Discussion of meeting on Statistical Aspects of Climate Change

Richard Smith,

University of North Carolina, USA

rls@email.unc.edu

Presented remotely at the Royal Statistical Society annual conference

Aberdeen, UK, September 13, 2022

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

Discussion of Clarkson et al.

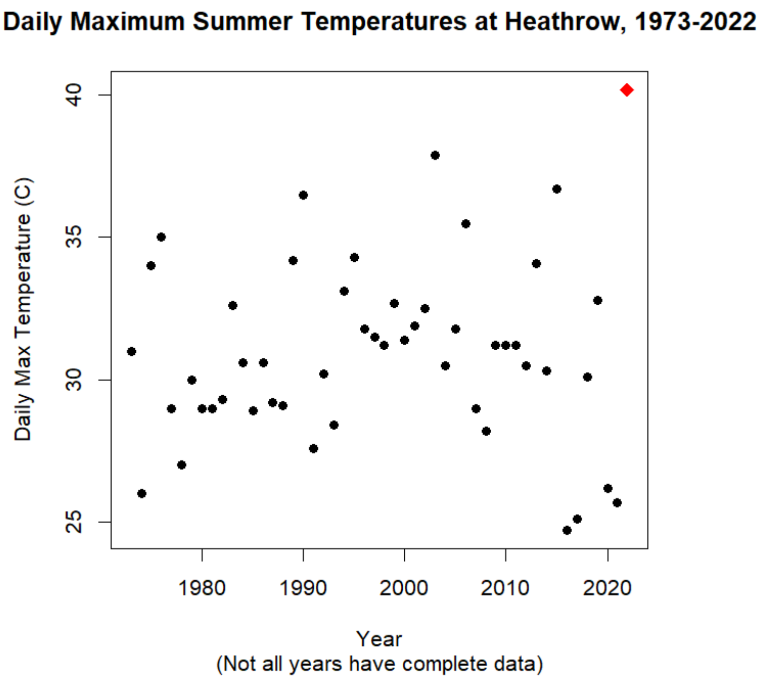
- This paper has proposed some alternative models for climate extremes, extending the classical GEV/GPD paradigm
- The proposed approach based on random effects has many parallels with Bayesian statistics
- A particular focus is the “zero probability problem”: in extreme value problems with $\xi<0$, there is the possibility that some observed event may be predicted to have probability zero
- This issue has been known for a long time, but has been given particular focus in recent papers of the “World Weather Attribution” group (WWA; Philip et al 2021, Zachariah et al 2022).
- The authors’ discussion of this is very thorough, but one simpler approach is to calculate the Bayes posterior mean instead of the MLE; many authors have argued this gives a more realistic assessment of very low event probabilities (Coles and Powell 1996)
- Here, I want to illustrate these issues in the broader context of integrating extreme value analysis with climate models
- As an example, I discuss the UK heatwave of July 2022

Questions

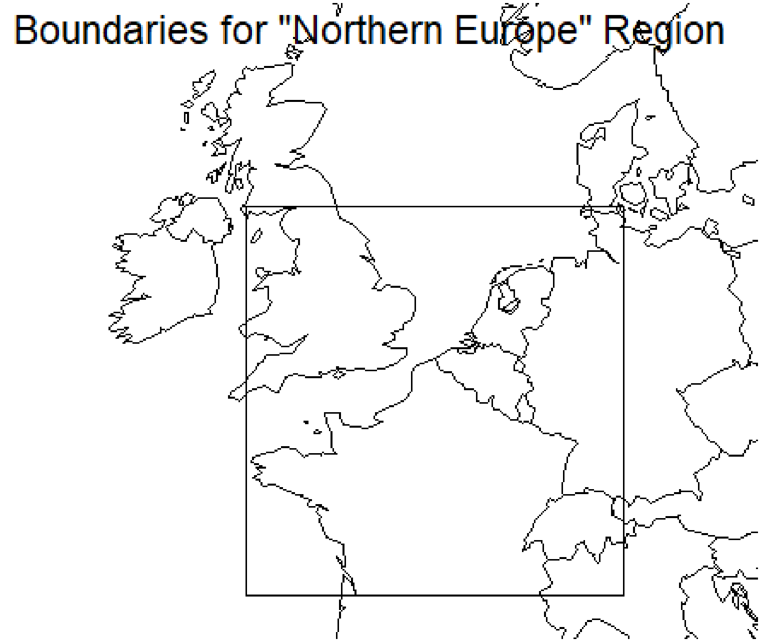
- On July 19, 2022, Heathrow Airport, London, recorded a daily high temperature of 40.2 C (104.4 F)
- Based on data available prior to 2022, what is the probability of a temperature in London exceeding 40.2C?
  - In 1950?
  - In 2022?
  - In 2100?

Approach

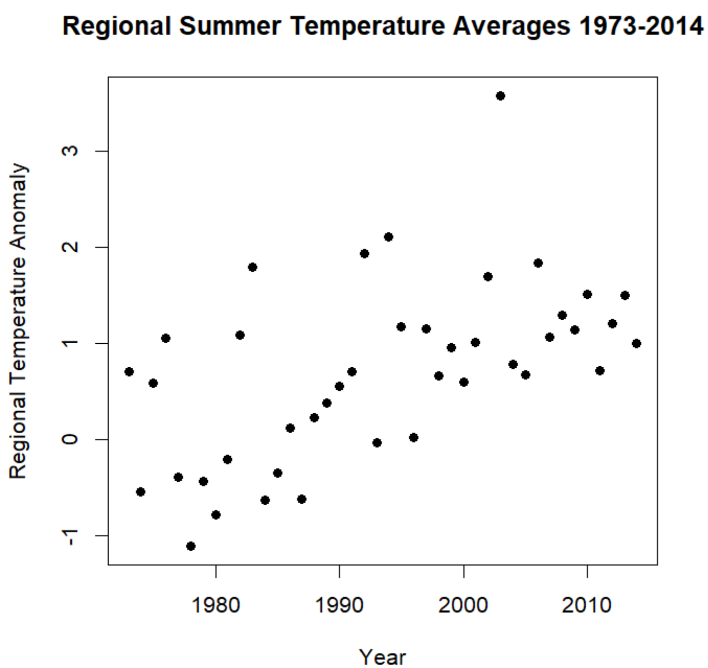
- Downloaded data from Heathrow back to the 1973; calculated annual temperature maxima (where available)
- Computed summer average temperatures across a region of southern UK and nearby parts of Europe (regional summer means)
- Downloaded data from 19 climate models (CMIP6) and calculated the equivalent model-based quantities



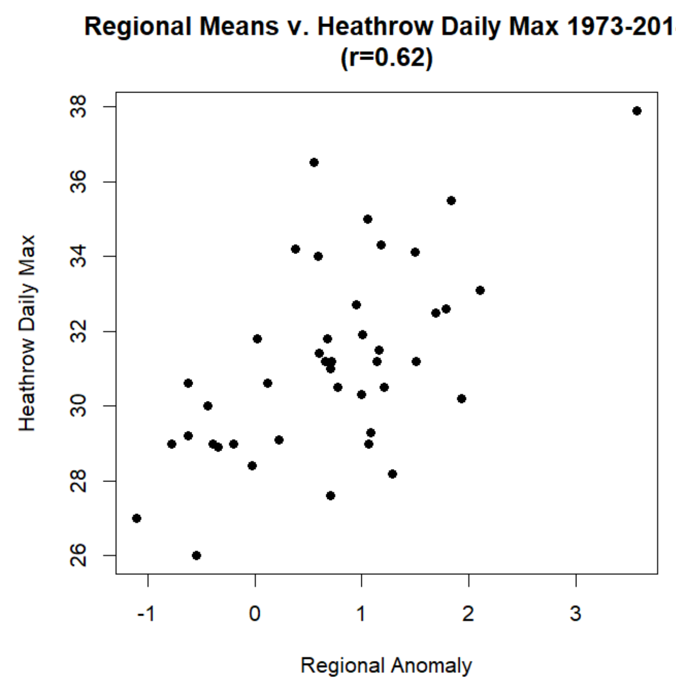
Daily Maximum Summer Temperatures at Heathrow, 1973-2022
Daily Max Temperature (C)
25
30
35
40
1980
1990
2000
2010
2020
Year
(Not all years have complete data)



Boundaries for "Northern Europe" Region



Regional Summer Temperature Averages 1973-2014
Regional Temperature Anomaly
-1
0
1
2
3
1980
1990
2000
2010
Year

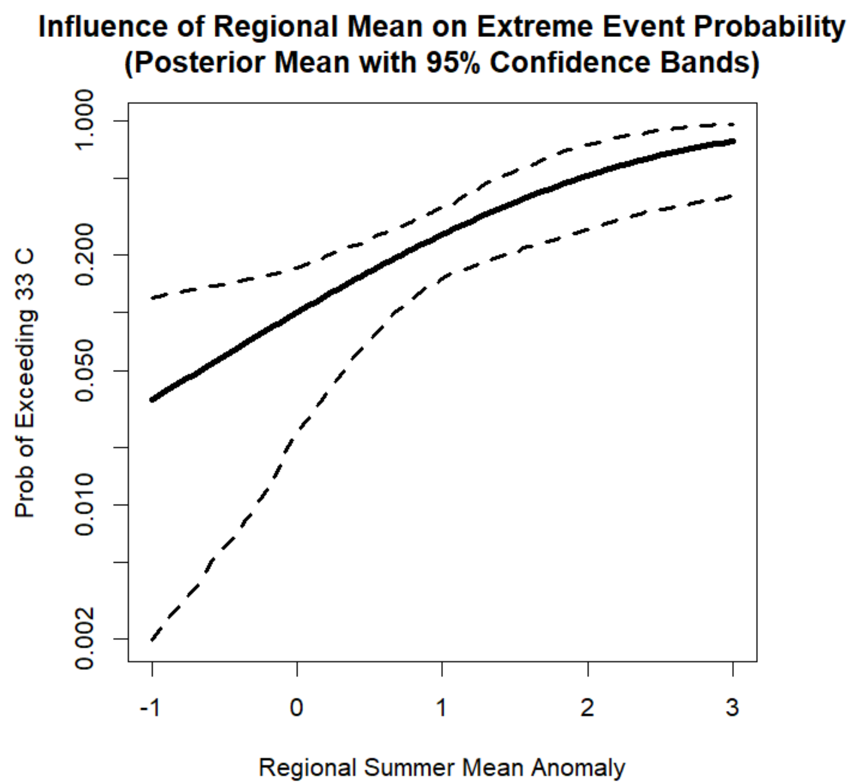


Regional Means v. Heathrow Daily Max 1973-2014
(r=0.62)
Heathrow Daily Max
26
28
30
32
34
36
38
-1
0
1
2
3
Regional Anomaly

Generalized Extreme Value (GEV) Analysis

- Let Y_t be annual max at Heathrow, Z_t regional summer anomaly, year t
- $\Pr\{Y_t \le y \mid Z_t\} = \exp\left\{-\left(1+\xi\frac{y-\beta_0-\beta_1 Z_t}{\sigma}\right)^{-1/\xi}\right\}$ so long as $(\cdot) > 0$
- MLE:

Parameter	Estimate	S.E.	t-value	p-value
β_0	29.2282	0.3889	75.1638	0
β_1	1.6239	0.3432	4.7315	2.2×10^{-6}
$\log \sigma$	0.6192	0.1176	5.2644	1.4×10^{-7}
ξ	-0.1751	0.0918	-1.9076	0.0564



Combining Observational and Model-Based Regional Averages

- Z_t observed regional averages 1850–2021 (from University of East Anglia)

- $M_{t,j},\ j = 1, 17$ results of 17 climate models 1850–2100 (from CMIP6); one of three scenarios for 2015–2100

  - ssp585: no systematic reduction in greenhouse gas emissions

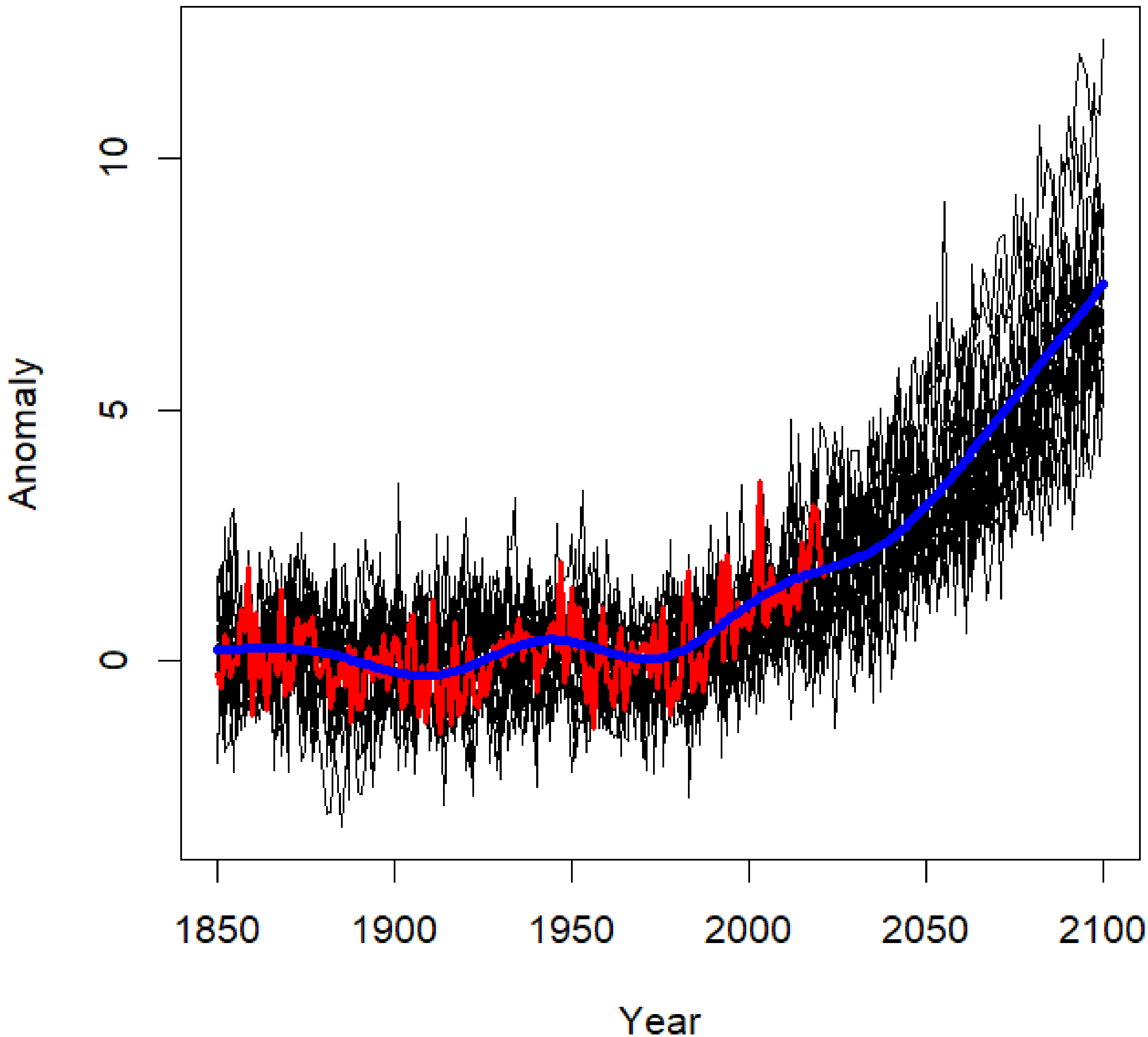
  - ssp126: sharp reductions in GHGs: temperatures will increase for a while and then level off

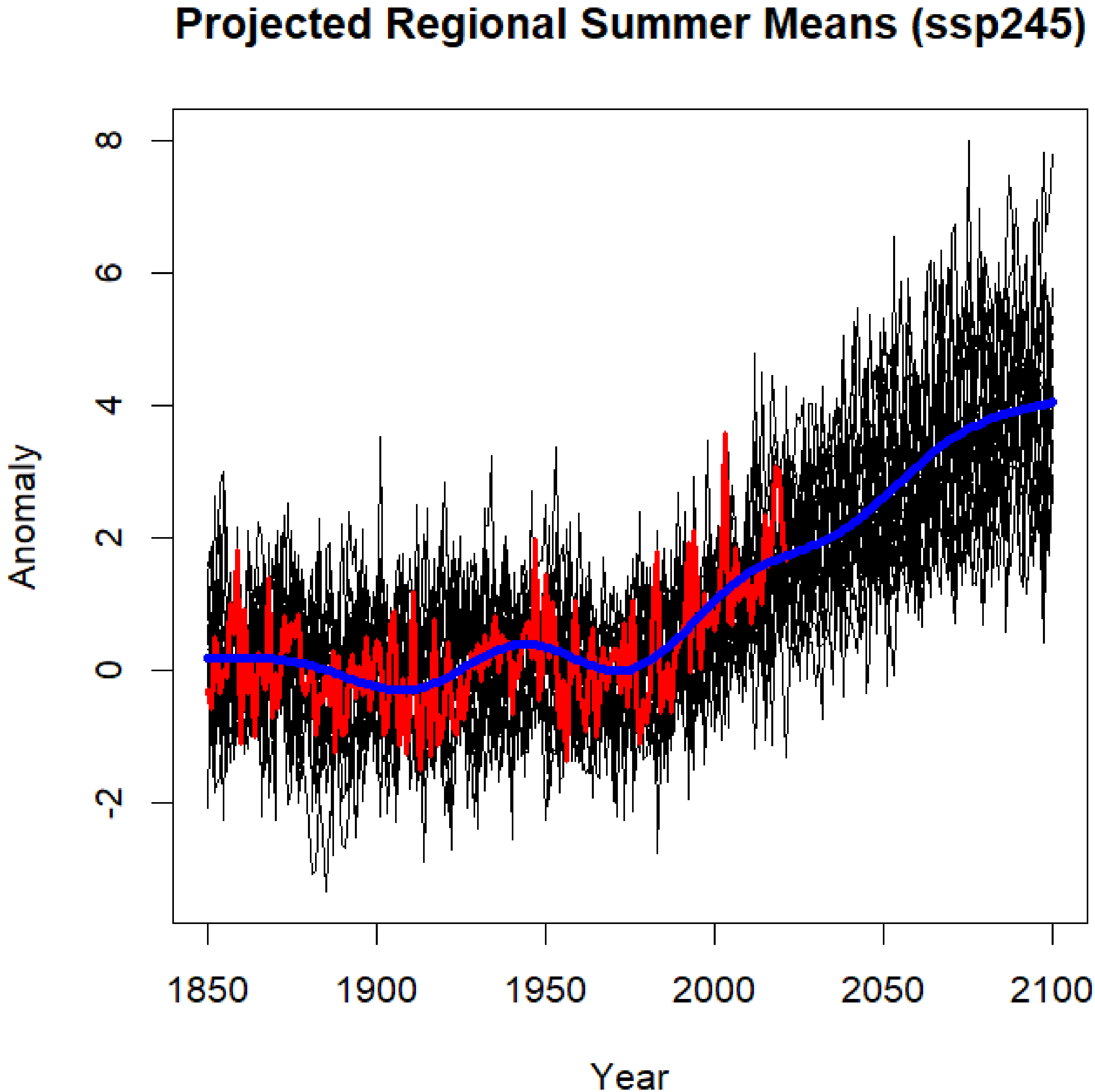
  - ssp245: intermediate, may be achievable with a concerted effort of many nations

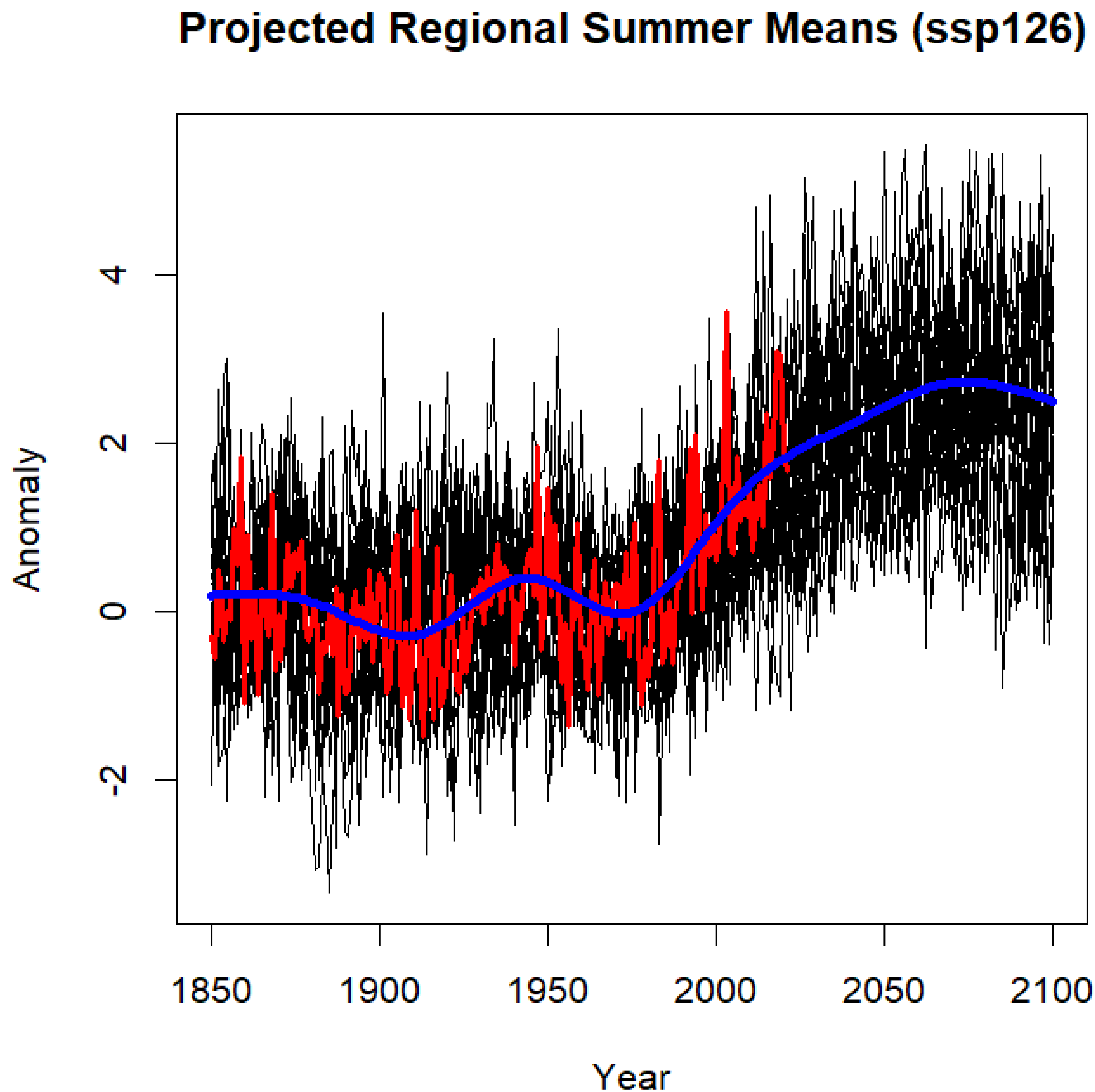
- Statistical analysis assumes a parametric μ_t function (splines) and that two PCs from the climate model and Z_t are each randomly distributed about μ_t with different variances (estimated)

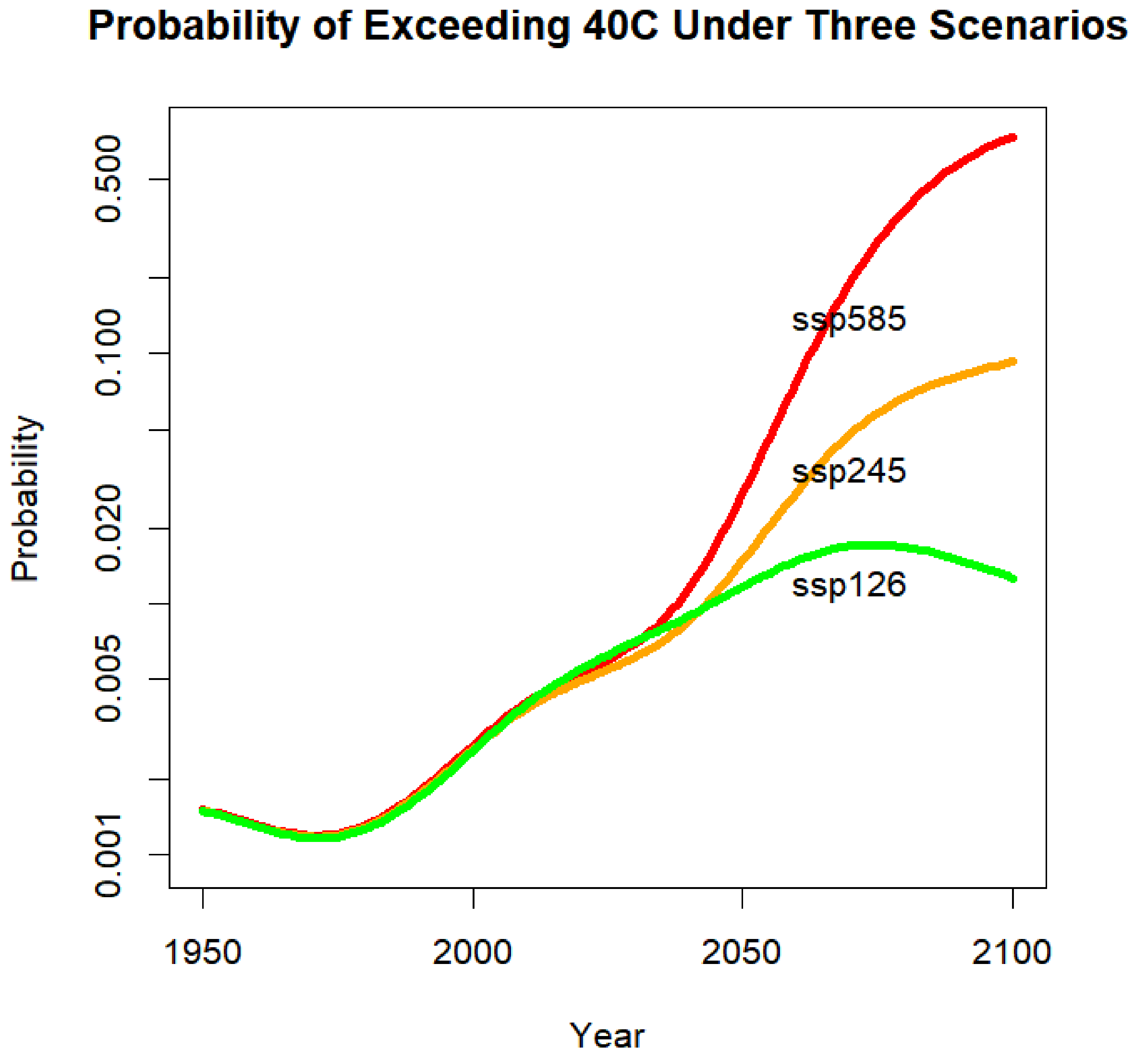
- In deference to the possibility of long-tailedness in Z_t, the model for that variable assumed t-distributed errors (though there is no evidence to contradict normality)

Projected Regional Summer Means (ssp585)




Projected Regional Summer Means (ssp245)
Anomaly
-2
0
2
4
6
8
1850
1900
1950
2000
2050
2100
Year


Projected Regional Summer Means (ssp126)
Anomaly
-2
0
2
4
1850
1900
1950
2000
2050
2100
Year



Probability of Exceeding 40C Under Three Scenarios
Probability
0.001
0.005
0.020
0.100
0.500
1950
2000
2050
2100
Year
ssp585
ssp245
ssp126

Conclusions and Further Remarks

- Figures like this show that extreme events are going to become more frequent under any scenario, but especially for scenarios in which emissions are unrestricted or only lightly regulated
- From a statistical point of view, I believe these analyses can be improved by incorporating more extensive data
- To complement the random effects model used by the authors, there is also a long literature on using similar models in a spatial context (Casson and Coles 1999,….,Russell, Risser, Smith and Kunkel 2020)
- Thank you for the opportunity to contribute this discussion.