EXTREME VALUE THEORY AND SINGLE-EVENT ATTRIBUTION

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

Statistical and Applied Mathematical Sciences Institute

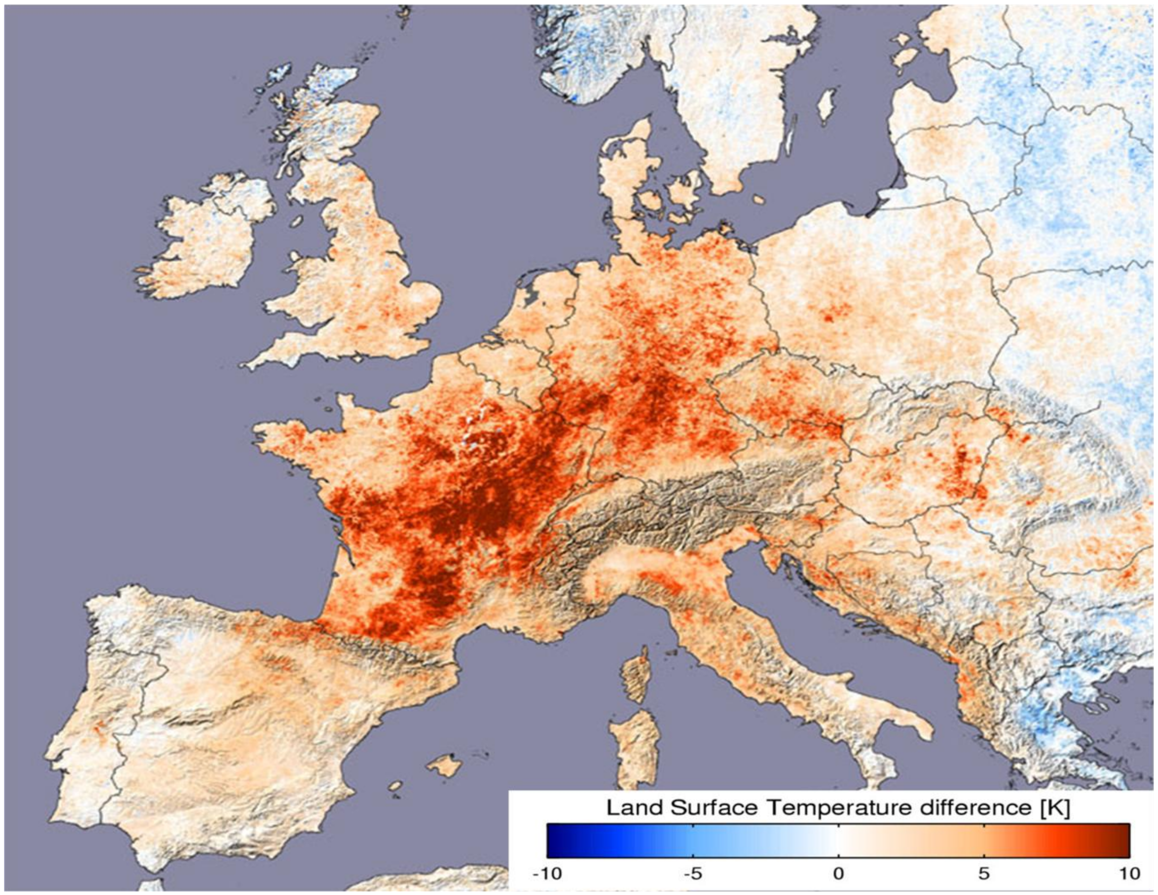
Joint Work with Michael Wehner
WCRP Open Science Conference
Denver, October 24 2011
rls@email.unc.edu

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

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

As an example, we consider the European heatwave of 2003.



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

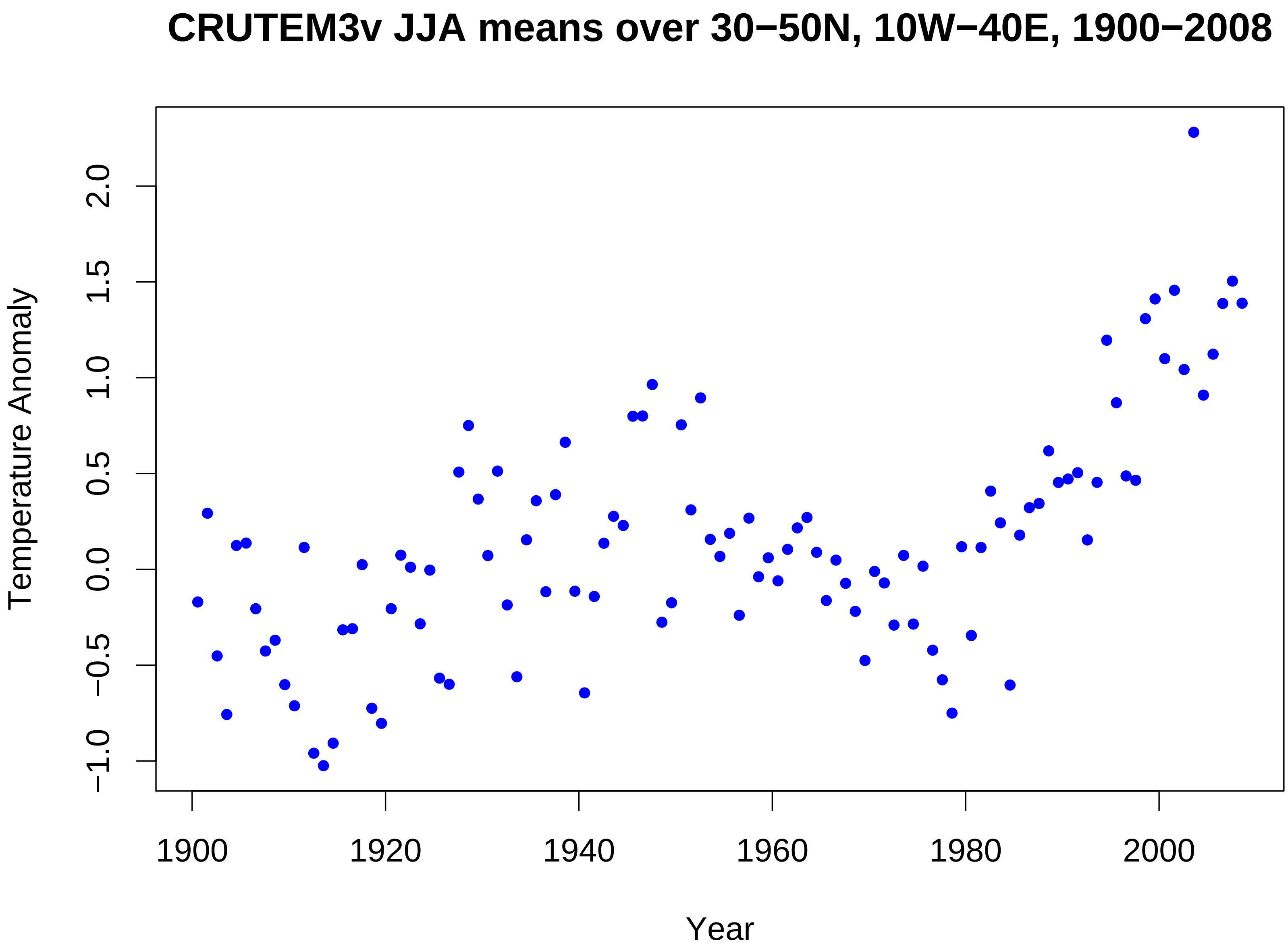
From NASA's MODIS - Moderate Resolution Imaging Spectrometer, courtesy of Reto Stöckli, ETHZ

Analysis of 2003 Heatwave

Stott, Stone and Allen (*Nature*, 2004) reduced the problem to a calculation of JJA land temperature annual averages over the region 30^{o}N to 50^{o}N, 10^{o}W to 40^{o}E.

They used 4 runs of the HadCM3 climate model, under both anthropogenic and natural (solar, volcanic only) forcings.

We have repeated this exercise but using anthropogenic and control run forcings from the public AR4 database at PCMDI.



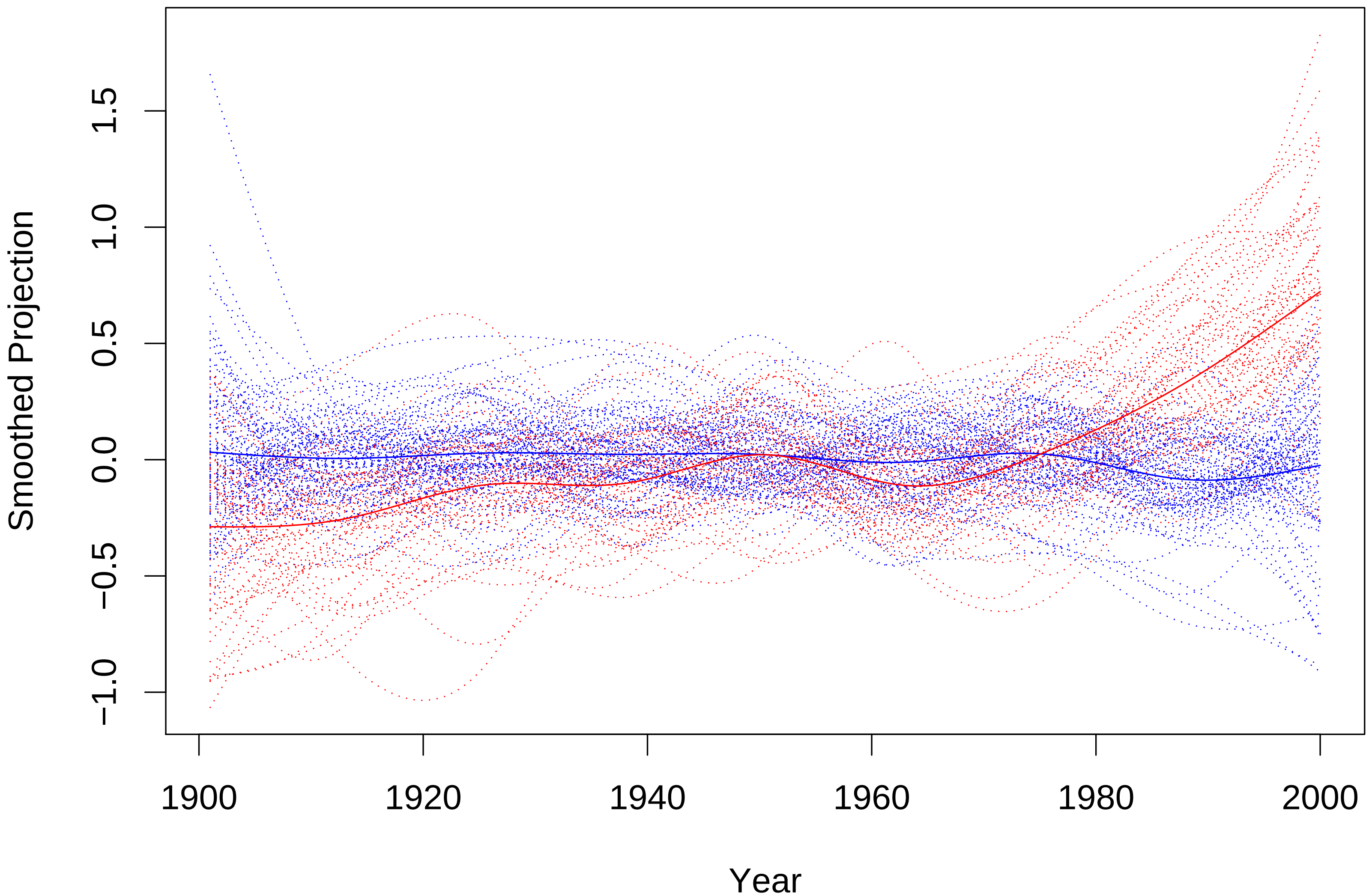
Observed anomalies for JJA mean temperatures for 1900–2008

Model	Number of Series	Final Year	Model	Number of Series	Final Year
ccsm3.0	7	1999	cgcm3.1_t47	5	2000
cgcm3.1_t63	1	unknown	cnrm	1	1999
csiro	1	2000	gfdl2.0	1	2000
gfdl2.1	1	2000	giss_aom	1	2000
giss_eh	5	1999	giss_er	5	2003[1]
hadcm3	1	1999	hadcrut2v	1	unknown
hadgem1	1	1999	iap_fgoals1_0_g	3	1999
inmcm3	1	2000	ipsl	1	2209[2]
miroc_hires	1	2000	miroc_medres	3	2000
miub_echo	5	2000	mpi_echam5	3	2050 or 2100[3]
mri_cgcm2_3_2a	5	2000	pcm	4	1999

Table 1. Models used for the twentieth-century analysis, together with the number of series from each model.

Model	Number of Series	Model	Number of Series
bccr_bcm2_0	2	ccsm3.0	5
cgcm3.1_t47	10	cgcm3.1_t63	3
cnrm	5	csiro_mk3_0	3
echam5	5	gfdl2.0	5
gfdl2.1	5	giss_aom	4
giss_eh	4	giss_er	5
hadcm3	3	iap	9
inmcm3	3	ipsl	5
miroc_hires	1	miroc_medres	5
miub_echo	3	pcm	10

Table 2. Models used for the control run analysis, together with the number of series from each model.



Smoothed JJA mean temperature anomalies based on 95 twentieth century model runs (red) and 57 control model runs (blue), and means over all model runs.

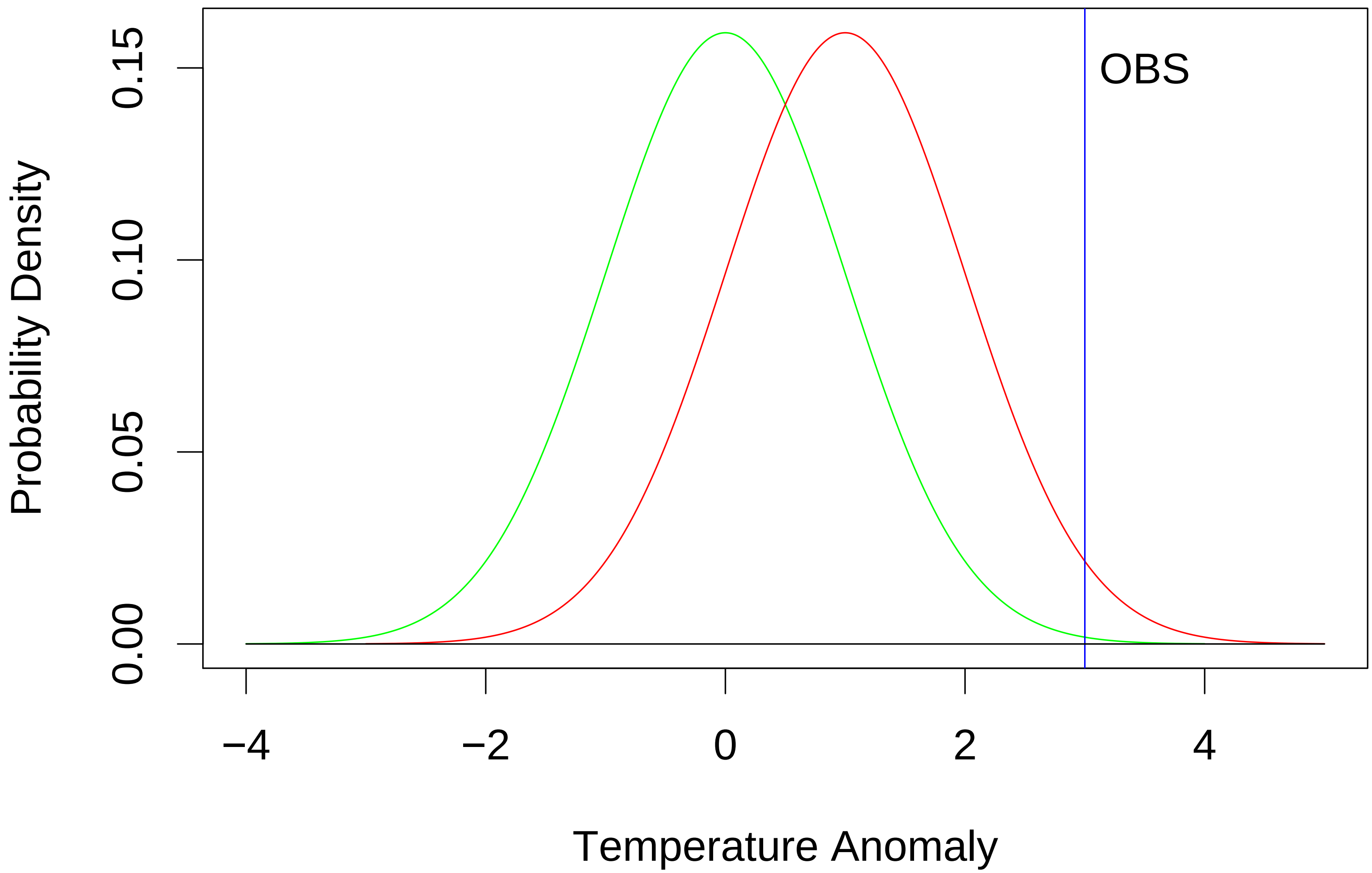
Statistical Methods

- Calculations based on normal distributions — not usually advisable when extreme values are of interest

- Nonparametric methods — avoids making unjustified distributional assumptions, but cannot be extrapolated beyong range of observed data

- Methods based on extreme value theory (GEV, GPD) (The method proposed here)

- Other families of probability distributions that include long-tailed cases? (A possible alternative)



Assumes normal probabilities; same variance, different means

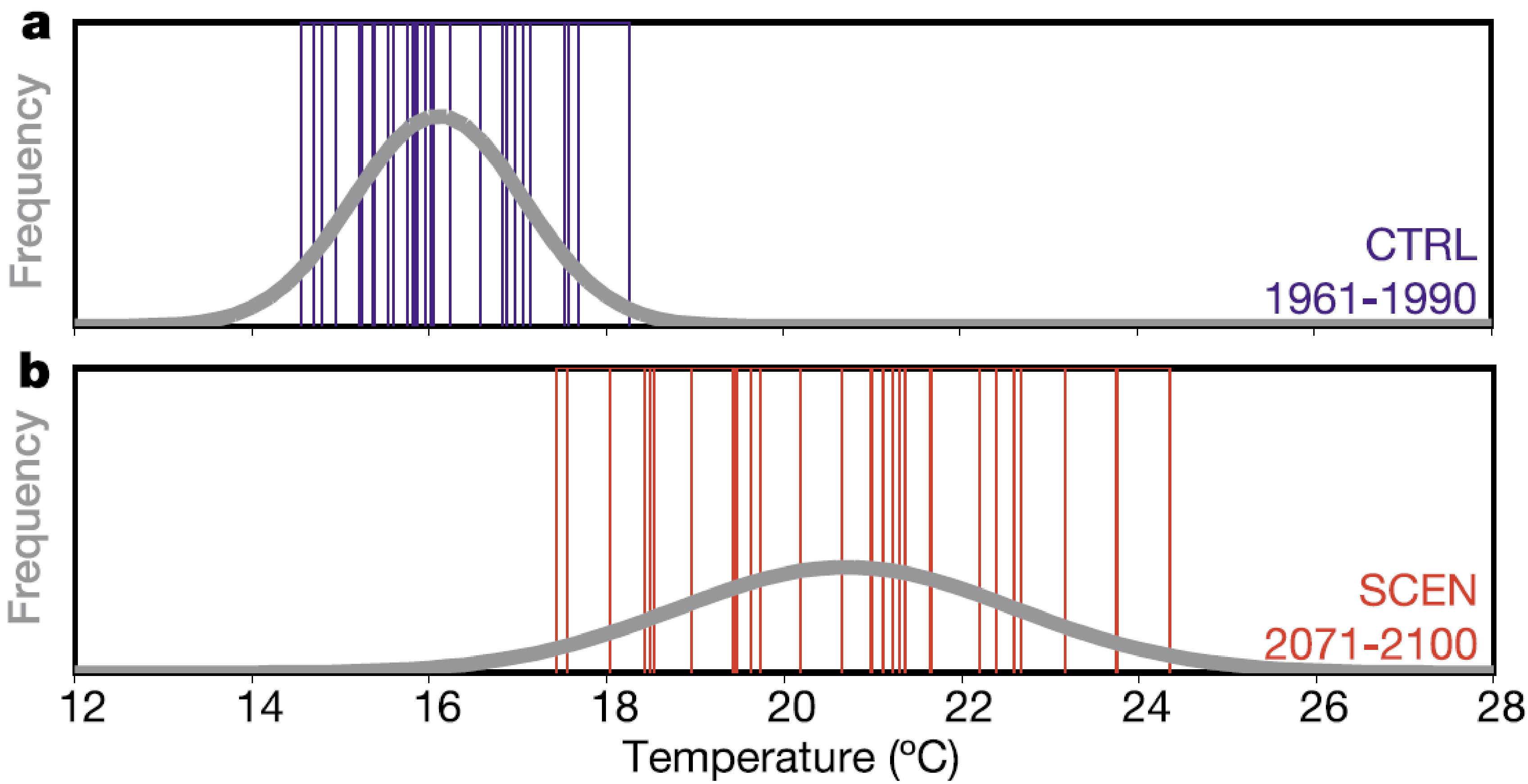
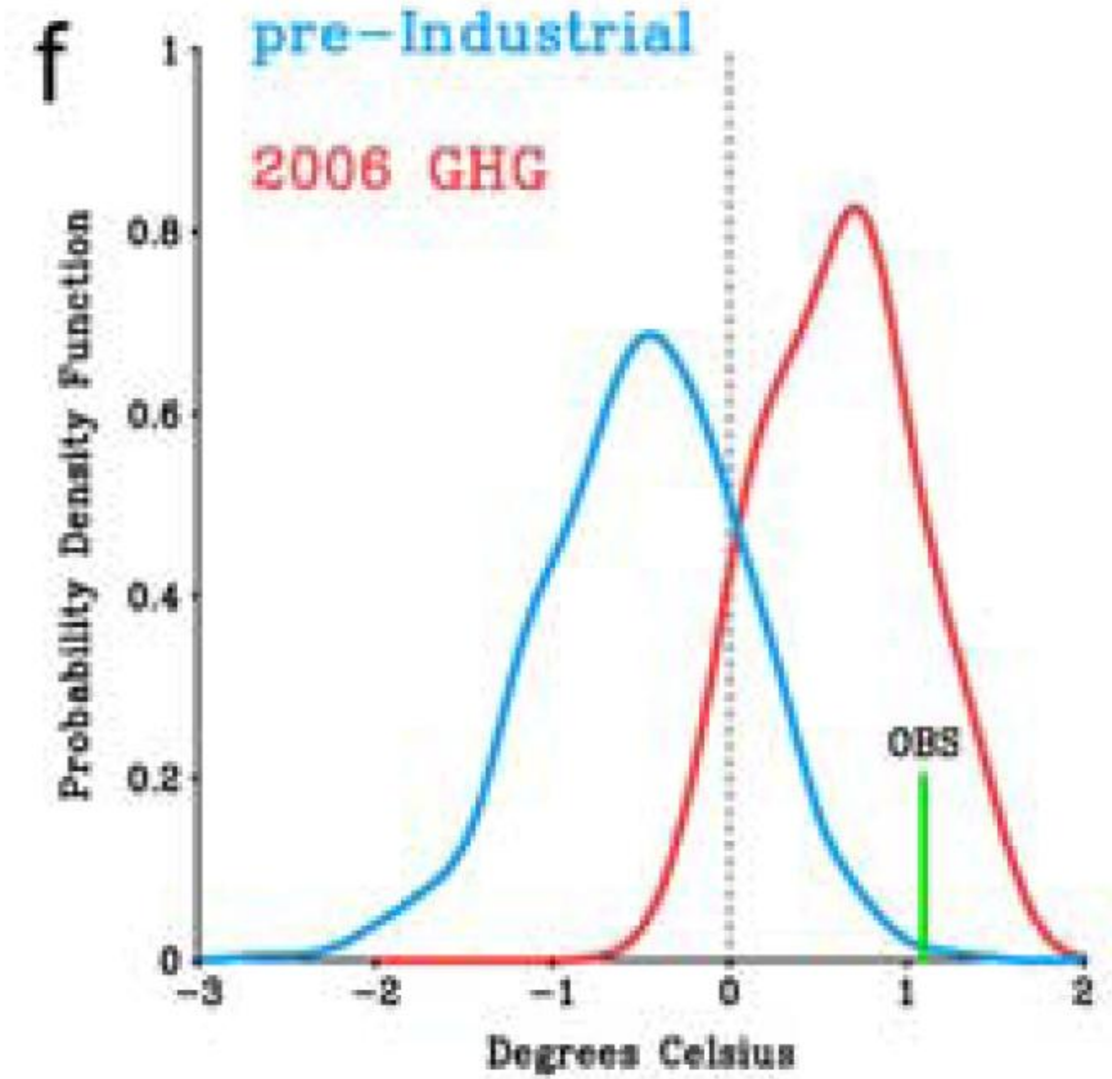


Figure 3 Results from an RCM climate change scenario representing current (CTRL 1961–90) and future (SCEN 2071–2100) conditions. **a**, **b**, Statistical distribution of summer temperatures at a grid point in northern Switzerland for CTRL and SCEN,

(Schär et al., Nature, 2004)



(Hoerling et al., GRL, 2007)

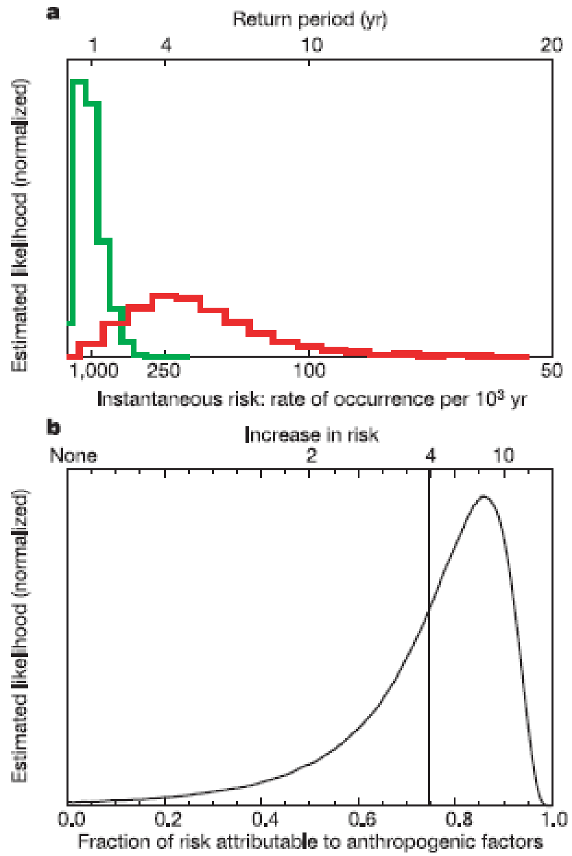


Figure 4 Change in risk of mean European summer temperatures exceeding the 1.6 K threshold. **a**, Histograms of instantaneous return periods under late-twentieth-century conditions in the absence of anthropogenic climate change (green line) and with anthropogenic climate change (red line). **b**, Fraction attributable risk (FAR). Also shown, as the vertical line, is the 'best estimate' FAR, the mean risk attributable to anthropogenic factors averaged over the distribution.

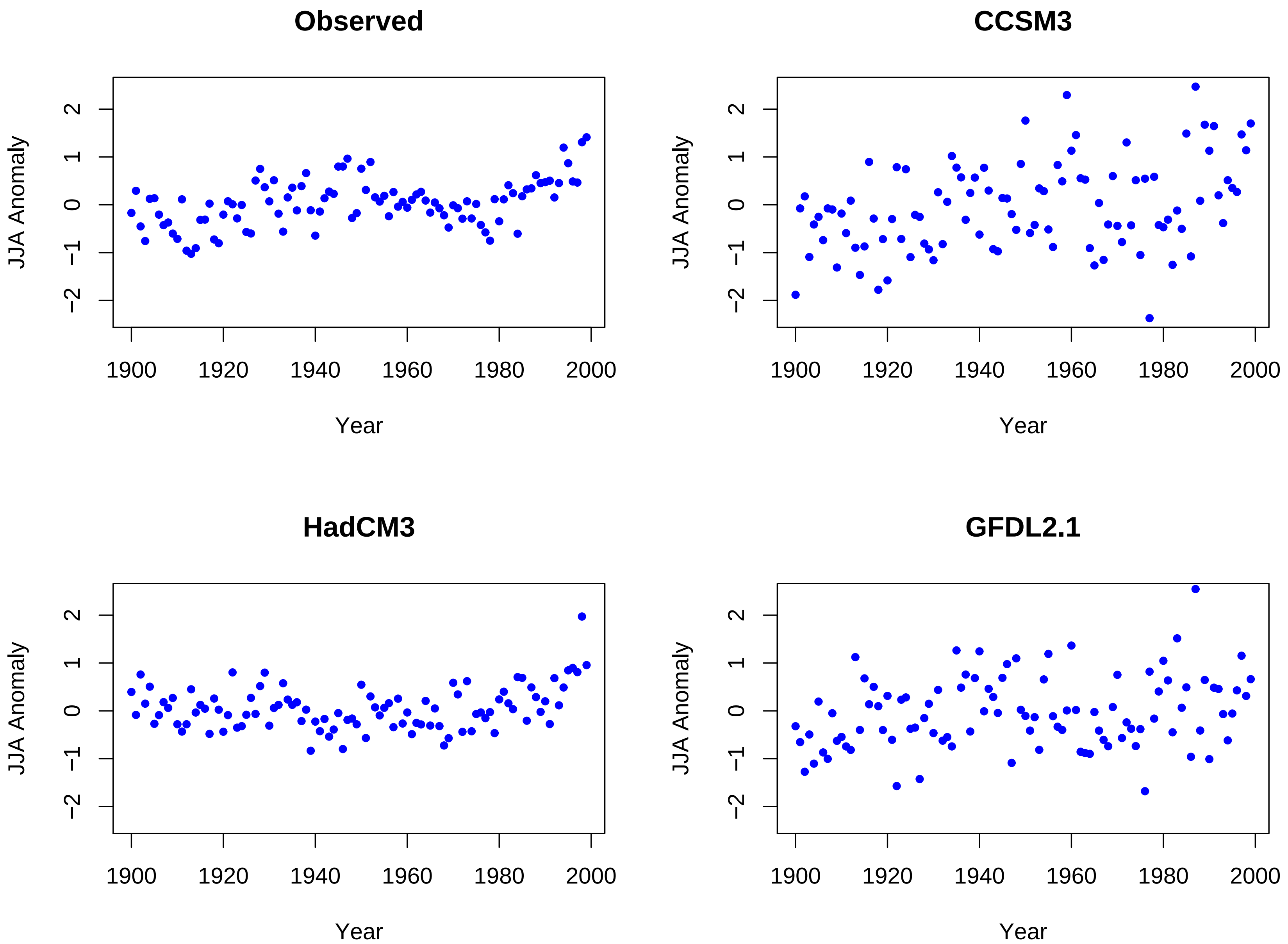
ıblishing Group NATURE | VOL 432 | 2 DECEMBER 2004 | www.nature.com/nature

Stott, Stone and Allen (2004)

Stott, Stone and Allen used the generalized Pareto distribution (GPD) for exceedances over a high threshold to estimate extreme value tail probabilities.

They didn't use the observed 2003 anomaly of 2.3K but instead 1.6K, justified as being close to the largest observed anomaly up to 1999.

However they also include a (conventional) detection and attribution step; appears to be necessary because of the *scale mismatch* problem between observations and model data.



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

The Challenge:

Find a statistically coherent approach that allows for the non-normal nature of extreme tail probabilities, that also takes into account the scale mismatch problem.

Proposed Statistical Model

Y_t: Real-data observation in year t ($1 \le t \le T$)

$Z_{m,i,t}$: Model data for model m ($1 \le m \le M$), run i ($1 \le i \le n_m$) and year t ($1 \le t \le T$)

Assume:

$$
\begin{aligned}
Y_t &= \mu_t + \epsilon_t, \\
\frac{Z_{m,i,t} - B_m}{A_m} &= \mu_t + \eta_{m,i,t}, \\
\mu_t &= \sum_{j=0}^{q} \beta_j x_{t,j}, \\
\epsilon_t,\ \eta_{m,i,t} &\sim F, \\
F(\epsilon) &\approx \exp\left\{ -\left(1 + \frac{\xi\epsilon}{\psi}\right)_+^{-1/\xi} \right\} \text{ as } F(\epsilon) \to 1.
\end{aligned}
$$

Statistical Model Parameters

$\beta_0, ..., \beta_q$ are regression parameters for spline-based trend

A_m and B_m are scale/location parameters for model m; reflect bias and scale mismatch

Common error cdf F for both the model and observational errors

Common GEV tail behavior of F; scale and shape parameters ψ and ξ

An extension: allow ψ and ξ parameters to be different for models and observations

Probability of Exceeding a Design Value

Estimate probability that some future value Y^*, for which the GEV parameters are μ^*, ψ^*, ξ^*, exceeds a design value u^*.

Assume $\psi^* = \psi$, $\xi^* = \xi$ and $\mu^* = \sum_{j=0}^{q} \beta_j x_j^*$ for given values of x_j^*, $0 \le j \le q$, usually defined so that $x_j^* = x_{T,j}$, the covariates that correspond to the final year T. The true probability is then

$$p^* = 1 - \exp\left\{ - \left(1 + \xi^* \frac{u^* - \mu^*}{\psi^*} \right)_+^{-1/\xi^*} \right\}$$

Estimation of p^*: Maximize the likelihood over all parameters

Profile Likelihood Approach: Maximize likelihood constrained to a fixed value of p^*. This is used to derive confidence intervals.

Example Model Fits

u^*	Threshold	q	P_1	P_0	FAR
1.6	85%	4	0.03461	0.00419	0.87909
1.6	85%	6	0.03103	0.00429	0.86161
1.6	80%	4	0.04283	0.00382	0.91083
1.6	80%	6	0.03806	0.00385	0.89891
1.6	75%	4	0.03329	0.00392	0.88236
1.6	75%	6	0.01982	0.00423	0.78659
1.6	70%	4	0.02254	0.00381	0.83099
1.6	70%	6	0.01588	0.00435	0.72583
2.3	85%	4	0.00024	0.00038	-0.53739
2.3	85%	6	0.00013	0.00023	-0.86897
2.3	80%	4	0.00333	0.00025	0.92631
2.3	80%	6	0.00294	0.00031	0.89444
2.3	75%	4	0.00159	0.00020	0.87264
2.3	75%	6	0.00092	0.00040	0.56705
2.3	70%	4	0.00117	0.00022	0.81168
2.3	70%	6	0.00068	0.00045	0.34080

Unfortunately, confidence intervals associated with these esitmated exceedance probabilities are very wide.

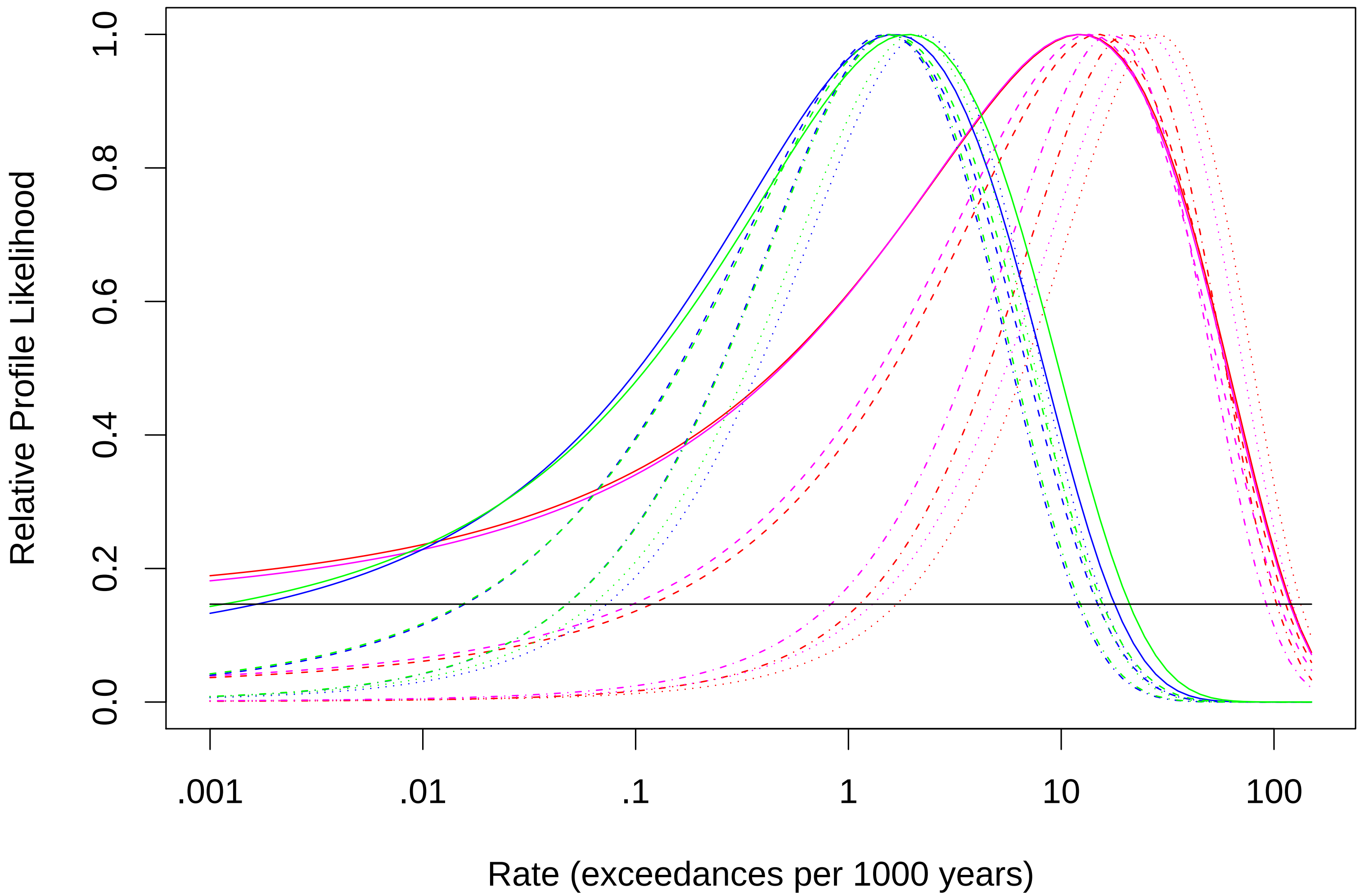
Example: for the model with 80% threshold, 4 degrees of freedom for the spline representation of the trend —

$u^* = 1.6$, estimated $p^* = .042$

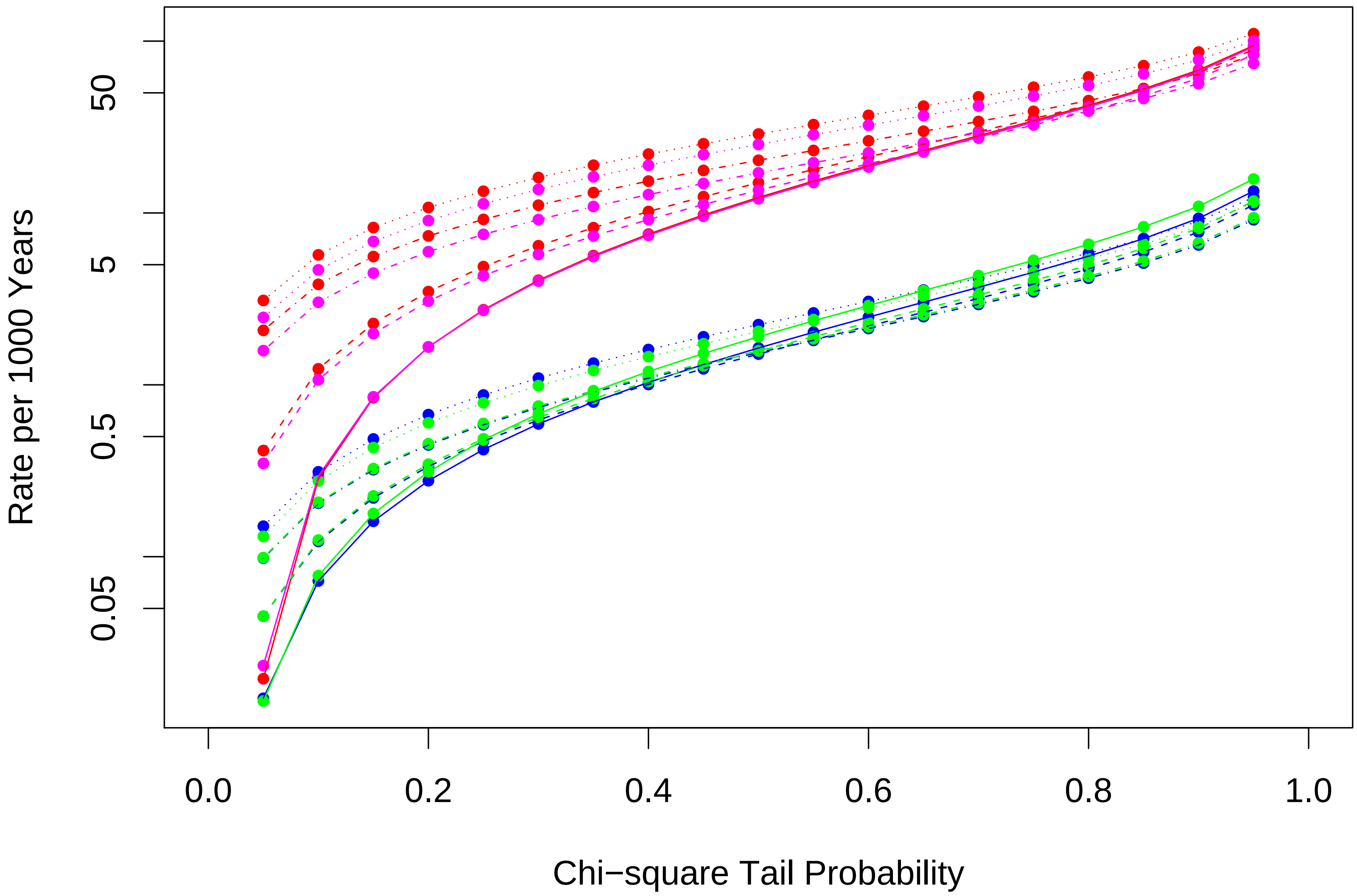
95% confidence interval from 0.006 to 0.2

$u^* = 2.3$, estimated $p^* = .0033$

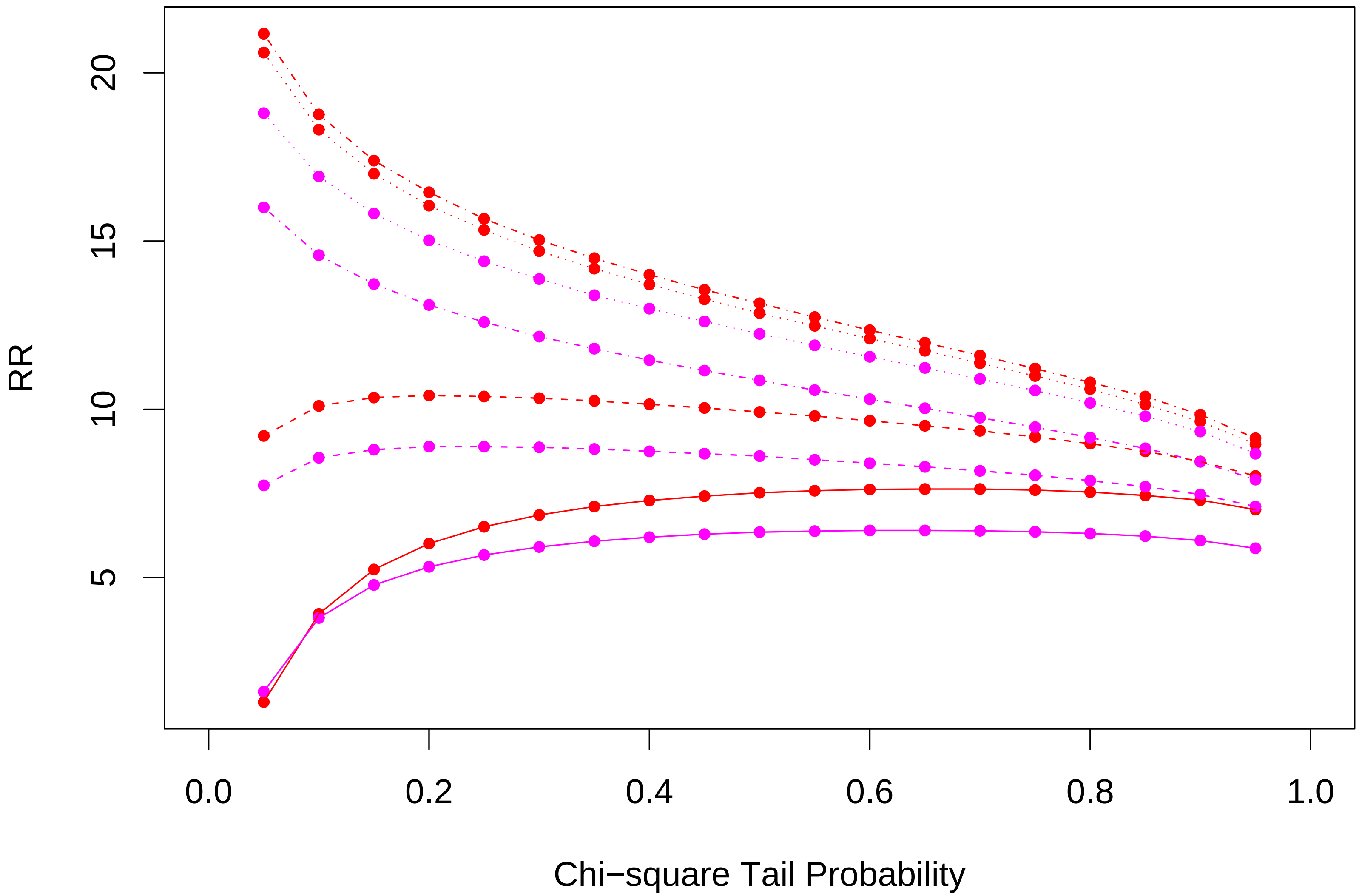
95% confidence interval from $< 10^{-6}$ to 0.036



Profile likelihoods corresponding to a design value of 1.6K (red/magenta curves correspond to anthropogenic forcing, blue/green curves correspond to control model runs)



Rate per 1000 years of exceeding the design value 1.6K, as a function of the chi-squared probability value.



Relative risks for 20th century versus control run scenarios at 1.6K, for each value of the chi-squared tail probability.

SUMMARY

- The results confirm there is a high FAR associated with the anthropogenic effect (of the order of 0.9 in several estimates)

- However, confidence intervals for P_0 and P_1 are very wide

- We still don't have a satisfactory method of computing a confidence interval for the FAR, that correctly takes account of all the unknown parameters