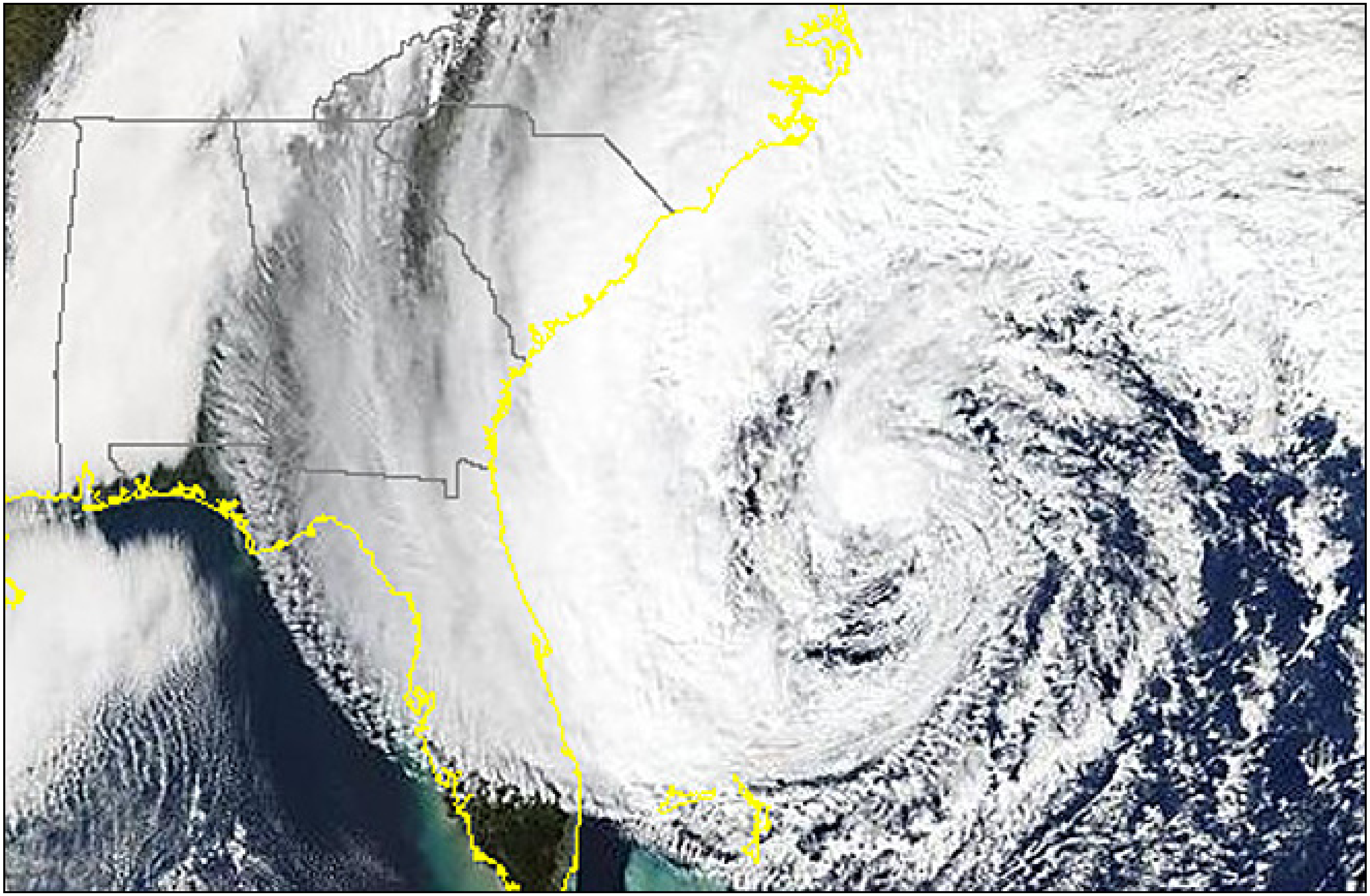
ATTRIBUTION OF EXTREME CLIMATIC EVENTS

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

Statistical and Applied Mathematical Sciences Institute

Wierman Lecture
Department of Applied Mathematics and Statistics
Johns Hopkins University
December 6, 2012

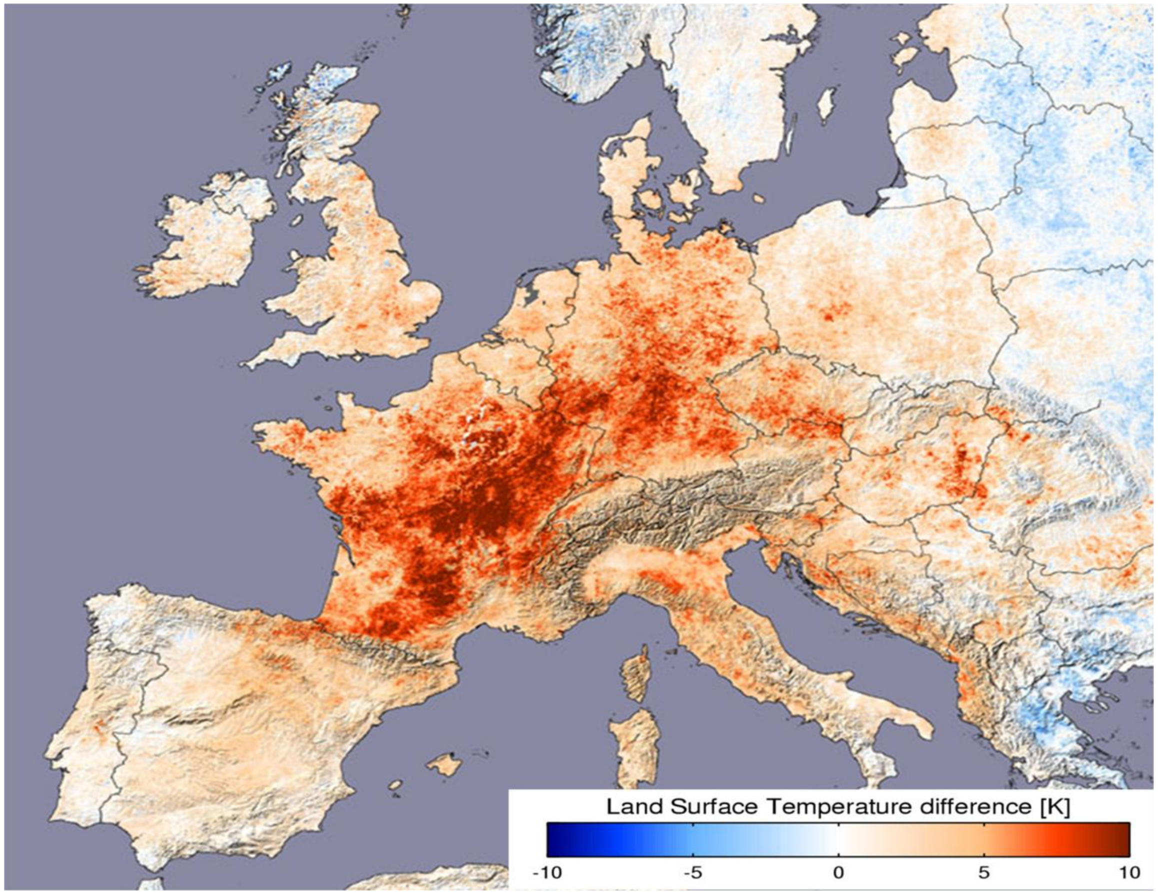
With Michael Wehner (Lawrence Berkeley Laboratory)



Superstorm Sandy on October 27 2012 (Scott Sistek)

Superstorm Sandy (www.guardian.co.uk; October 30, 2012)

Superstorm Sandy (www.cnn.com; October 31, 2012)



European temperatures in early August 2003, relative to 2001-2004 average

From NASA's MODIS - Moderate Resolution Imaging Spectrometer, courtesy of Reto Stöckli, ETHZ

I Introduction

II Literature Review

III Why is Extreme Value Modeling Hard?

IV Data

V The Generalized Extreme Value Distribution (GEV)

VI Analysis of a Single Time Series

VII Hierarchical Models

VIII Results

IX Conclusions

I Introduction

Concept of Extreme Event Attribution

- 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* $\frac{P_1}{P_0}$.

A recent report of the NRC:

This PDF is available from The National Academies Press at http://www.nap.edu/catalog.php?record_id=14682

Climate and Social Stress: Implications for Security Analysis

ISBN
978-0-309-27856-0

280 pages
6 x 9
PAPERBACK (2012)

John D. Steinbruner, Paul C. Stern, and Jo L. Husbands, Editors; Committee on Assessing the Impact of Climate Change on Social and Political Stresses; Board on Environmental Change and Society; Division of Behavioral and Social Sciences and Education; National Research Council

Question: How will probabilities of extreme events change over the next ten years?

Statement of the Problem

- Find a statistically defensible strategy for estimating the FAR or RR based on time series of observational or model data

- Characterize the uncertainty (e.g. through confidence or Bayesian credible intervals)

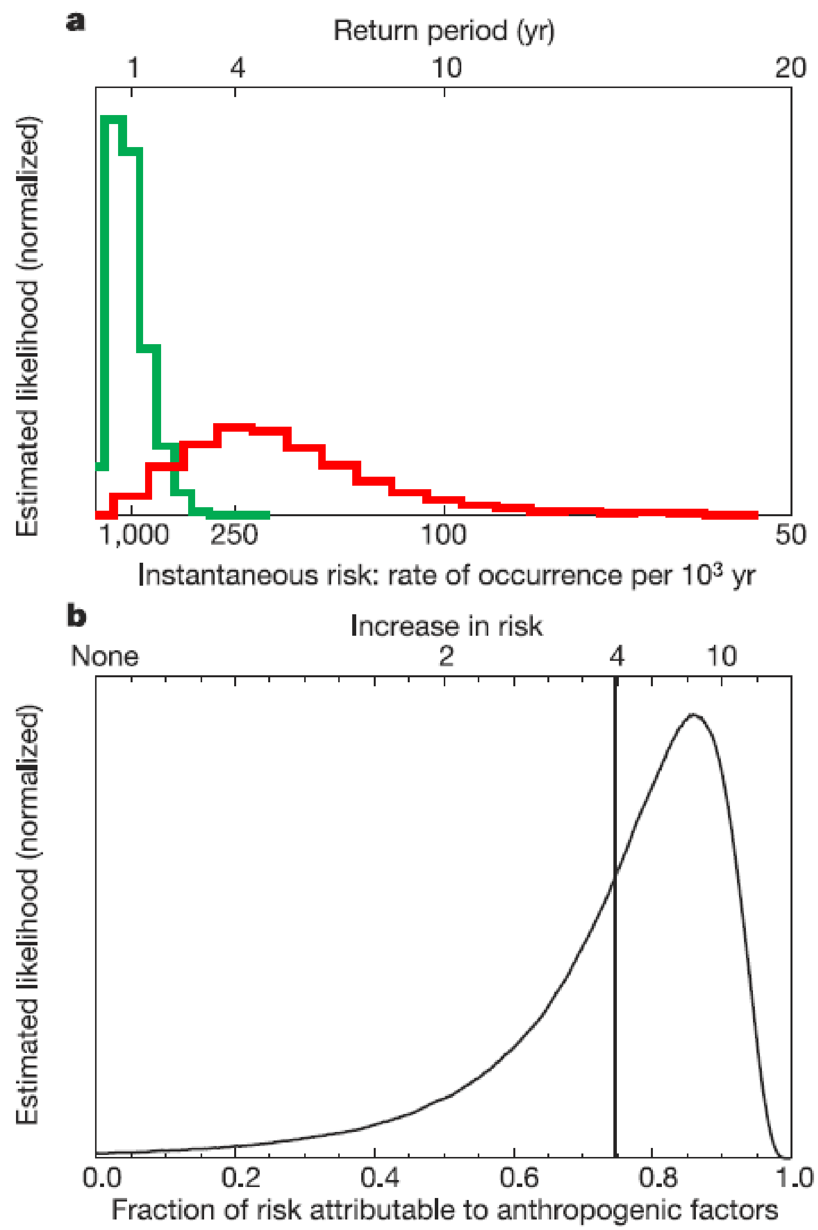
- The objective is to provide portable R software that is applicable to public databases

- Also: find ways of projecting extreme event probabilities into the future

- We illustrate these issues with regard to three recent heatwave events: the European heatwave of 2003, the Russian heatwave of 2010, the central USA heatwave of 2011.

II Literature Review

Stott, Stone and Allen (Nature, 2004) used the GPD plus bootstrapping to estimate the probability of an exceedance of 1.6K using both natural forcings (green curve shows pdf of estimated return period) and anthropogenic forcings (red curve). The results show an estimated probability of around 1/250 per year under anthropogenic forcing and around 1/1000 per year under natural forcing, for a risk ratio of 4 or a FAR of 1-1/4=0.75. The bottom curve expresses the pdf of the estimated FAR. They concluded that there is at least a 90% chance that the FAR is >0.5 (risk ratio >2).



LETTER

doi:10.1038/nature09762

Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000

Pardeep Pall[1,2]†, Tolu Aina[3], Dáithí A. Stone[1,4], Peter A. Stott[5], Toru Nozawa[6], Arno G. J. Hilberts[7], Dag Lohmann[7] & Myles R. Allen[1,4]

Interest in attributing the risk of damaging weather-related events to anthropogenic climate change is increasing[1]. Yet climate models used to study the attribution problem typically do not resolve the weather systems associated with damaging events[2] such as the UK floods of October and November 2000. Occurring during the wettest autumn in England and Wales since records began in 1766[3,4], these floods damaged nearly 10,000 properties across that region, disrupted services severely, and caused insured losses estimated at £1.3 billion (refs 5, 6). Although the flooding was deemed a 'wake-up call' to the impacts of climate change at the time[7], such claims are typically supported only by general thermodynamic arguments that suggest increased extreme precipitation under global warming, but fail[8,9] to account fully for the complex hydrometeorology[4,10] associated with flooding. Here we present a multi-step, physically based 'probabilistic event attribution' framework showing that it is very likely that global anthropogenic greenhouse gas emissions substantially increased the risk of flood occurrence in England and Wales in autumn 2000. Using publicly volunteered distributed computing[11,12], we generate several thousand seasonal-forecast-resolution climate model simulations of autumn 2000 weather, both under realistic conditions, and under conditions as they might have been had these greenhouse gas emissions and the resulting large-scale warming never occurred. Results are fed into a precipitation-runoff model that is used to simulate severe daily river runoff events in England and Wales (proxy indicators of flood events). The precise magnitude of the anthropogenic contribution remains uncertain, but in nine out of ten cases our model results indicate that twentieth-century anthropogenic greenhouse gas emissions increased the risk of floods occurring in England and Wales in autumn 2000 by more than 20%, and in two out of three cases by more than 90%.

Pall et al. (2011), nonparametric, very data-intensive

Perceptions of Climate Change: The New Climate Dice

James Hansen[a1], Makiko Sato[a], Reto Ruedy[b]

[a]NASA Goddard Institute for Space Studies and Columbia University Earth Institute, [b]Sigma Space Partners, New York, NY 10025

"Climate dice", describing the chance of unusually warm or cool seasons relative to climatology, have become progressively "loaded" in the past 30 years, coincident with rapid global warming. The distribution of seasonal mean temperature anomalies has shifted toward higher temperatures and the range of anomalies has increased. An important change is the emergence of a category of summertime extremely hot outliers, more than three standard deviations (σ) warmer than climatology. This hot extreme, which covered much less than 1% of Earth's surface in the period of climatology, now typically covers about 10% of the land area. We conclude that extreme heat waves, such as that in Texas and Oklahoma in 2011 and Moscow in 2010, were "caused" by global warming, because their likelihood was negligible prior to the recent rapid global warming. We discuss practical implications of this substantial, growing climate change.

GEOPHYSICAL RESEARCH LETTERS, VOL. 38, L06702, doi:10.1029/2010GL046582, 2011

Was there a basis for anticipating the 2010 Russian heat wave?

Randall Dole,[1] Martin Hoerling,[1] Judith Perlwitz,[2] Jon Eischeid,[2] Philip Pegion,[2] Tao Zhang,[2] Xiao-Wei Quan,[2] Taiyi Xu,[2] and Donald Murray[2]

[1] The 2010 summer heat wave in western Russia was extraordinary, with the region experiencing the warmest July since at least 1880 and numerous locations setting all-time maximum temperature records. This study explores whether early warning could have been provided through knowledge of natural and human-caused climate forcings. Model simulations and observational data are used to determine the impact of observed sea surface temperatures (SSTs), sea ice conditions and greenhouse gas concentrations. Analysis of forced model simulations indicates that neither human influences nor other slowly evolving ocean boundary conditions contributed substantially to the magnitude of this heat wave. They also provide evidence that such an intense event could be produced through natural variability alone. Analysis

Anatomy of an Extreme Event

Martin Hoerling[1], Arun Kumar[2], Randall Dole[1], John W. Nielsen-Gammon[3], Jon Eischeid,[1,4], Judith Perlwitz[1,4], Xiao-Wei Quan[1,4], Tao Zhang[1,4], Philip Pegion[1,4], and Mingyue Chen[2]

[1]*NOAA Earth System Research Laboratory, Boulder, Colorado*

[2] *NOAA Climate Prediction Center, Camp Springs, MD*

[3] *Department of Atmospheric Sciences, Texas A&M University, College Station, Texas*

[4] *University of Colorado, Cooperative Institute for Research in Environmental Sciences, Boulder, Colorado*

The record-setting 2011 Texas drought/heat wave is examined to identify physical processes, underlying causes, and predictability. October 2010-September 2011 was Texas's driest 12-month period on record. While the summer 2011 heat wave magnitude (2.9°C above the 1981-2010 mean) was larger than the previous record, events of similar or larger magnitude appear in pre-industrial control runs of climate models. The principal factor contributing to the heat wave magnitude was a severe rainfall deficit during antecedent and concurrent seasons related to anomalous sea surface temperatures (SSTs) that included a La Niña event. Virtually

III Why is Extreme Value Modeling Hard?

To provide some context for the results and modeling methods to follow, we give some elementary theoretical calculations that show why we cannot expect to get very precise results based on small numbers of climate model runs.

Suppose we try to estimate the probability of an extreme events as the proportion X/n from a set of n climate model runs where X contain the event in question (this is essentially the technique of Pall *et al.*).

Suppose $X \sim \text{Bin}(n, p)$ and we are interested in testing $H_0: \ p = p_0$ versus $H_1: \ p = kp_0$ for some $k > 1$.

If we fix the size of the test to be 0.05 and the power to be 0.8, how big a sample size n do we need?

p_0	k								
	2	3	4	5	6	7	8	9	10
0.05	169	52	27	16	14	12	7	6	5
0.025	339	104	54	43	28	24	14	13	11
0.01	905	301	137	110	71	60	53	33	29
0.0075	1207	402	223	146	94	81	71	44	39
0.005	1811	604	335	220	142	122	106	66	59
0.0025	3623	1209	671	440	285	244	213	133	119
0.001	9061	3024	1679	1102	713	611	534	332	299
0.00075	12082	4032	2239	1470	950	814	713	443	399
0.0005	18124	6049	3360	2205	1426	1222	1069	665	598
0.00025	36249	12099	6720	4411	2852	2445	2139	1330	1197
0.0001	90624	30250	16802	11029	7131	6112	5348	3327	2994

Sample size required to reject null hypothesis $p = p_0$ against alternative hypothesis $p = kp_0, \ k > 1$ in a binomial experiment of size 0.05 and power 0.8

Conclusion:

By a simple argument of counting threshold exceedances in repeated climate model runs, we would expect to need sample sizes from a few tens up to several hundred to distinguish extreme event probabilities that are typical in these discussions.

Most published climate model runs contain between 1 and 5 replications of the same model, so direct estimation is unlikely to work.

IV Data

Climate model runs have been downloaded from the WCRP CMIP3 Multi-Model Data website (http://esg.llnl.gov:8080/index.jsp)

Three kinds of model runs:

- Twentieth-century
- Pre-industrial control model runs (used a proxy for natural forcing)
- Future projections (A2 scenario)

We also took observational data ($5^o \times 5^o$ gridded monthly temperature anomalies) from the website of the Climate Research Unit of the University of East Anglia (www.cru.uea.ac.uk — HadCRUT3v dataset)

Number	Model	Control runs	20C runs	A2 runs
1	bccr_bcm2_0	2	1	1
2	cccma_cgcm3_1	10	5	5
3	cnrm_cm3	5	1	1
4	csiro_mk3_0	3	3	1
5	gfdl_cm2_1	5	3	1
6	giss_model_e_r	5	9	1
7	ingv_echam4	1	1	1
8	inmcm3_0	3	1	1
9	ipsl_cm4	7	1	1
10	miroc3_2_medres	5	3	3
11	mpi_echam5	5	4	3
12	mri_cgcm2_3_2a	3	5	5
13	ncar_ccsm3_0	7	5	5
14	ukmo_hadcm3	3	2	1

List of climate models, including numbers of runs available under three scenarios

We calculated summer (JJA) averages of temperature anomaly for each of three regions:

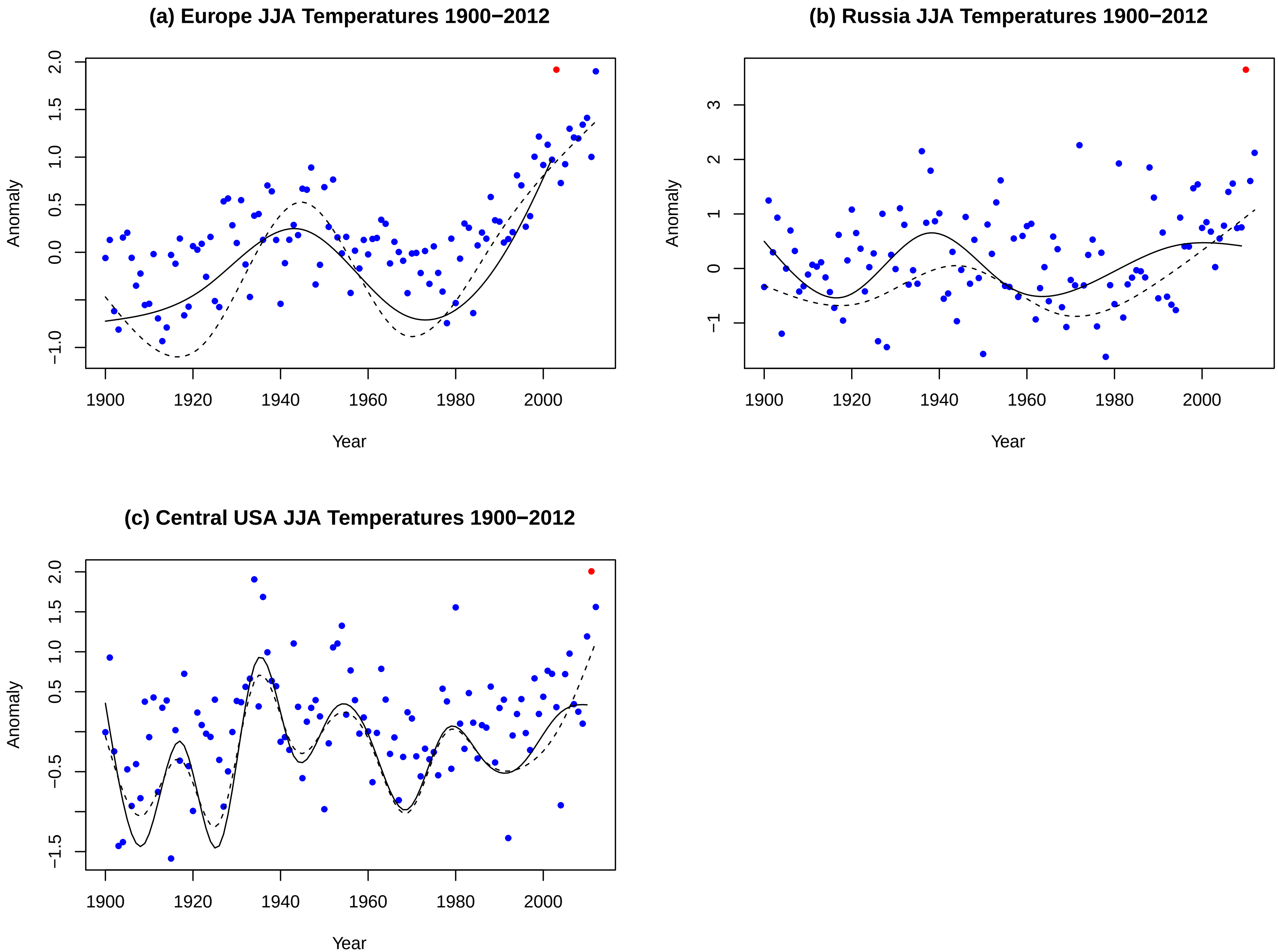
- Europe — 10^o W to 40^o E, 30^o—50^o N

- Russia — 30^o to 60^o E, 45^o—65^o N

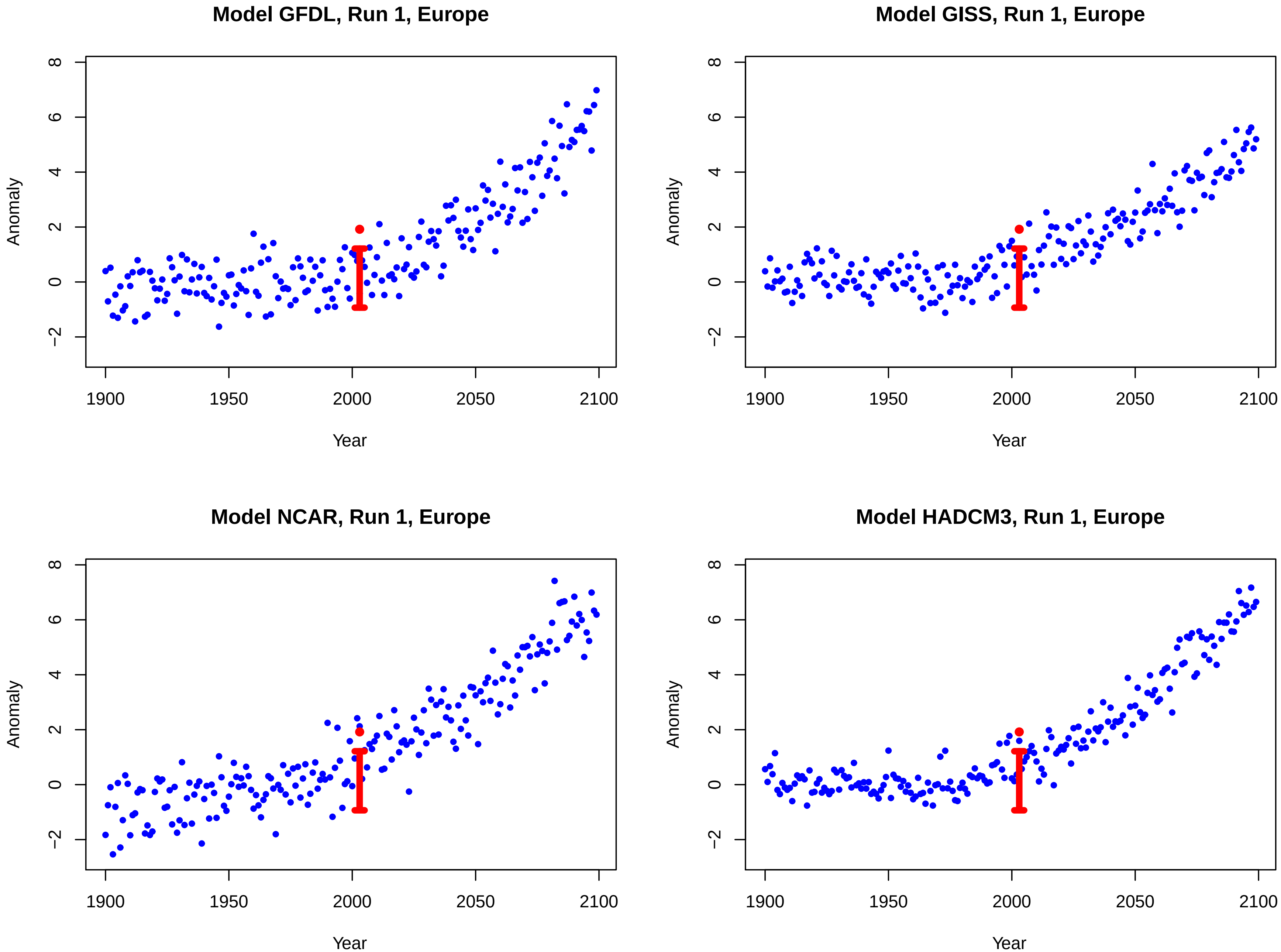
- Central USA — 90^o to 105^o W, 25^o—45^o N

Plots of the time series show both increasing and decreasing trends

We also note a "scale mismatch" problem — the variances of the observational and model time series are typically different, meaning that we cannot expect to use the model data directly to calculate extreme event probabilities for the osbervations



Plot of three time series for 1900–2012, with fitted trend curves.



Scale mismatch: 4 model runs (range of observations in red).

V The Generalized Extreme Value Distribution (GEV)

The Generalized Extreme Value Distribution (GEV)

- Three-parameter distribution, derived as the general form of limiting distribution for extreme values (Fisher-Tippett 1928, Gnedenko 1943)
- μ, σ, ξ known as location, scale and shape parameters
- $\xi > 0$ represents long-tailed distribution, $\xi < 0$ short-tailed

Formula:

$$\Pr\{Y \leq y\} \;=\; \exp\left[-\left\{1+\xi\left(\frac{y-\mu}{\sigma}\right)\right\}_{+}^{-1/\xi}\right].$$

- *Peaks over threshold* approach implies that the GEV can be used generally to study the tail of a distribution: assume GEV holds exactly above a threshold u and that values below u are treated as left-censored

- Time trends by allowing μ, σ, ξ to depend on time

- *Example:* Allow $\mu_t = \beta_0 + \sum_{k=1}^{K} \beta_k x_{kt}$ where $\{x_{kt},\ k = 1, ..., K,\ t = 1, ..., T\}$ are spline basis functions for the approximation of a smooth trend from time 1 to T with K degrees of freedom

- Critical questions:
  - Determination of threshold and K
  - Estimating the probability of exceeding a high value such as 2.3K

VI Analysis of a Single Time Series

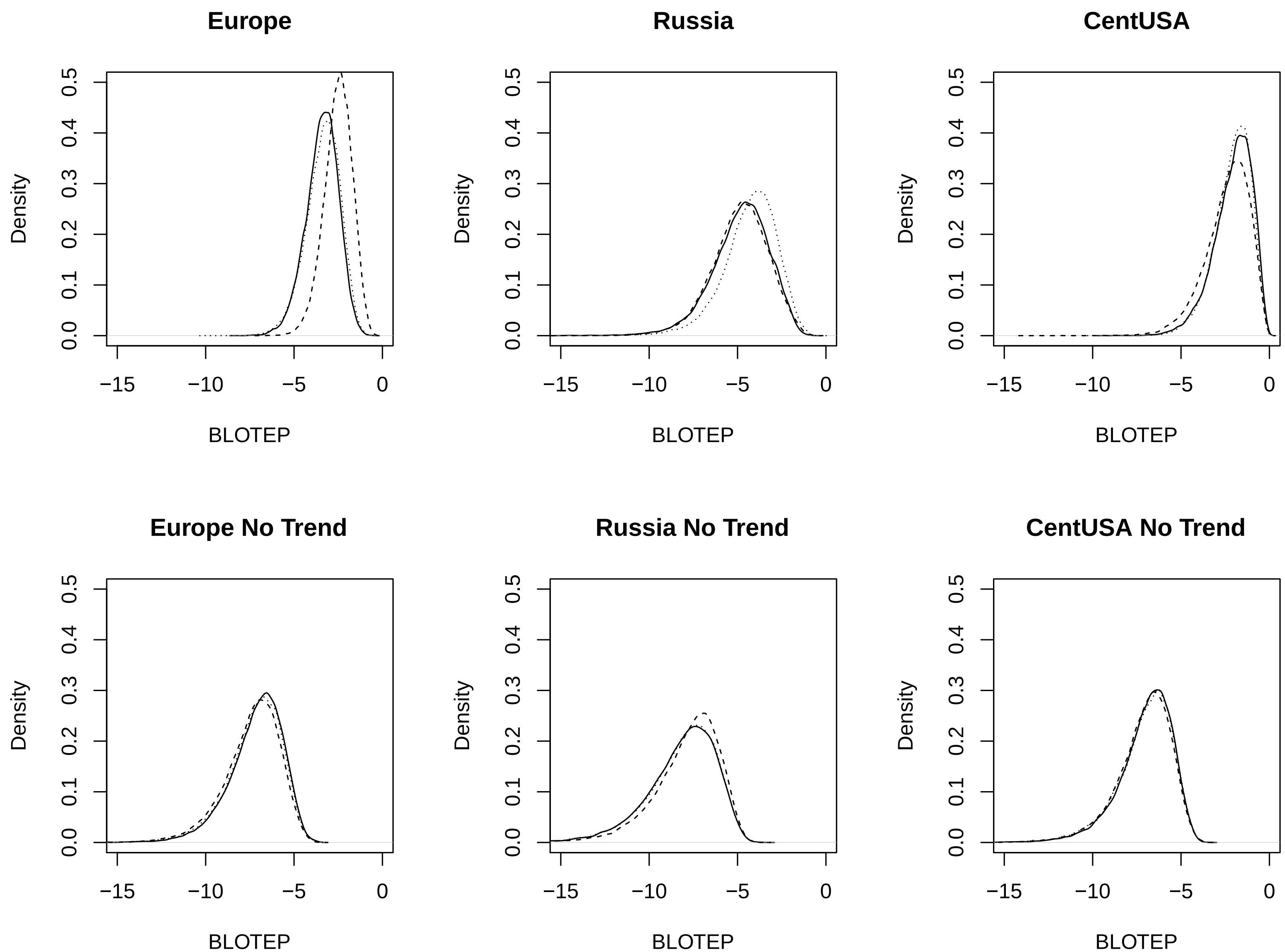
- GEV with trend fitted to three observational time series
- Threshold was chosen as fixed quantile — 75th, 80th or 85th percentile
- AIC was used to help select the number of spline basis terms K
- Estimate probability of extreme event by maximum likelihood (MLE) or Bayesian method
- Repeat the same calculation with no spline terms
- Use full series or part?
- Examine sensitivity to threshold choice through plots of the posterior densities.

K Threshold	Europe 75%	80%	85%	Russia 0.75	0.8	0.85	Texas 0.75	0.8	0.85
2	97.9	87.7	67.5	149.8	131.2	110.4	146.6	131.3	108.8
3	75.7	68.5	60.5	145.8	135.4	112.7	142.6	125.0	105.5
4	76.1	66.2	44.9	148.1	137.8	113.8	144.6	126.8	103.6
5	74.1	64.6	54.6	147.0	134.1	121.2	144.1	126.5	104.9
6	74.2	74.3	61.6	146.8	133.6	113.1	143.8	125.5	106.1
7	77.9	75.2	59.8	146.6	135.1	114.0	133.4	126.4	106.8
8	86.2	77.4	65.9	148.0	137.1	122.1	138.9	128.4	108.1
9	86.8	74.6	67.1	149.4	138.7	113.3	148.6	130.6	110.2
10	88.7	94.8	54.2	150.8	140.4	125.1	128.2	122.9	105.7
11	90.6	73.4	73.5	153.1	142.6	125.7	144.2	127.8	110.5
12	79.1	98.6	59.3	152.8	140.8	126.4	135.1	119.7	105.8
13	95.3	79.6	59.1	156.1	144.2	127.4	136.2	116.9	104.2
14	77.5	78.6	54.6	157.5	142.4	128.7	138.9	121.8	107.9
15	97.6	85.5	77.9	157.2	143.1	129.5	136.8	122.5	109.6

AIC values for different values of K, at three different thresholds, for each dataset of interest. In each column, the smallest three AIC values are indicated in red, green and blue respectively.

Dataset	Endpoint	K	Threshold	MLE	Posterior Mean	Posterior Quantiles 0.05	0.5	0.95
Europe	2002	5	80%	.021	.076	0	.057	.217
Europe	2012	5	80%	.0027	.113	.031	.098	.246
Europe	2002	0	80%	0	.0004	0	0	.002
Europe	2012	0	80%	.0044	.011	.001	.0081	.029
Russia	2009	6	80%	.0013	.010	0	.004	.040
Russia	2012	5	80%	.010	.058	.005	.039	.181
Russia	2009	0	80%	0	.0011	0	0	.0069
Russia	2012	0	80%	.0019	.0067	.0003	.0043	.021
CentUSA	2010	13	80%	.0007	.072	.003	.045	.234
CentUSA	2012	13	80%	.089	.300	.058	.268	.653
CentUSA	2010	0	80%	.0023	.0078	.00007	.0052	.024
CentUSA	2012	0	80%	.005	.012	.001	.0092	.031

Results of extreme value analysis applied to observational datasets. For three datasets (Europe, Russia, Central USA), different choices of the endpoint of the analysis, spline degrees of freedom K, and threshold, we show the maximum likelihood estimate (MLE) of the probability of the extreme event of interest, as well as the posterior mean and three quantiles of the posterior distribution.



Posterior densities of the BLOTEP, with (top) and without (bottom) spline-based trends. Based on 80% (solid curve), 75% (dashed) and 85% (dot-dashed) thresholds.

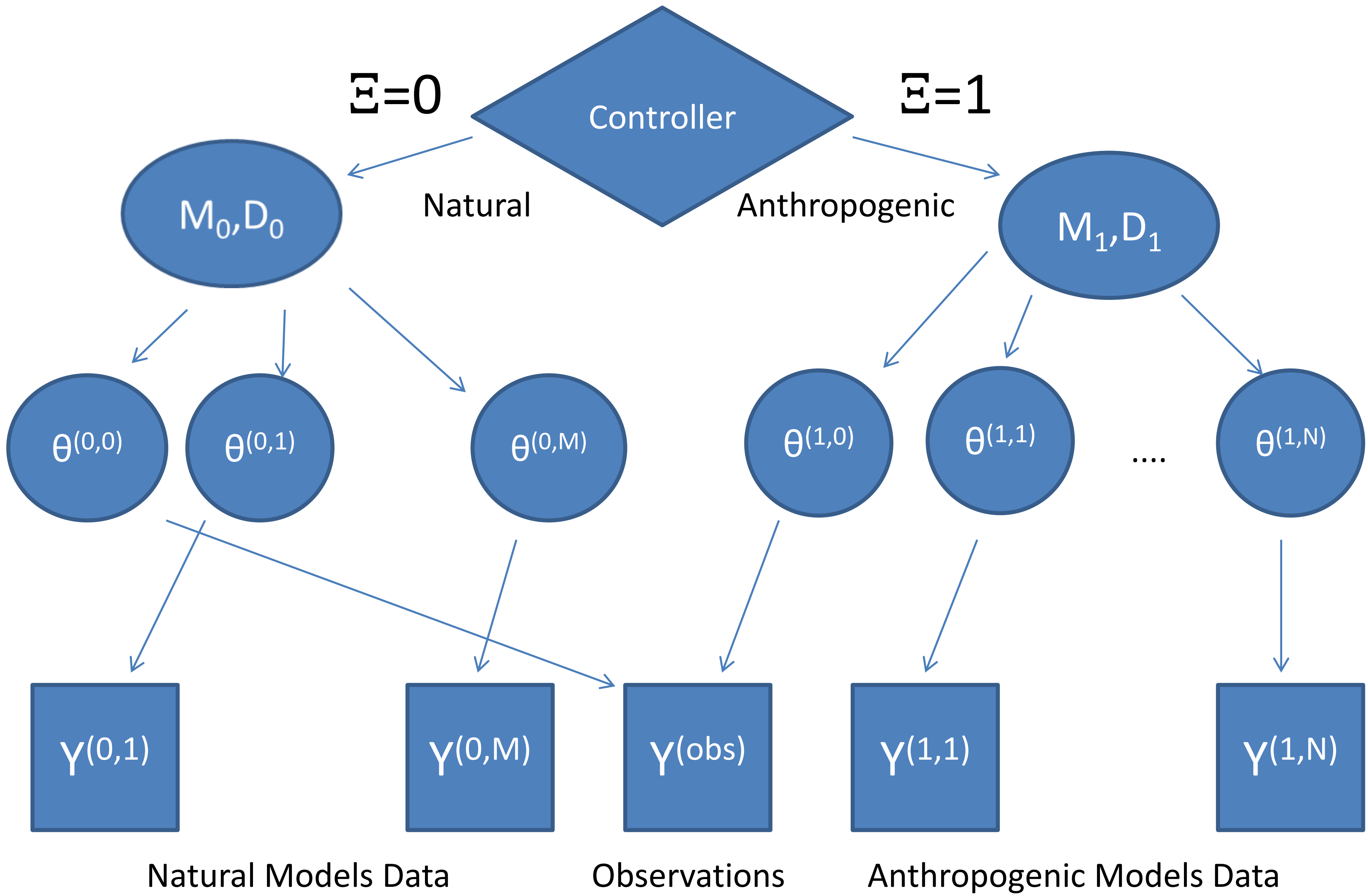
Summary So Far:

- Estimate extreme event probabilities by GEV with trends
- Bayesian posterior densities best way to describe uncertainty
- Two major disadvantages:
  - No way to distinguish anthropogenic climate change effects from other short-term fluctations in the climate (El Niños and other circulation-based events; the 1930s dust-bowl in the US)
  - No basis for projecting into the future

To go further, we need to find a way to combine observational and climate model data in a way that takes account of the scale mismatch issue noted earlier. We now propose a *hierarchical modeling* approach to do this.

VII Hierarchical Models

Hierarchical Model



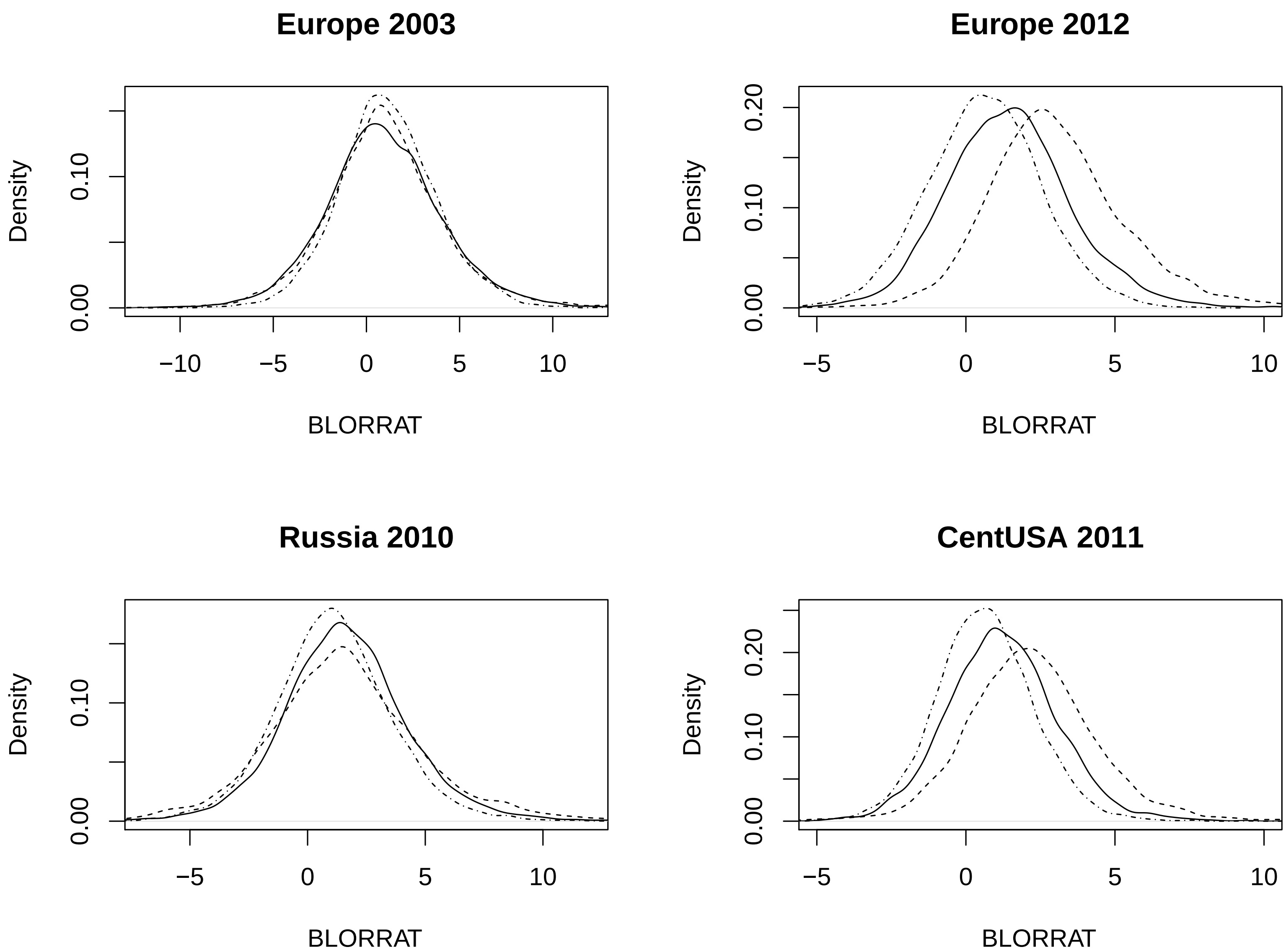
Features of the Present Approach

- A "switching" variable Ξ deterines whether climate change is natural or anthropogenic
- Conditional on $\Xi = 0$ or 1, define mean vector (M_0 or M_1) and precision matrix (D_0 or D_1) for GEV parameters $\theta^{(i,j)}$
- Bottom level of hierarchy contains observational and model data
- Normal-Wishart conjugate prior for (M_0, D_0) or (M_1, D_1) — allows Gibbs updating
- Metropolis updating for GEV parameters
- Ultimately use posterior density for observation GEV parameters ($\theta^{(0,0)}$ or $\theta^{(1,0)}$) to calculate extreme event probabilities
- A refinement — multiply the precision matrix for $\theta^{(0,0)}$ or $\theta^{(1,0)}$ by ψ to allow non-exhangeability.

VIII Results

For each of the three time series, we calculate the binary log risk ratio (BLORRAT) for the extreme event of interest, corresponding to the European, Russian and Central USA heatwaves of 2003, 2010 and 2011 respetively.

Compute posterior densities under variety of assumptions, including varying ψ.

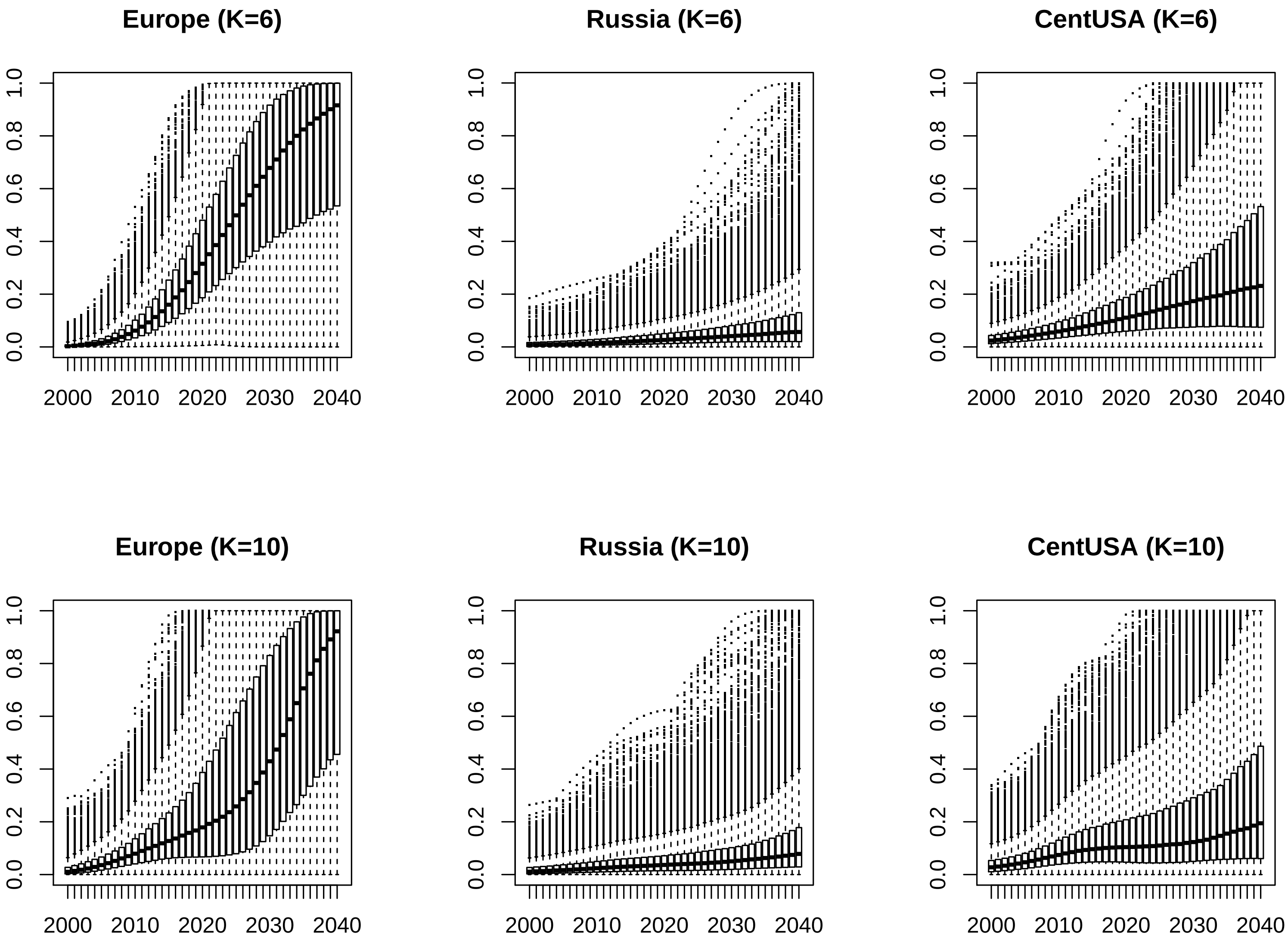


Posterior densities of the BLORRAT. Solid curve for $\psi = 1$; dashed for $\psi = 4$; dot-dashed for $\psi = 1/4$.

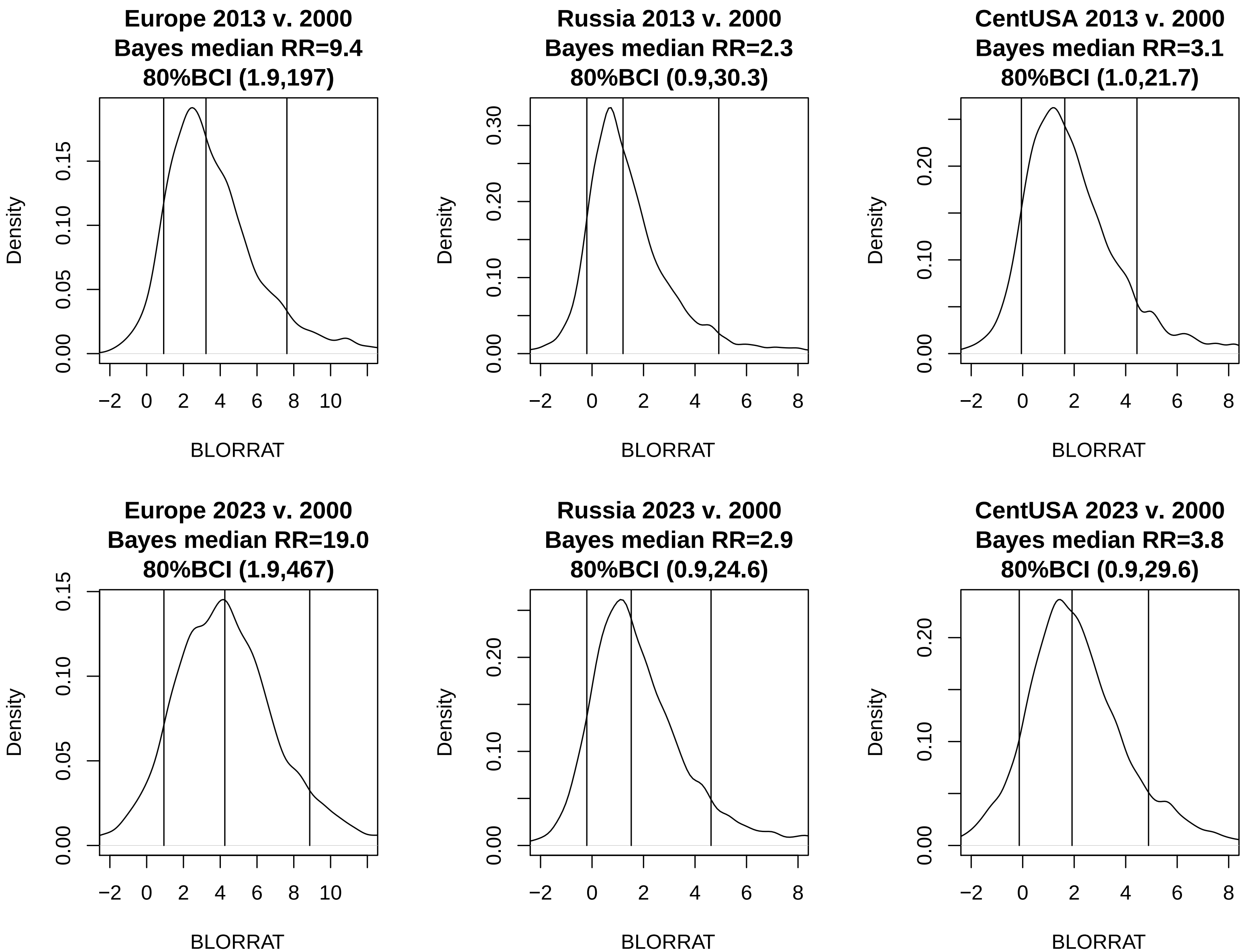
Posterior Quantiles of the Risk Ratio

Series	ψ	Percentile of Posterior Density				
		5	25	50	75	95
Europe 2003	0.25	0.11	0.68	2.06	6.88	47.40
Europe 2003	1	0.06	0.49	1.78	7.03	77.69
Europe 2003	4	0.06	0.53	1.85	6.84	85.61
Europe 2012	0.25	0.17	0.64	1.54	3.66	13.71
Europe 2012	1	0.29	1.03	2.66	6.66	34.86
Europe 2012	4	0.80	2.91	7.14	20.46	159.48
Russia 2010	0.25	0.14	0.73	2.07	6.02	36.62
Russia 2010	1	0.16	0.95	2.89	9.00	68.17
Russia 2010	4	0.08	0.76	2.75	10.90	186.06
Central USA 2011	0.25	0.25	0.74	1.53	3.23	10.42
Central USA 2011	1	0.31	1.00	2.25	5.21	19.24
Central USA 2011	4	0.45	1.75	4.37	10.98	55.46

We also calculate boxplots of the projected extreme event probabilities up to 2040, and risk ratios based on those (from 20C+A2 model runs)



Boxplots of the posterior distribution of exceedance probability for 2000–2040, for three regions and $K = 6$ or 10 DF in the spline representation of trend.



Posterior Densities for the Risk Ratio at Different Time Points

IX Conclusions

- For comparing extreme event probabilities for anthropogenic *versus* control conditions in climate models, we typically find estimated risk ratios of about 2, but with very wide credible intervals

- These calculations remain problematic given the difficulty of estimating extreme event probabilities as well as the scale mismatch problem

- However, projections of extreme event probabilities into the future show notable increases, *especially for Europe*

- We can also compare the risk ratios for current or future event probabilities versus past probabilities (2000 used as reference), and find much stronger evidence for change (but still very wide BCIs)