CLIMATE EXTREMES: ATTRIBUTIONS AND FUTURE PROJECTIONS

Richard L Smith

University of North Carolina and SAMSI

(Joint with Michael Wehner, Lawrence Berkeley Lab)

Ninth Conference on Extreme Value Analysis

Ann Arbor, June 19, 2015

THE UNIVERSITY *of* NORTH CAROLINA *at* CHAPEL HILL

samsi

NSF • Duke • NCSU • UNC • NISS

BERKELEY LAB

Bringing Science Solutions to the World

Raleigh News and Observer, this week

Weather JUNE 16, 2015

Heat sets record in the Triangle; bus stops bake

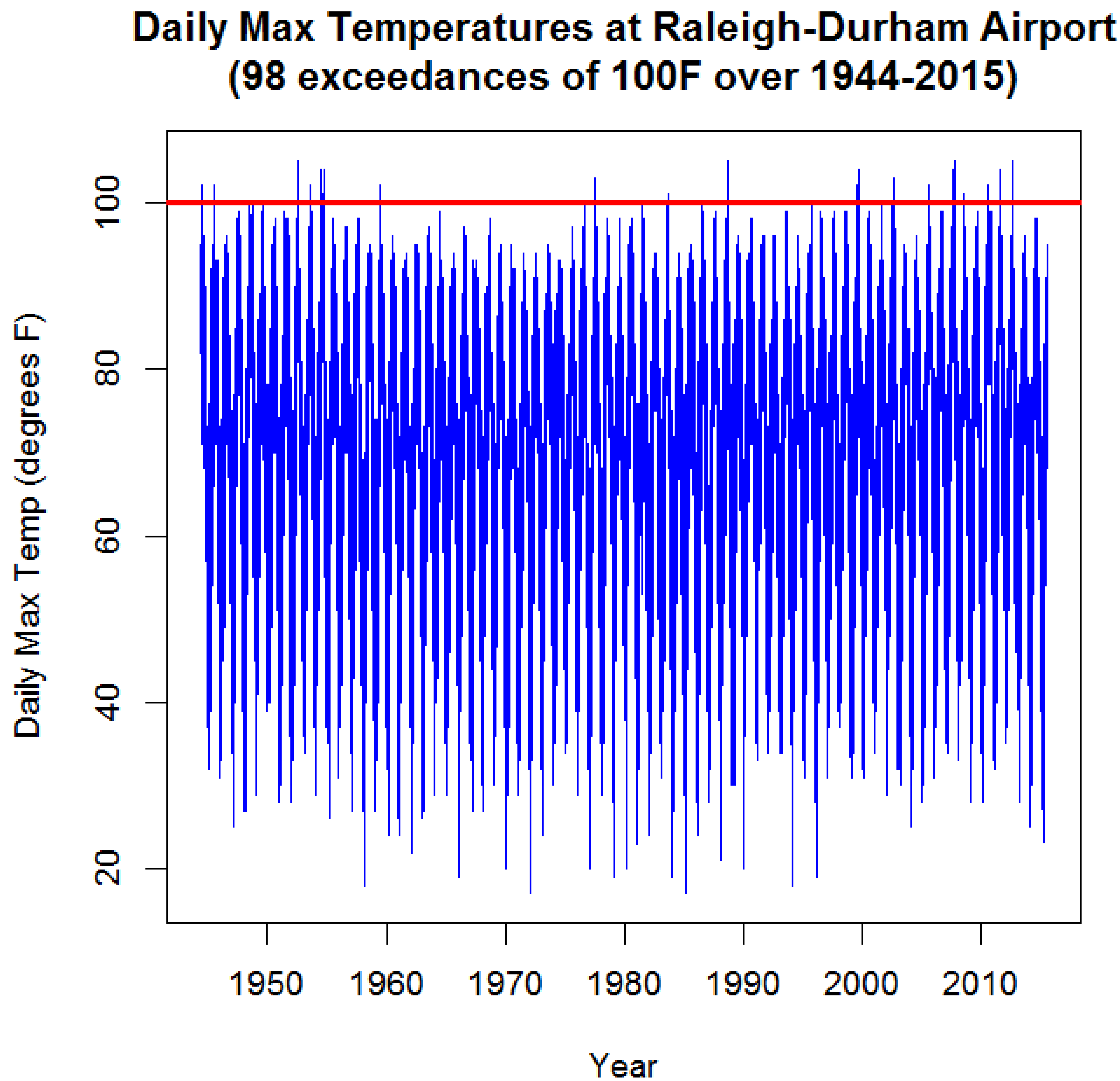
One year old Elena Gonzales beats the oppressive heat a little differently than her family as she rests comfortably in a hammock in the shade along the Eno River on Tuesday. The West Point on the Eno park in Durham was quite the popular place as record breaking temperatures drew families to enjoy the cool waters. **Chuck Liddy** - cliddy@newsobserver.com

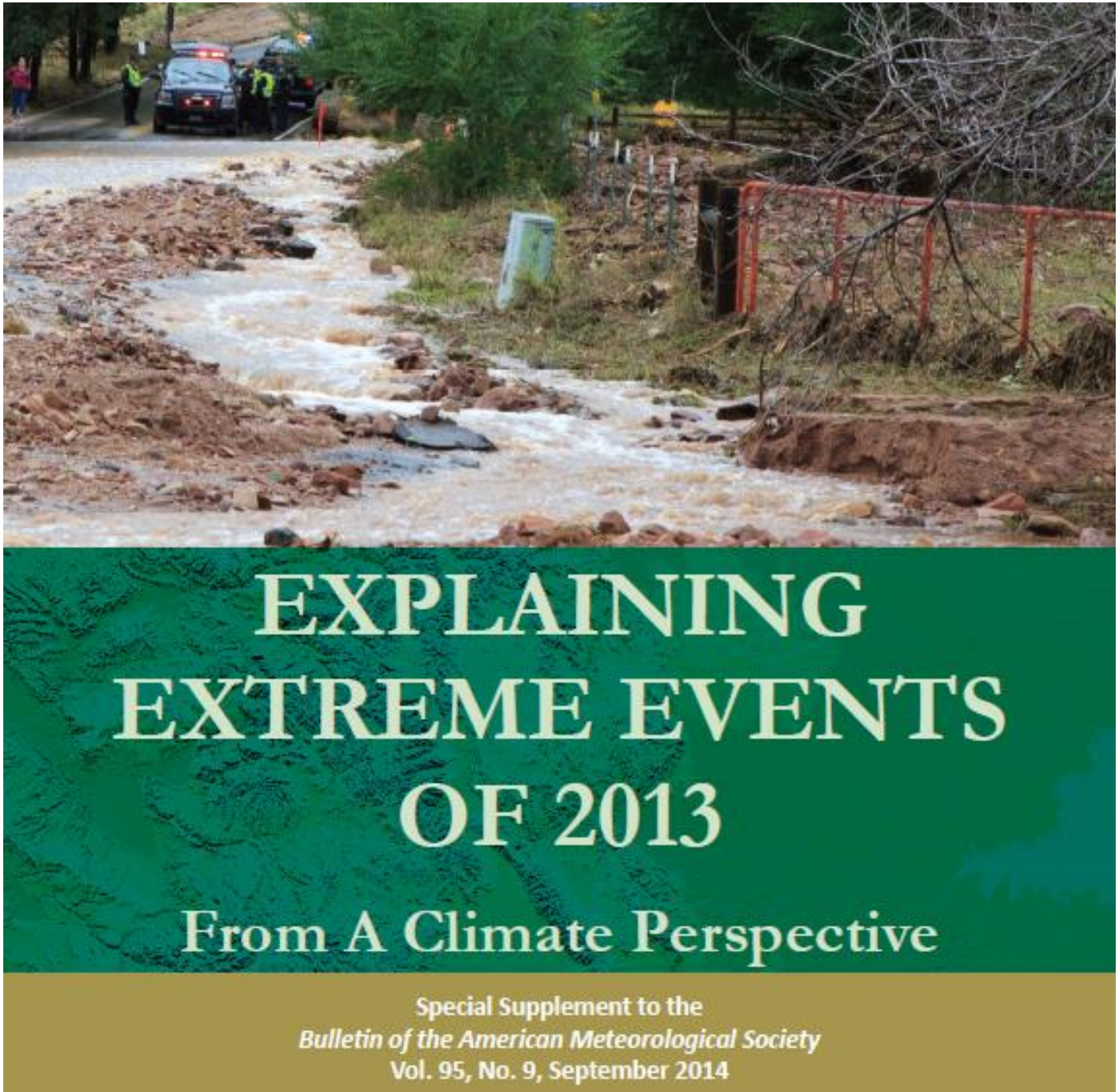
Single-day record for June 16: 100°F in 2015

(and for June 15: 99°F)

about 3 hours ago

Dangers of leaving children unat


Daily Max Temperatures at Raleigh-Durham Airport
(98 exceedances of 100F over 1944-2015)
Daily Max Temp (degrees F)
20
40
60
80
100
1950
1960
1970
1980
1990
2000
2010
Year


EXPLAINING
EXTREME EVENTS
OF 2013
From A Climate Perspective
Special Supplement to the
Bulletin of the American Meteorological Society
Vol. 95, No. 9, September 2014

LETTERS
PUBLISHED ONLINE: 27 APRIL 2015 | DOI: 10.1038/NCLIMATE2617

nature climate change

Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes

E. M. Fischer* and R. Knutti

Climate change includes not only changes in mean climate but also in weather extremes. For a few prominent heatwaves and heavy precipitation events a human contribution to their occurrence has been demonstrated[1-5]. Here we apply a similar framework but estimate what fraction of all globally occurring heavy precipitation and hot extremes is attributable to warming. We show that at the present-day warming of 0.85 °C about 18% of the moderate daily precipitation extremes over land are attributable to the observed temperature increase since pre-industrial times, which in turn primarily results from human influence[6]. For 2 °C of warming the fraction of precipitation extremes attributable to human influence rises to about 40%. Likewise, today about 75% of the moderate daily hot extremes over land are attributable to warming. It is the most rare and extreme events for which the largest fraction is anthropogenic, and that contribution increases nonlinearly with further warming. The approach introduced here is robust owing to its global perspective, less sensitive to model biases than alternative methods and informative for mitigation policy, and thereby complementary to single-event attribution. Combined with information on vulnerability and exposure, it serves as a scientific basis for assessment of global risk from extreme weather, the discussion of mitigation targets, and liability considerations.

changed, and FAR indicates the fraction attributable to humans. 'Fraction of events' throughout the text should be interpreted as an anthropogenic contribution to the probability of such events, rather than some events being anthropogenic and some not. We base our estimates on well-defined percentiles of daily temperature and precipitation derived from long pre-industrial control runs of 25 CMIP5 models (see models in Supplementary Table 1).

In response to increasing global temperatures, models project more heavy precipitation days, as illustrated by histograms aggregating daily precipitation (Fig. 1) across Northern Europe and North America (see Methods). The simulated occurrence of heavy precipitation days under present-day warming of 0.85 °C (blue lines) is only slightly higher than in pre-industrial conditions. At a warming of 2 °C (red lines) the probability of the most extreme cases, exceeding the pre-industrial 99.99%-quantile, increases by about a factor of 1.5 to 3 depending on region and model (lower panels). This implies that on average across the area an event expected once every 10,000 days (about 30 years), in pre-industrial conditions, is expected every 10 to 20 years at a 2 °C warming. The wet tail of the precipitation distribution becomes fatter; thus, the PR increases most rapidly for the most intense and rarest events (Fig. 1) at the expense of days with moderate, low or no precipitation. This is consistent with the finding that in some cases mean precipitation decreases (primarily owing to large-scale

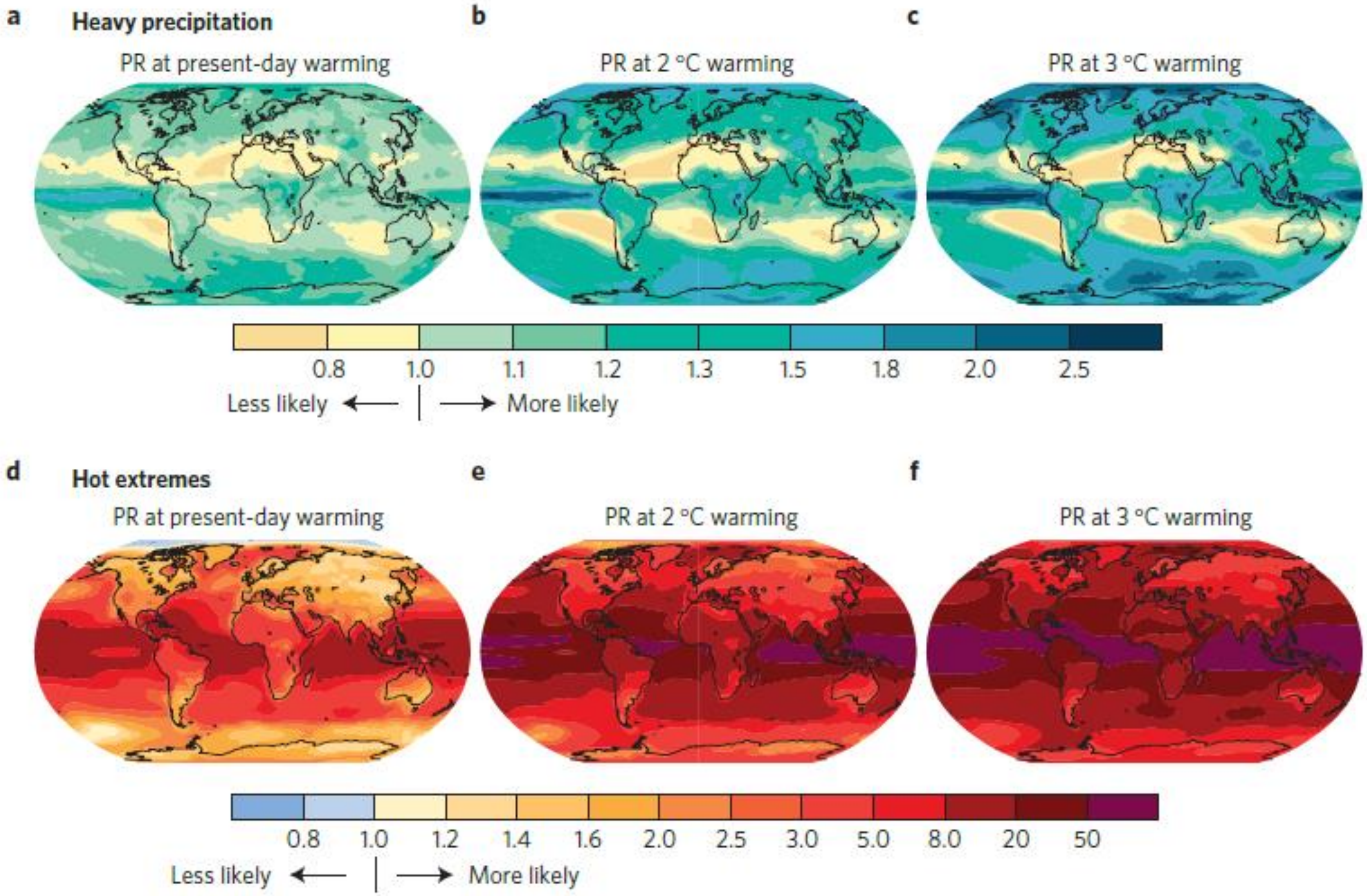


Figure 3 | Change in probability of heavy precipitation and hot extremes. a–f, Multi-model mean probability of exceeding the pre-industrial 99th percentile of daily precipitation (**a**–**c**) and temperature (**d**–**f**), relative to pre-industrial. Ratios are shown for 30-year periods in which the global mean temperatures warmed 0.85 °C (present-day) (**a**,**d**), 2 °C (**b**,**e**) and 3 °C (**c**,**f**) above pre-industrial conditions.

Causes and predictability of the 2011 to 2014 California drought

Richard Seager *

Lamont Doherty Earth Observatory of Columbia University, Palisades, New York

Martin Hoerling

NOAA Earth System Research Laboratory

Siegfried Schubert, Hailan Wang

NASA Goddard Space Flight Center

Bradfield Lyon,

International Research Institute for Climate and Society

Arun Kumar,

NOAA Climate Prediction Center

Jennifer Nakamura, Naomi Henderson

Lamont Doherty Earth Observatory of Columbia University, Palisades, New York

"…the recent drought was dominated by natural variability, a conclusion framed by discussion of differences between observed and modeled tropical SST trends over past decades."

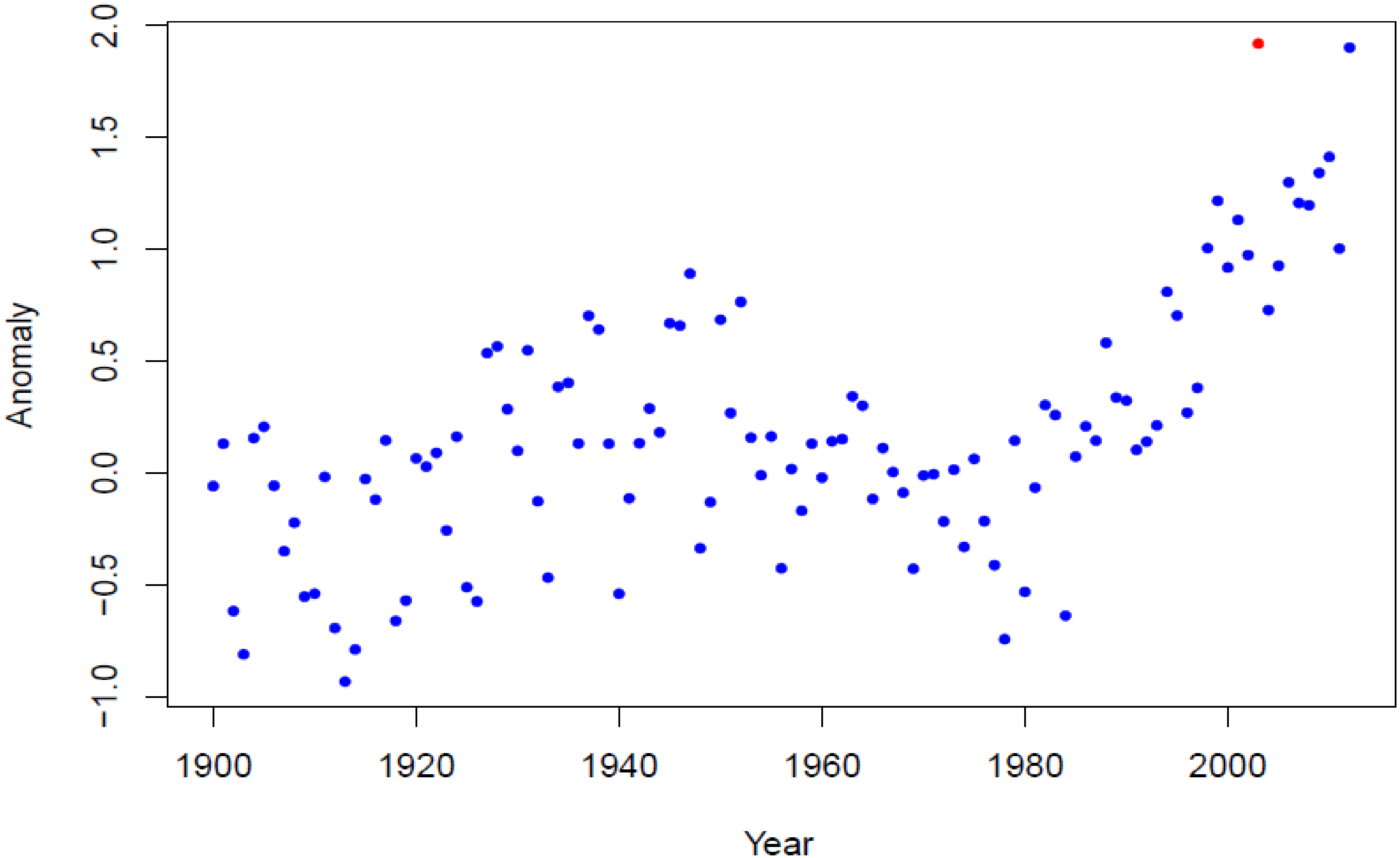
Methods

- Original paper by Stott, Stone and Allen (2004)
  - Defined FAR = 1 – P_0/P_1 where P_0 and P_1 are probabilities of extreme event under natural, anthropogenic scenarios
  - Could also consider risk ratio RR=P_1/P_0
  - Method used a combination of extreme value theory (GPD) and normal-theory detection and attribution techniques
- New Method by Pall et al. (2011)
  - No EVT – counting threshold exceedances in customized climate model
  - Data intensive
- A variety of methods in recent climatology papers
  - Fischer-Knutti 2015: different perspective, not trying to compare observations with models, didn't evaluate uncertainty of estimates
- Other statistical approaches
  - Soyoung Jeon at this meeting
- We also highlight the problem of *projecting future extreme events* (Christidis, Jones, Stott 2014; NRC report on *Climate and Social Stress*, 2012)
- Not much using methods of extreme value theory – objective of this talk is to examine where that could take us

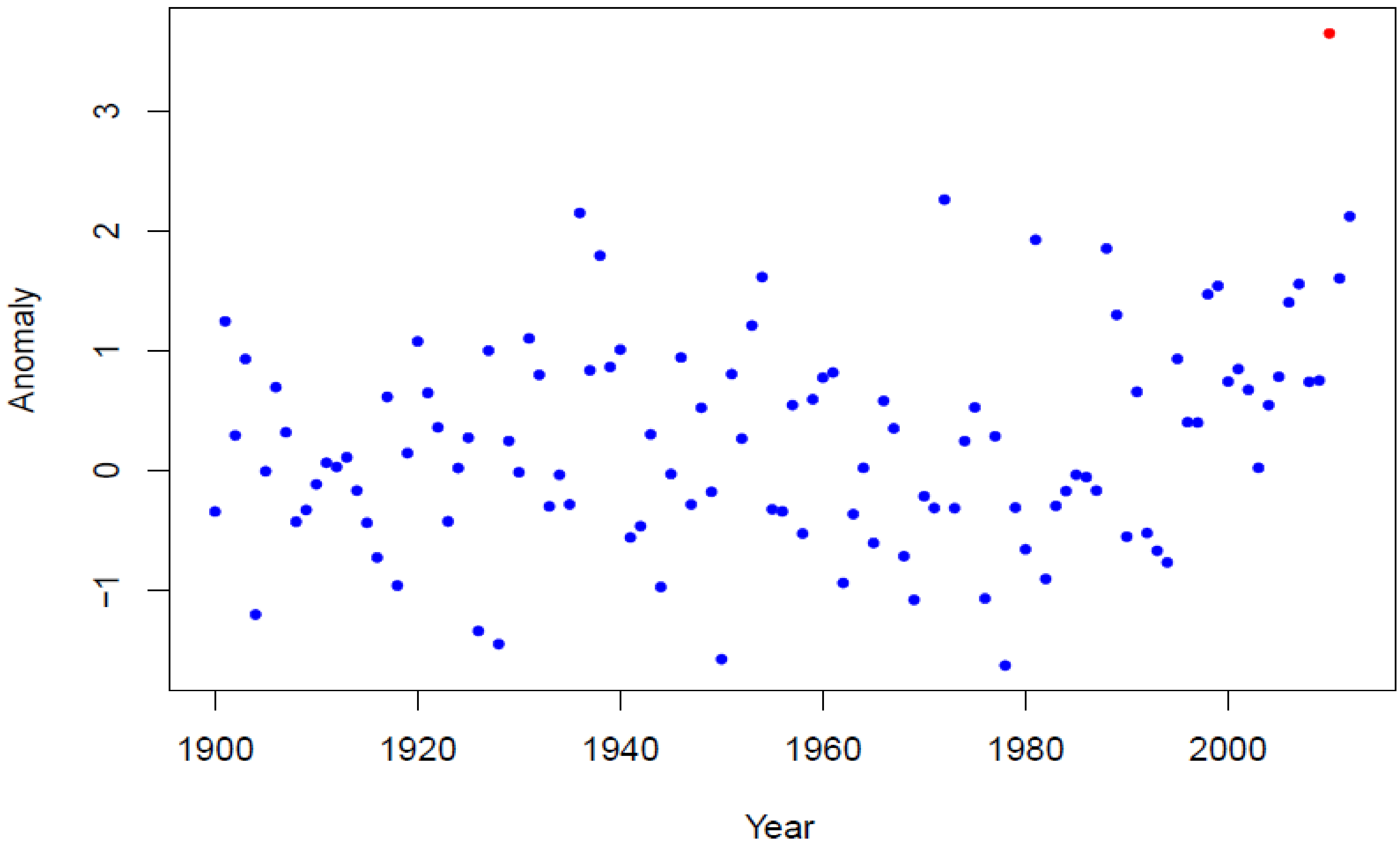
Data

- Observational data to 2012 from CRU (Climate Research Unit, University of East Anglia, UK) – monthly averages on $5^o x 5^o$ grid boxes, aggregated to JJA average anomalies over
  - Europe: spatial averages over 10^oW-40^oE, 30^oN-50^oN (2003 value was 1.92K but 2012 almost the same)
  - Russia: spatial averages over 30^oE-60^oE, 45^oN-65^oN (2010 value 3.65K)
  - Central USA (including Texas and Oklahoma): spatial averages over 90^oW-105^oW, 25^oN-45^oN (2011 value 2.01K)
- Climate model data from CMIP3
  - 14 climate models
  - Total of 64 control runs, 44 twentieth century runs, 34 future projections under A2 scenario
  - Same spatial regions as observational data, converted to anomalies

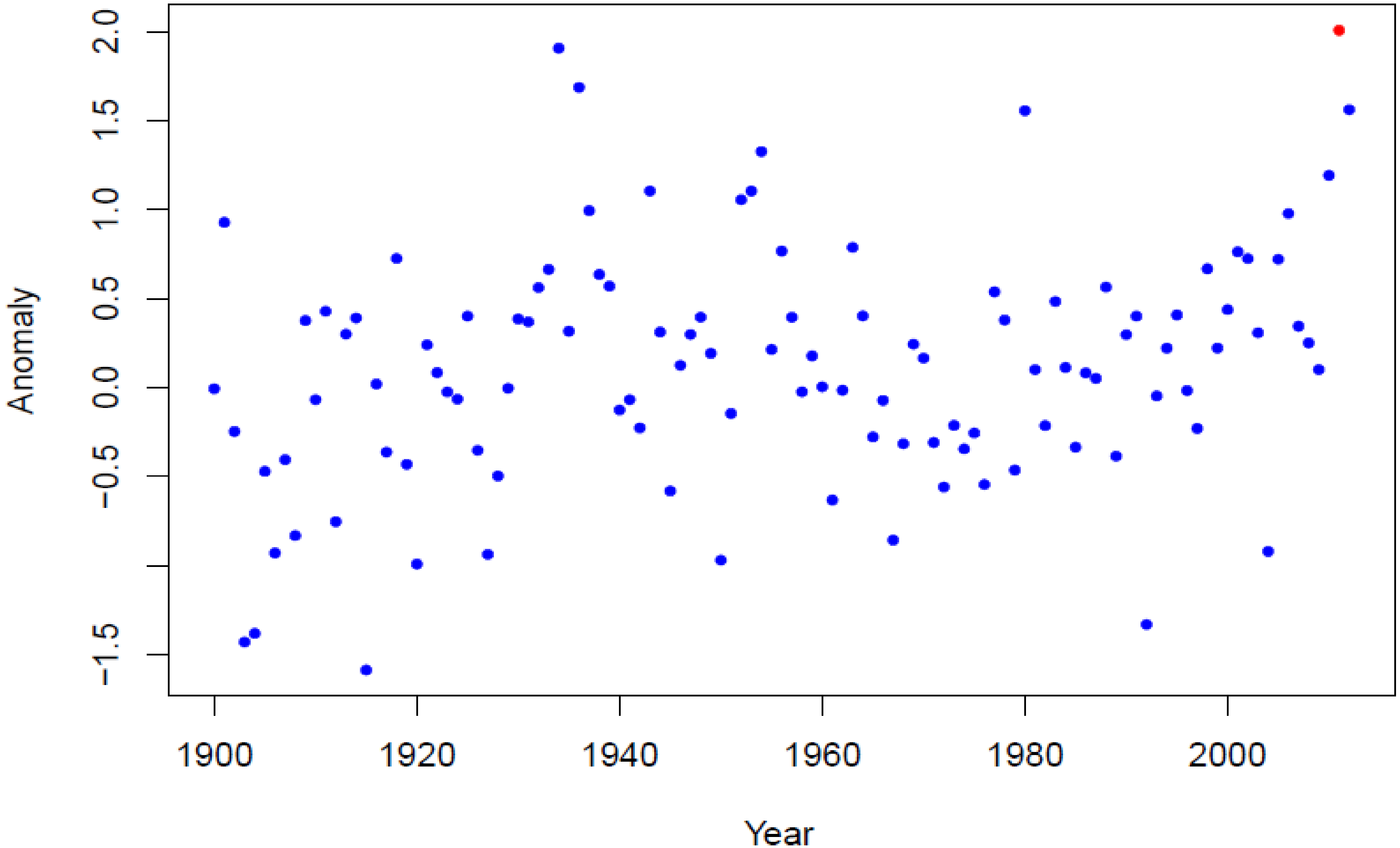
Europe Summer Mean Temperatures



Russia Summer Mean Temperatures



Central USA Summer Mean Temperatures



Approach for a Single Series

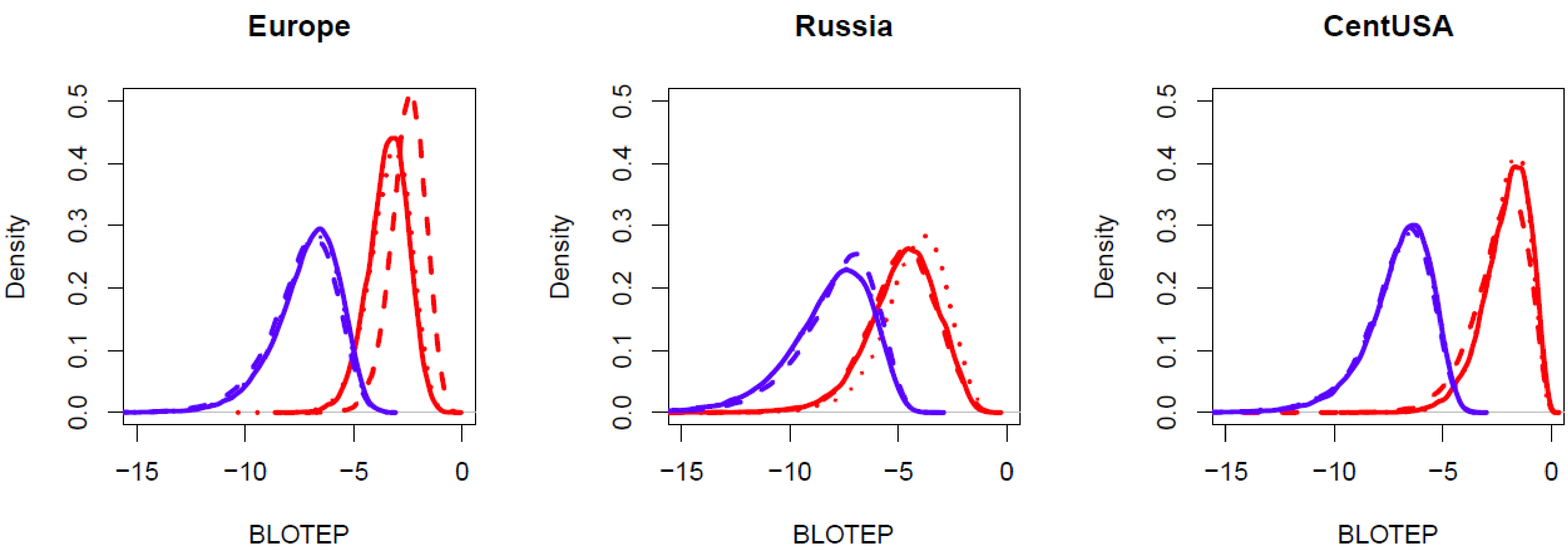
Generalized Extreme Value Distribution (GEV) with covariates to represent trend

$$\begin{aligned}
\Pr\{Y_t \le y\} &= \exp\left[-\left\{1+\xi_t\left(\frac{y-\mu_t}{\sigma_t}\right)\right\}_+^{-1/\xi_t}\right], \\
\mu_t &= \beta_0 + \sum_{k=1}^{K} \beta_k x_{kt}, \\
\sigma_t &= \sigma_0, \\
\xi_t &= \xi_0,
\end{aligned}$$

- *Peaks over threshold* approach: fit GEV to exceedances over threshold u, treat $Y_t < u$ as censored
- u chosen as one of 75th, 80th or 85th percentile
- Covariates $\{x_{kt},\ 1 \le k \le K\}$ chosen to represent spline basis functions
- K chosen by AIC, no formal selection of threshold but run different thresholds for comparison
- *Bayesian predictive analysis: use MCMC to calculate a posterior density for the probability of crossing a given high level in a given year*

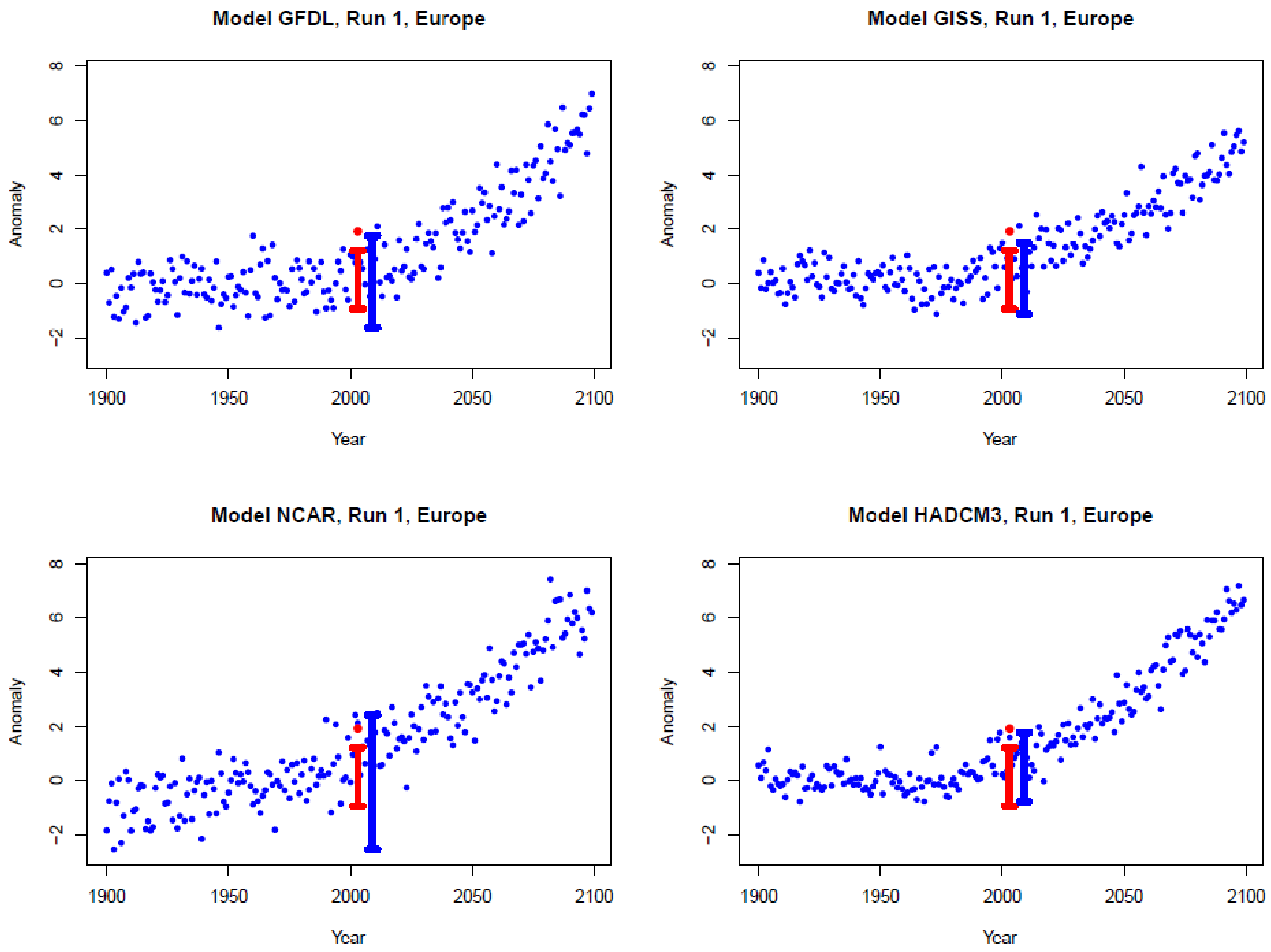
Bayesian Calculations

- Focus on binary log of exceedance probability of threshold (BLOTEP)
- Use models both with (red) and without (blue) trends
- Use 80th (solid curve), 75th (dashed) and 85th (dot-dashed) percentiles for thresholds

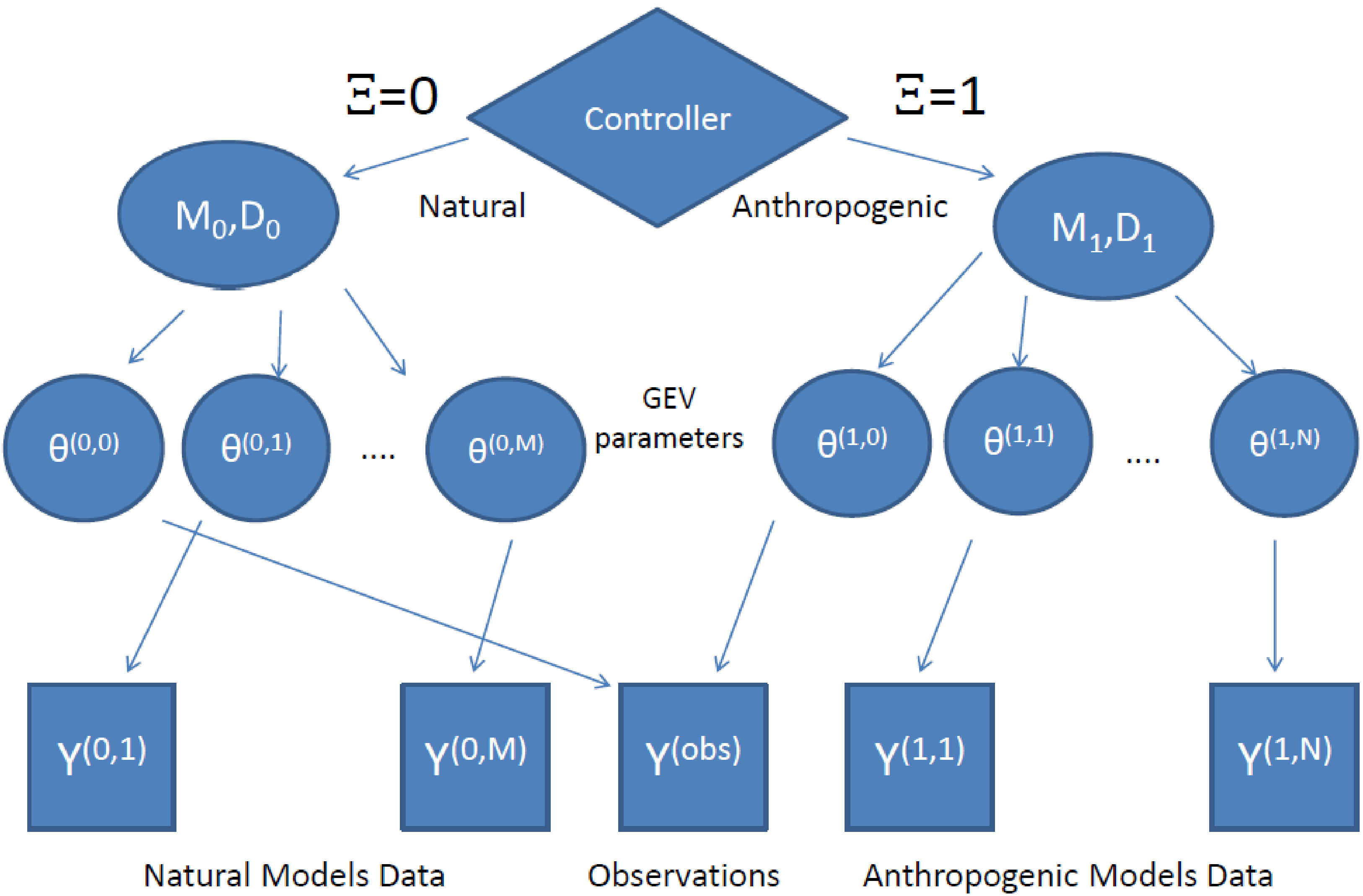


What's Next?

- Obvious strategy at this point is to rerun the GEV calculation on the model data
- But this runs into the *scale mismatch problem*: data plots shows that the models and observations are on different scales, so we should expect the extreme value parameters to be different as well
- Requires a more subtle approach – *hierarchical modeling*



Proposed Hierarchical Model



Bayesian Statistics Details

Model Specification

- $(M_1, D_1) \sim WN_q(A, m, M^*, F)$, Wishart-Normal prior with density $\propto |D_1|^{(m-q)/2} \exp\left[-\frac{1}{2}\text{tr}\left\{D_1\left(A + F(M_1 - M^*)(M_1 - M^*)^T\right)\right\}\right]$.
- Given M_1, D_1, $\theta^{(1,0)}, ..., \theta^{(1,N)}$ are IID $\sim N_q(M_1, D_1^{-1})$.
- Given $\theta^{(1,j)}$, $\mathbf{Y}^{(1,j)}$ generated by GEV with parameters $\theta^{(1,j)}$ ($\mathbf{Y}^{(\text{obs})}$ for $j = 0$, if $\Xi = 1$)
- Similar structure for M_0, D_0 etc.
- We can expand this model by defining $\theta^{(1,0)} \sim N_q(M_1, (\psi D_1)^{-1})$ where ψ represents departure from exchangeability ($\psi = 1$ is exchangeable). However, ψ is not identifiable — we can only try different values as a sensitivity check.

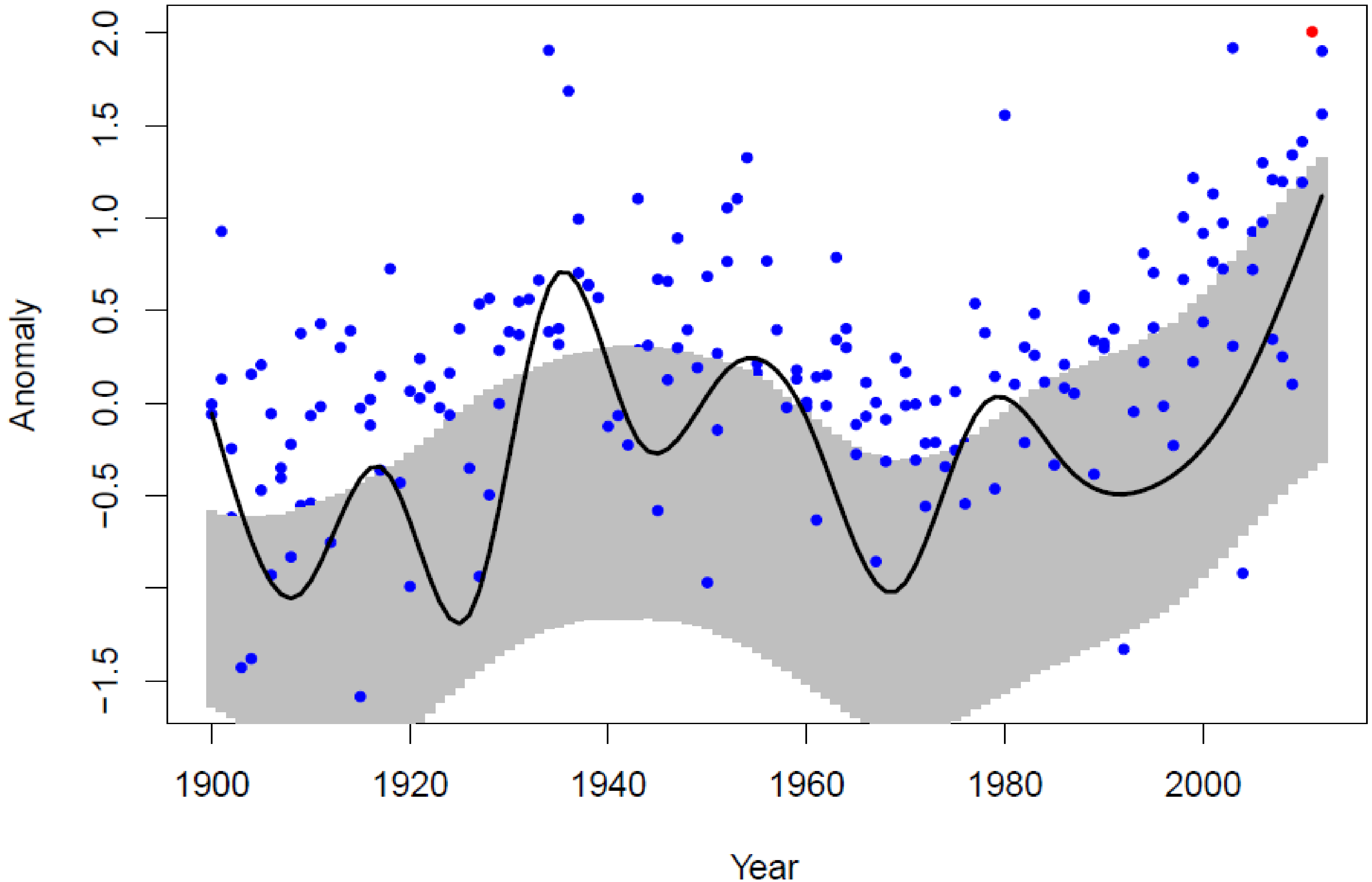
Computation

- $(M_1, D_1) \mid \theta^{(1,1)}, ..., \theta^{(1,N)} \sim WN_q(\tilde{A}, \tilde{m}, \tilde{M}^*, \tilde{F})$, where $\tilde{m} = m + N, \tilde{F} = F + N, \tilde{M}^* = \left(FM^* + \sum_{j=1}^N \theta^{(1,j)}\right)/\tilde{F}, \tilde{A} = A + FM^*M^{*T} + \sum_{j=1}^N \theta^{(j)}\theta^{(j)T} - \tilde{F}\tilde{M}^*\tilde{M}^{*T}$.
- Metropolis update for $\theta^{(1,1)}, ..., \theta^{(1,N)}$ given M_1, D_1 and $\mathbf{Y}$'s
- Metropolis update for $\theta^{(1,0)}$ based on conditional density

$$\exp\left\{-\frac{\psi}{2}\left(\theta^{(1,0)} - M_1\right)^T D_1 \left(\theta^{(1,0)} - M_1\right)\right\} \cdot L\left(\theta^{(1,0)}; \mathbf{Y}^{(\text{obs})}\right)$$

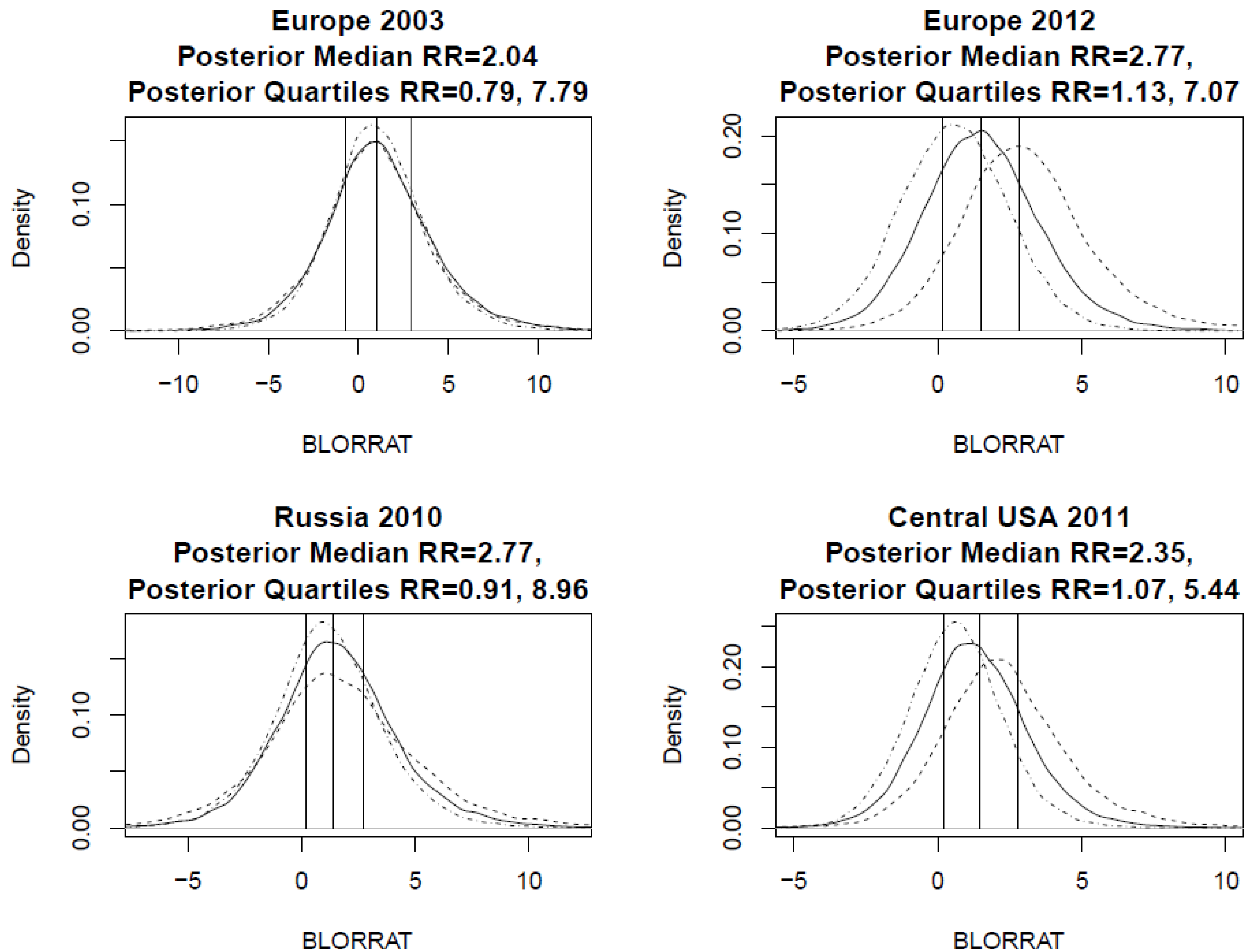
  where L is likelihood for $\theta^{(1,0)}$ given data $\mathbf{Y}^{(\text{obs})}$ and $\Xi = 1$
- Similar updates for $\Xi = 0$ side of picture; up to 1,000,000 iterations

Central USA Summer Mean Temperatures With Trend and Central 50% of Hierarchical Model Distribution

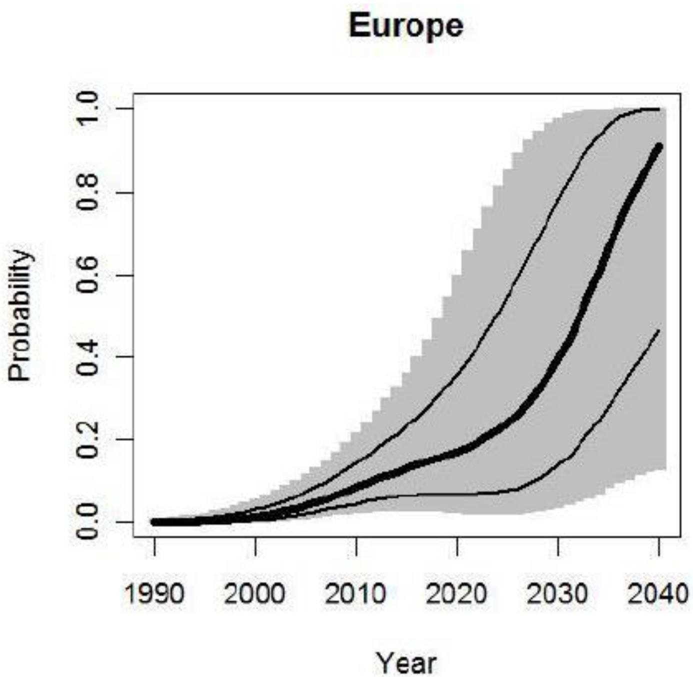


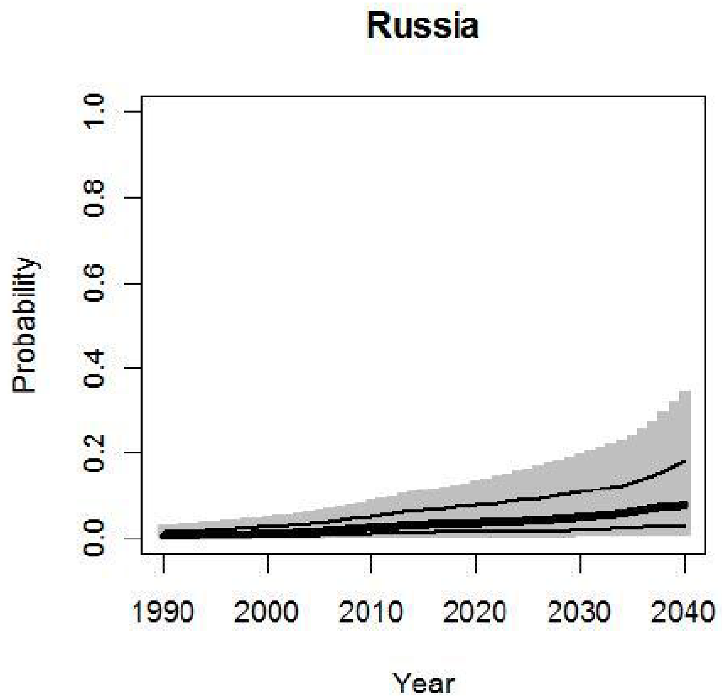
Post. Dens. for Binary Log Risk Ratio (BLORRAT)

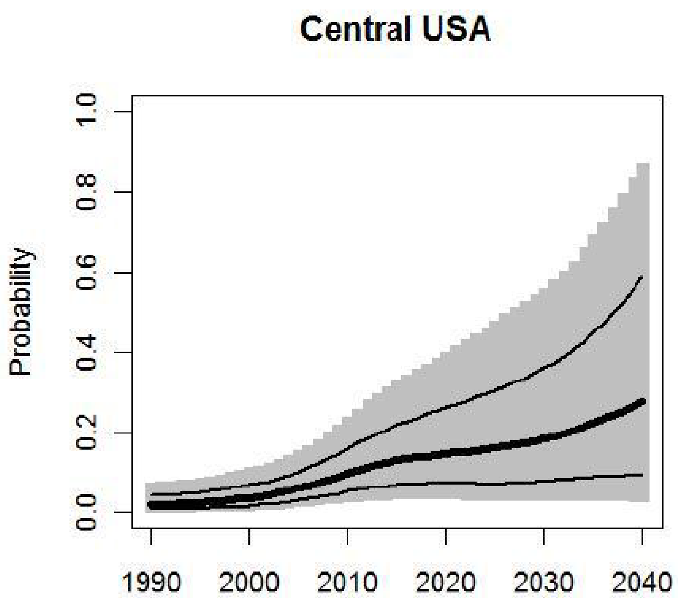
(numbers are for solid curves and equal weights; dashed curves allow for different weights between climate models and observations)



Changes in Projected Extreme Event Probabilities Over Time







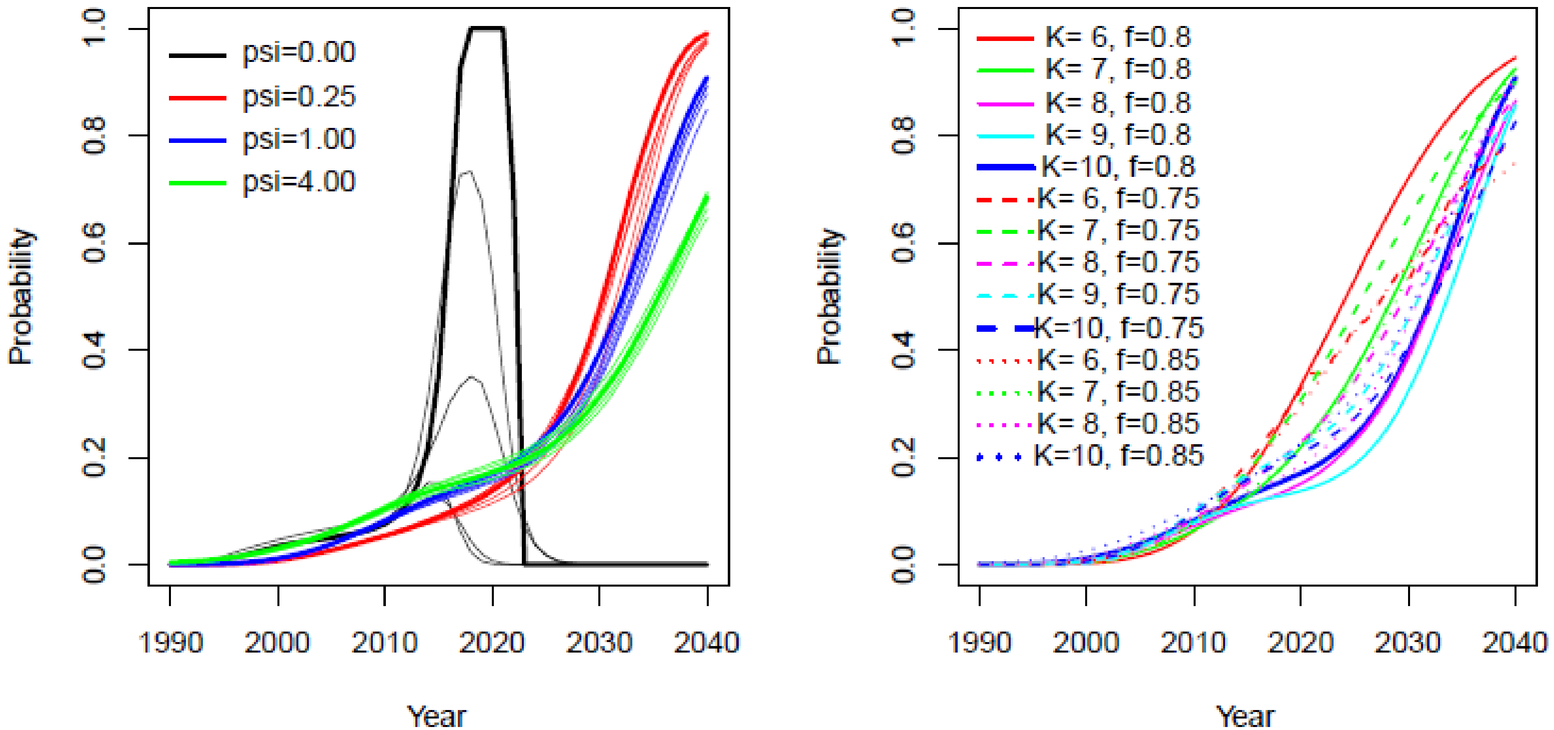
Central Solid Curve: Posterior Median

Thin Outer Curves: Posterior Quartiles

Outer Limits of Shaded Region:

Posterior 10th and 90th percentiles

Sensitivity Plots



Sensitivity plots for Europe. Left-hand figure: Plots of the posterior median probability of the extreme event for various weightings between models and observations, represented by psi, and with the Monte Carlo procedure repreated several times. Right-hand figure: Plots of the posterior median probability of the extreme event with various choices of the smoothness of the trend and the threshold of the distribution fit.

A Proposal for Bivariate Extension

- Motivation: dependence is important as well as marginal distributions
  - Different variables, e.g. temperature and precipitation
  - The same variable measured at different places
  - Using one meteorological variable as explanatory for another
- We give preliminary analyses of two examples
  - Precipitation patterns in the south-west USA are highly correlated with those of a region of South America (Herweijer and Seager 2008)
  - There may have been a connection between the 2010 Russian heatwave and the 2010 Pakistan floods (Lau and Kim 2012)
  - Focus here is on characterizing tail dependence

Russian Temperatures and Pakistan Precipitation

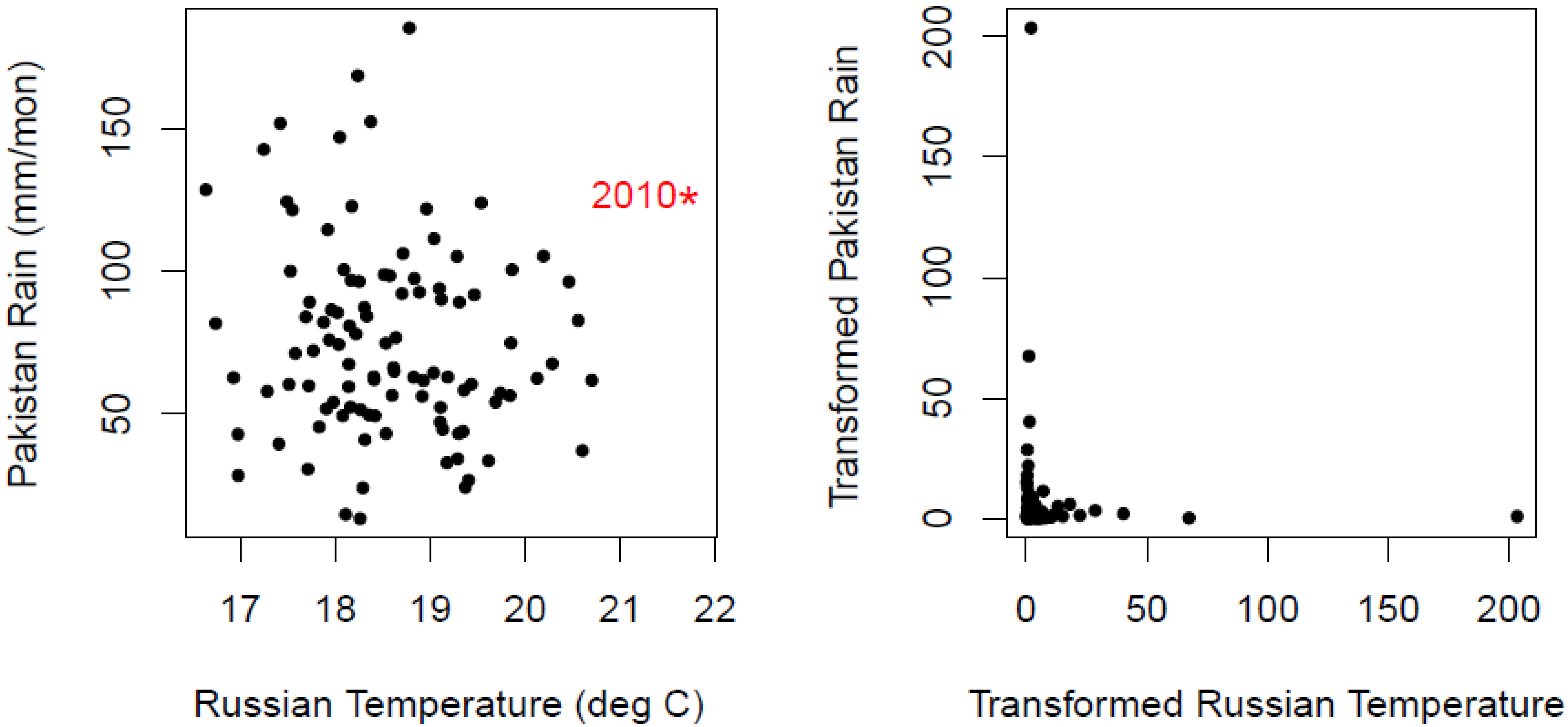


Figure 2. Left: Plot of JJA Russian temperature means against Pakistan JJA precipitation means, 1901-2002. Right: Same data with empirical transformation to unit Fréchet distribution. Data from CRU, as in Figure 1. The Russian data were averaged over 45-65°N, 30-60°E, while the Pakistan data were averaged over 32-35°N, 70-73°E, same as in Lau and Kim (2012).

SW-USA and Uruguay/Argentina Precipitation

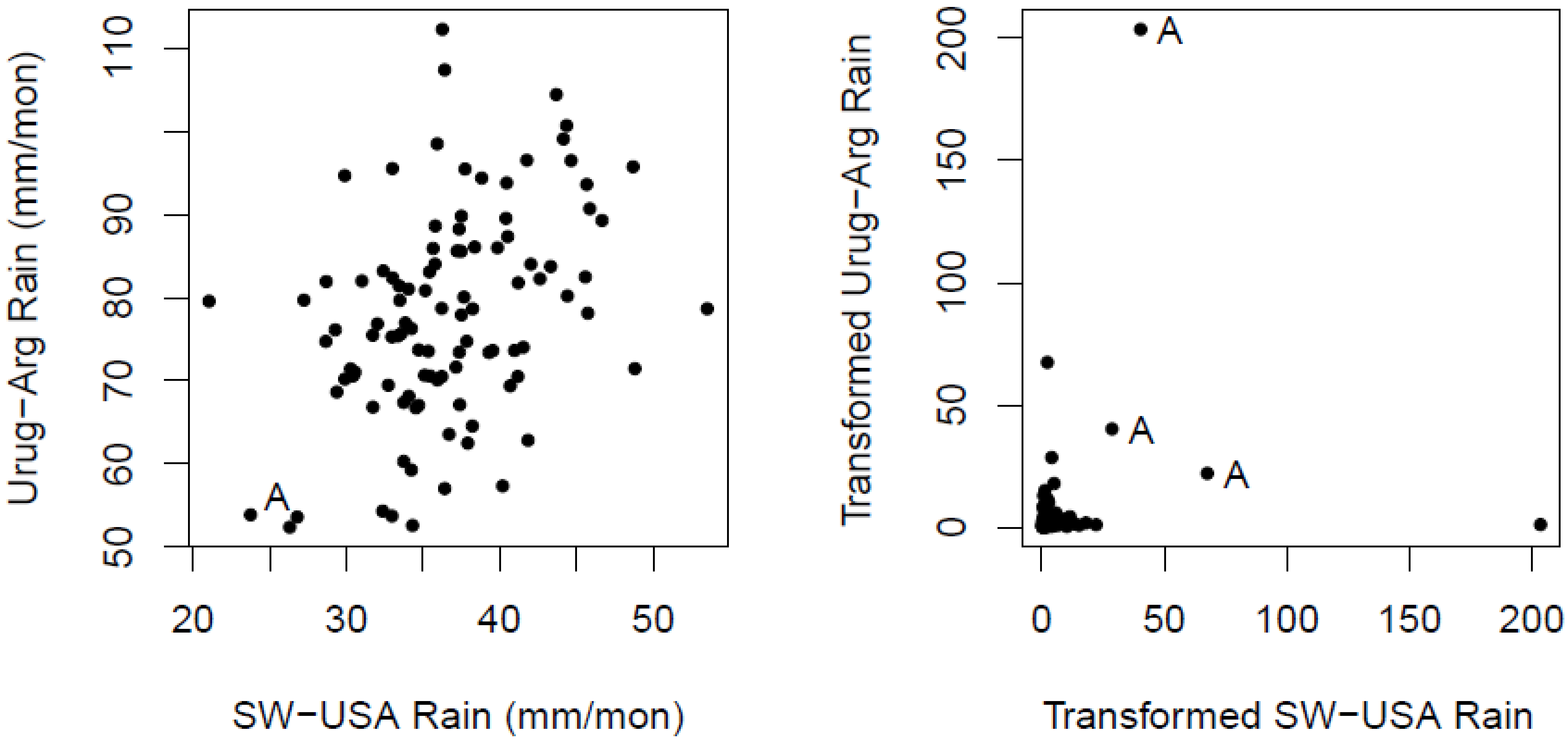


Figure 1. Left: Plot of USA annual precipitation means over latitudes 25-35°N, longitudes 95-120°W, against Argentina annual precipitation means over latitudes 30-40°S, longitudes 50-65°W, 1901-2002. Right: Same data with empirical transformation to unit Fréchet distribution. Data from gridded monthly precipitation means archived by the Climate Research Unit of the University of East Anglia.

Methods

- **Focus on the proportion by which the probability of a joint exceedance is greater than what would be true under independence.**
- **Method: Fit a joint bivariate model to the exceedances above a threshold on the unit Fréchet scale**
- **Two models:**
  - ***Classical logistic dependence model (Gumbel and Mustafi 1967; Coles and Tawn 1991)***
  - ***The η-asymmetric logistic model (Ramos and Ledford 2009)***

Ref: Alexandra Ramos and Anthony Ledford (2009), A new class of models for bivariate joint tails, *J.R. Statist. Soc. B* **71**, 219-241.

3.2. Smooth H_η

The model that is detailed here is based on a modified version of the asymmetric logistic dependence structure of classical bivariate extremes (Tawn, 1988). Suppose that H_η has density h_η given by

$$h_\eta(w) = \frac{\eta - \alpha}{\alpha \eta^2 N_\rho} \left\{ (\rho w)^{-1/\alpha} + \left(\frac{1-w}{\rho} \right)^{-1/\alpha} \right\}^{\alpha/\eta - 2} \{w(1-w)\}^{-(1+1/\alpha)} \tag{3.1}$$

for $0 < w < 1$ where $N_\rho = \rho^{-1/\eta} + \rho^{1/\eta} - (\rho^{-1/\alpha} + \rho^{1/\alpha})^{\alpha/\eta}$ and $\eta \in (0,1]$, $\alpha > 0$ and $\rho > 0$. It is straightforward to show that h_η as defined above satisfies the normalizing condition (2.5) and so, by equation (2.3), we have

$$\bar{F}_{ST}(s,t) = N_\rho^{-1} \left[(\rho s)^{-1/\eta} + \left(\frac{t}{\rho} \right)^{-1/\eta} - \left\{ (\rho s)^{-1/\alpha} + \left(\frac{t}{\rho} \right)^{-1/\alpha} \right\}^{\alpha/\eta} \right] \tag{3.2}$$

where $(s,t) \in [1,\infty) \times [1,\infty)$. The marginal survivor functions of S and T, as given by equations (2.6), have leading terms that behave like powers of s or t. For example, $\Pr(S > s)$ behaves like $s^{-1/\eta}$ if $\alpha \leqslant \eta$ and $s^{-1/\alpha}$ if $\alpha > \eta$ for large s. Thus the marginal tails of (S,T) can be heavier or of the same heaviness as the joint survivor function, depending on the relative sizes of α and η.

Results

	Logistic Model		Ramos-Ledford Model	
	Estimate	90% CI	Estimate	90% CI
10-year	2.7	(1.2 , 4.2)	2.9	(1.2 , 5.0)
20-year	4.7	(1.4 , 7.8)	4.9	(1.2 , 9.6)
50-year	10.8	(2.1 , 18.8)	9.9	(1.4 , 23.4)

Table 1. Estimates of the increase in probability of a joint extreme event in both variables, relative to the probability under independence, for the USA/Uruguay-Argentina precipitation data. Shown are the point estimate and 90% confidence interval, under both the logistic model and the Ramos-Ledford model.

	Logistic Model		Ramos-Ledford Model	
	Estimate	90% CI	Estimate	90% CI
10-year	1.01	(1.00 , 1.01)	0.33	(0.04 , 1.4)
20-year	1.02	(1.00 , 1.03)	0.21	(0.008 , 1.8)
50-year	1.05	(1.01 , 1.07)	0.17	(0.001 , 2.9)

Table 2. Similar to Table 1, but for the Russia-Pakistan dataset.

What Next?

- We have demonstrated an empirical tail dependence for the North and South American precipitations data, but not for Russian temperatures and Pakistan precipitation
- To use this for future climate projections, need to show that the same phenomenon occurs in climate model
  - Likely to raise similar issues of "scale mismatch" (or "dependence mismatch")
  - Possibility of using a hierarchical model in this case
- The advantage of the Ramos-Ledford approach is that it's a parametric model going beyond classical BEVT (but there are other approaches, such as the Heffernan-Tawn conditional approach)

Conclusions

- Extreme value theory provides a viable method for estimating extreme event probabilities in the presence of a trend
- For combining observations with climate models, we propose a hierarchical model that allows for systematic discrepancies between models and observations
- For each of Russia 2010, Central USA 2011 and Europe 2012 events, estimated risk ratio is at least 2.3, and it's *likely* (probability at least .66) that the risk ratio is >1.5.
- We also computed future projections of extreme event probabilities; sharp increase for Europe; much less so for the other two regions studied
- Likely future work will involve bivariate or spatial dependence