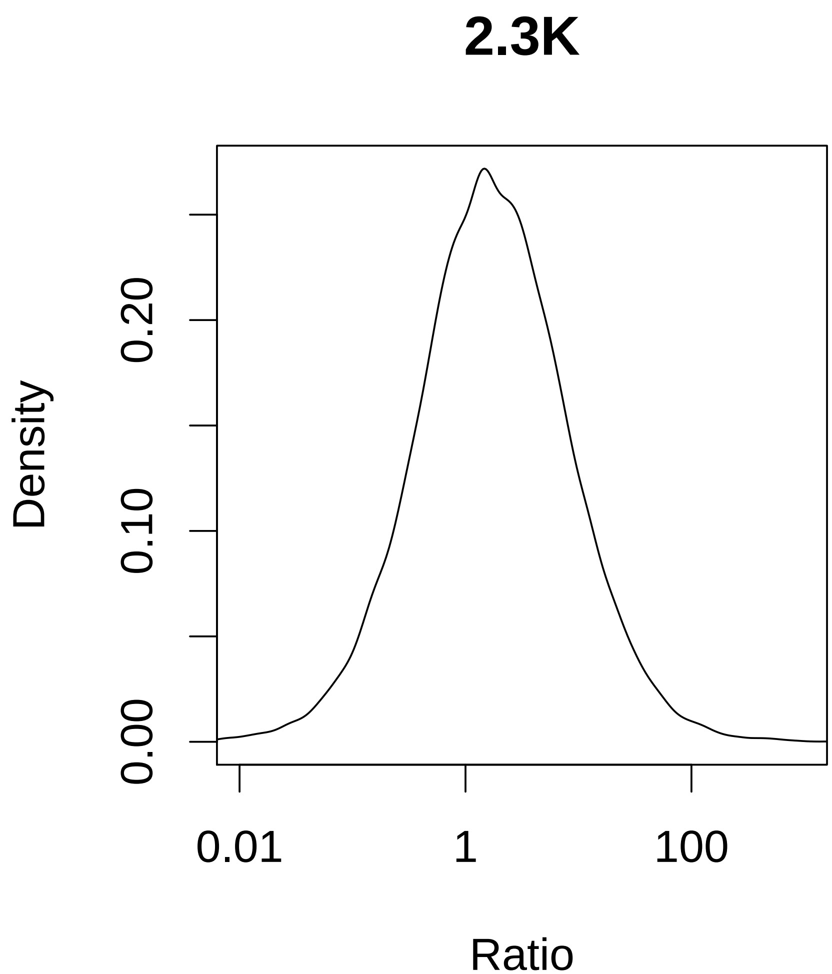
Posterior Probability Ratios

Run	Exceedances of 2.3K				Exceedances of 1.6K			
	Posterior	Posterior Quantiles			Posterior	Posterior Quantiles		
	GMean	.25	.5	.75	GMean	.25	.5	.75
1	1.73	0.64	1.73	4.59	2.50	1.21	2.47	5.03
2	1.70	0.63	1.68	4.49	2.53	1.23	2.47	5.10
3	1.63	0.59	1.65	4.51	2.41	1.18	2.41	4.94
4	1.69	0.63	1.66	4.55	2.43	1.18	2.40	4.90
5	1.87	0.66	1.85	5.24	2.93	1.43	2.91	5.96
6	1.86	0.65	1.88	5.37	2.06	0.95	2.04	4.46
7	1.61	0.58	1.63	4.43	1.80	0.83	1.79	3.85
8	2.10	0.74	2.09	6.01	2.82	1.30	2.79	6.00
9	1.81	0.62	1.85	5.33	1.90	0.85	1.87	4.21

Table 3. Summaries of posterior geometric means and posterior quantiles for the ratios of post-event probabilities of exceeding a high level under anthropogenic and natural forcing.

POSTERIOR DENSITIES OF RATIOS





However there are still some objections to this.

- Exchangability among the climate models and observations not necessarily appropriate

- Analysis may give too much weight to the observations over the climate models — possible that there will still be a positive estimated trend for the control run scenario even though we know the climate models contain no trend in this case (an objection voiced by Chris Paciorek)

- However one can also make the opposite objection — real data typically have more variability that climate models and the statistical analysis should take account of that (a comment made to me by Tony O'Hagan)

- As a way of extending the analysis I propose a *variance expanding approach* — the conditional variances for the parameters in the observational data are assumed multiplied by a factor R compared with those for climate models

Posterior probability ratios with varying weights on the climate models

R	Exceedances of 2.3K				Exceedances of 1.6K			
	Posterior GMean	Posterior Quantiles .25	.5	.75	Posterior GMean	Posterior Quantiles .25	.5	.75
0.25	2.10	0.73	2.11	5.93	1.67	0.87	1.64	3.14
0.5	1.86	0.67	1.83	5.11	2.11	1.06	2.07	4.12
1	1.51	0.57	1.49	3.98	2.67	1.28	2.63	5.48
2	1.68	0.67	1.63	4.19	3.40	1.64	3.29	6.79
4	1.72	0.62	1.71	4.65	4.22	2.03	4.06	8.35

Table 4. Summaries of posterior geometric means and posterior quantiles for the ratios of post-event probabilities of exceeding a high level under anthropogenic and natural forcing. Based on DF=4; threshold=80%; varying R parameter

CONCLUSIONS

- Original estimates by Stott, Stone and Allen suggested a risk ratio (anthropogenic to natural) of about 4:1, wtih relatively narrow confidence bounds

- I believe their result was too optimistic, especially regarding the width of the confidence bands (but the $R = 4$ result for 1.6K essentially agrees with theirs)

- Nevertheless, the new analysis still confirms a risk ratio of about 2:1 in many analyses, and this is robust to a number of model selection aspects

- The bottom line — *yes, there is a human influence on extreme events as well as on means!*

- Next steps: analyze Russian heatwave 2010, Texas heatwave 2011, and other extreme events