A Conditional Approach to Extreme Event Attribution

Richard Smith,

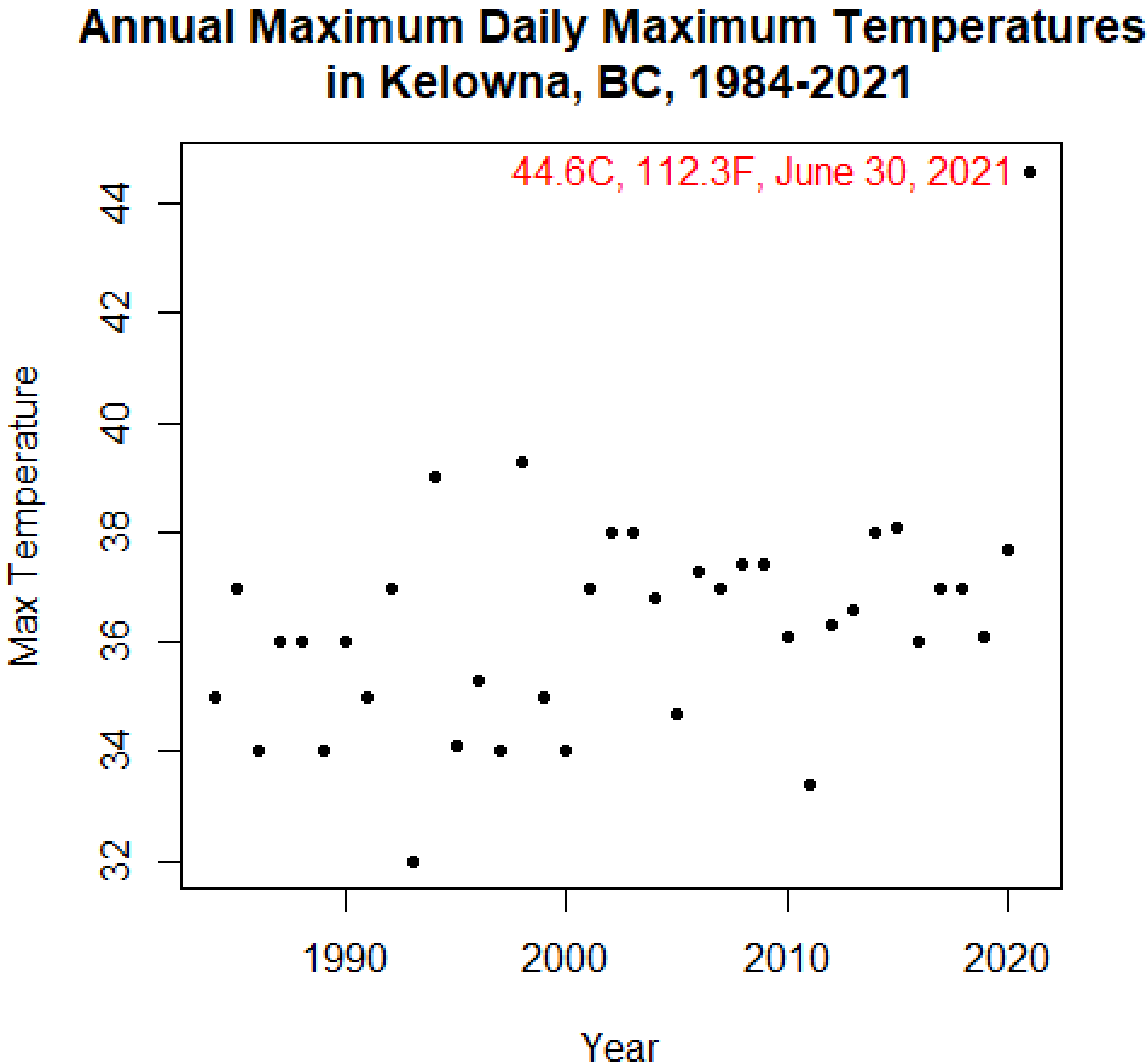
University of North Carolina, USA

rls@email.unc.edu

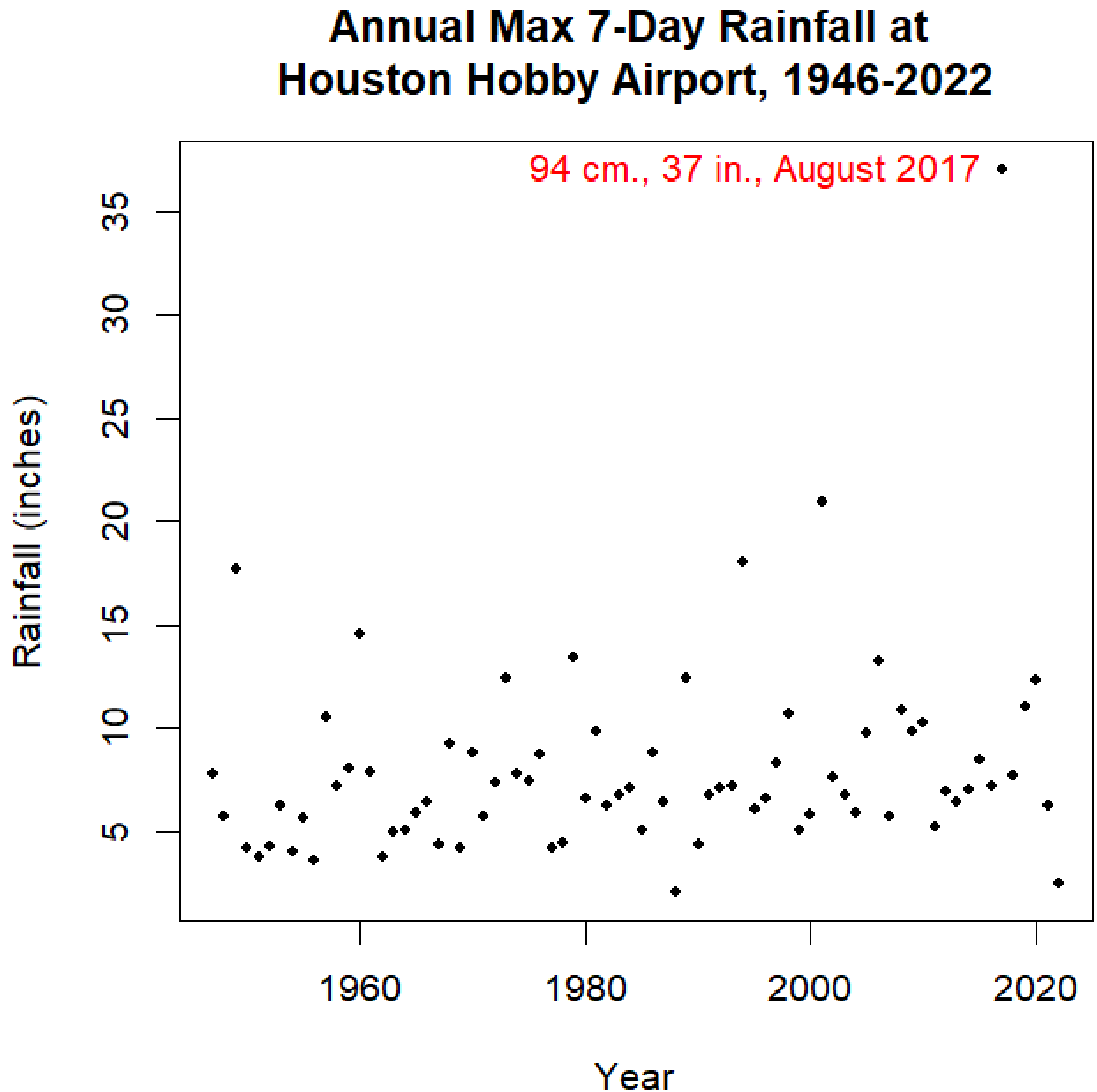
IMSI Workshop on Detection and Attribution of Climate Change

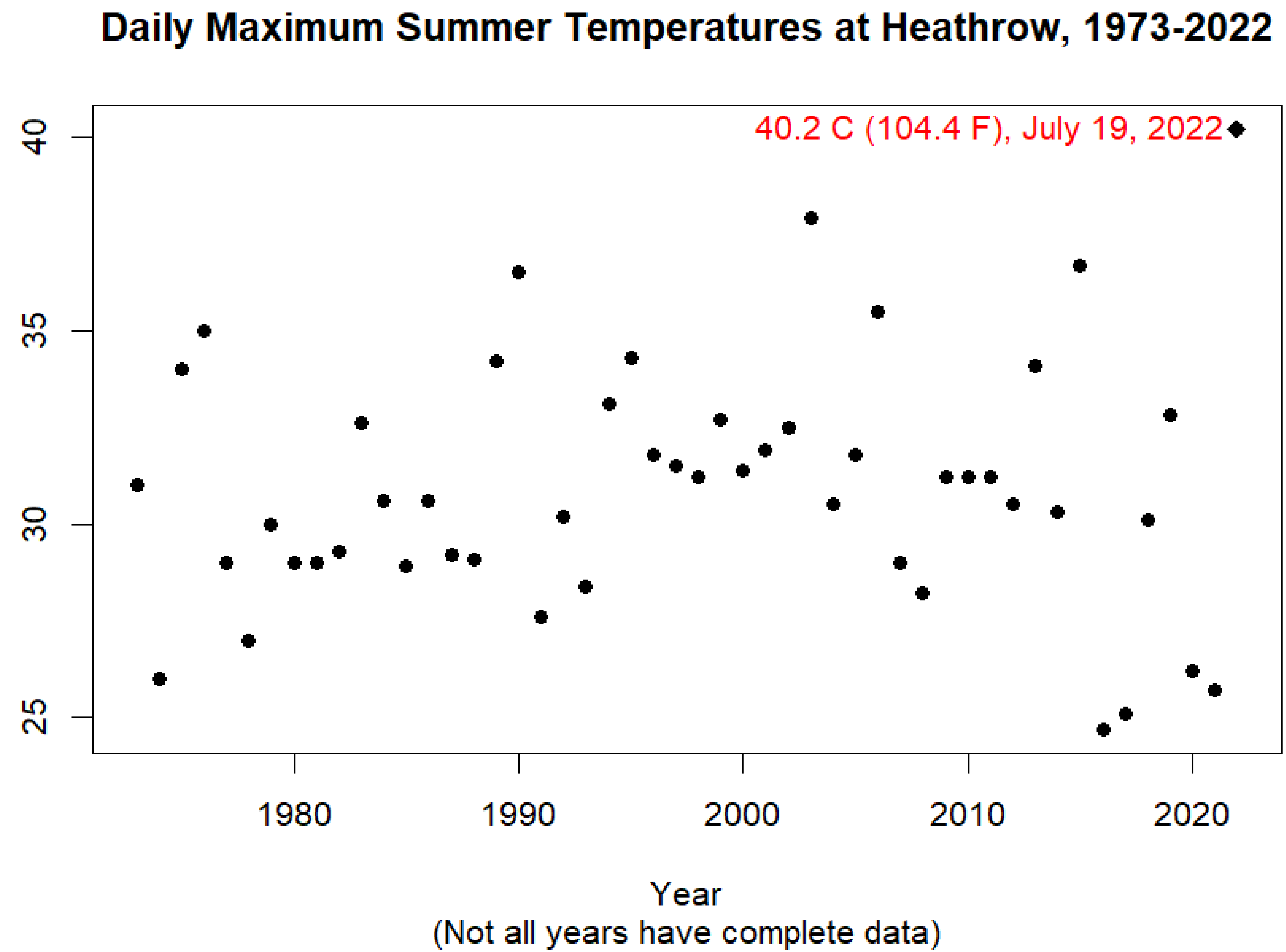
University of Chicago, October 18, 2022

THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



Annual Maximum Daily Maximum Temperatures
in Kelowna, BC, 1984-2021
44.6C, 112.3F, June 30, 2021
Max Temperature
32
34
36
38
40
42
44
1990
2000
2010
2020
Year


Annual Max 7-Day Rainfall at
Houston Hobby Airport, 1946-2022
94 cm., 37 in., August 2017
Rainfall (inches)
5
10
15
20
25
30
35
1960
1980
2000
2020
Year



Daily Maximum Summer Temperatures at Heathrow, 1973-2022
40.2 C (104.4 F), July 19, 2022
40
35
30
25
1980
1990
2000
2010
2020
Year
(Not all years have complete data)

Questions for Each Example

- How extreme was this event?
  - Based on data available prior to the event, what was the probability of its occurrence?
- To what extent was it “attributable” to human influence?
- How will the probability of this event change in the future under various climate scenarios (ssp126, ssp245, ssp585)?

Background on Extreme Event Attribution

(Hammerling, Katzfuss, Smith in *Handbook of Environmental and Ecological Statistics*, 2019)

- Fraction of Attributable Risk (Stott, Stone and Allen 2004)

$$FAR = 1 - \frac{p_0}{p_1}$$

where p_1 is factual probability of observed event, p_0 a counter-factual

$$\text{Equivalently}: \quad RR = \frac{p_1}{p_0}.$$

- Probabilities of Necessary and Sufficient Causation (Hannart, Pearl, Otto and Ghil, 2016)

$$\begin{aligned} PN &= \max\left\{1 - \frac{p_0}{p_1}, 0\right\}, \\ PS &= \max\left\{1 - \frac{1-p_0}{1-p_1}, 0\right\}, \\ PNS &= \max\{p_1 - p_0, 0\}. \end{aligned}$$

- Other "causal inference" ideas (e.g. Ebert-Uphoff & Deng 2012, Hannart & Naveau 2018)

Other “Framing” Issues

(National Academy of Sciences, *Attribution of Extreme Weather Events in the Context of Climate Change*, 2016)

- Choice of Weather Event
- Changes in frequency or changes in magnitude?
- Conditioning?
- Selection Bias?

Statistical Methods

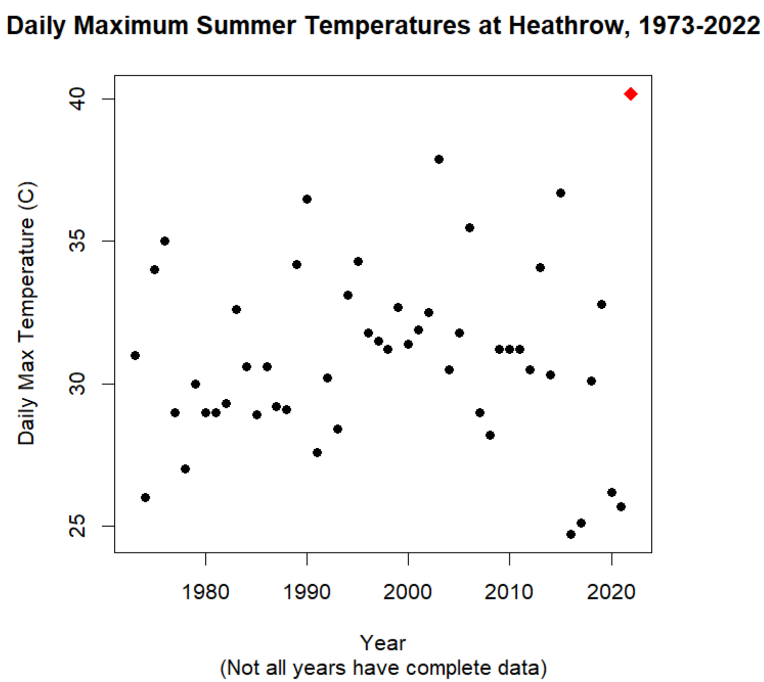
- Methods based on the normal distribution (not recommended)
- Methods adapting conventional detection and attribution methodology (e.g. Min, Zhang, Zwiers and Hegerl, 2011)
- Methods based on counting exceedances in long climate model runs (e.g. Pall et al. 2011, Emanuel 2017)
- Methods based on extreme value theory to assess probabilities of individual extreme events (Stott et al. 2004, ..., Risser and Wehner 2017; van Oldenborgh et al. (2021) review of methodology for World Weather Attribution project)

So, What's New?

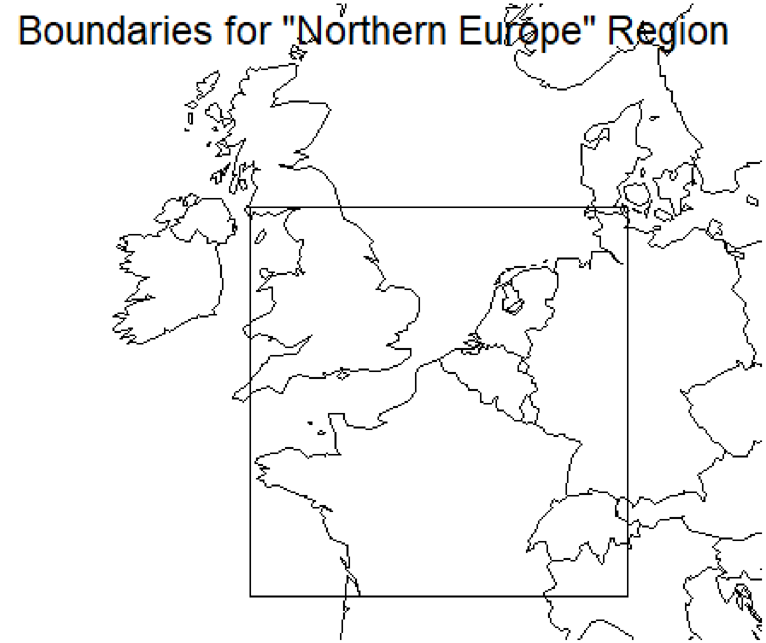
- Recent papers by WWA group (Philip et al on 2021 Pacific Northwest heatwave, Zachariah et al on 2022 UK heatwave) have drawn renewed attention to the "zero probability" problem, or its converse
- Spatial smoothing methods improve estimation of GEV parameters (Russell et al 2020; papers by Risser, L. Zhang, Shaby, Huser, Wadsworth…)
- Other methods based on threshold exceedances and/or random parameters (Clarkson et al, Applied Statistics, to appear)
- Here, I want to draw attention to what I call the *conditioning approach* – make the extreme value distributions dependent on a climate variable that can be well simulated by a climate model

Approach

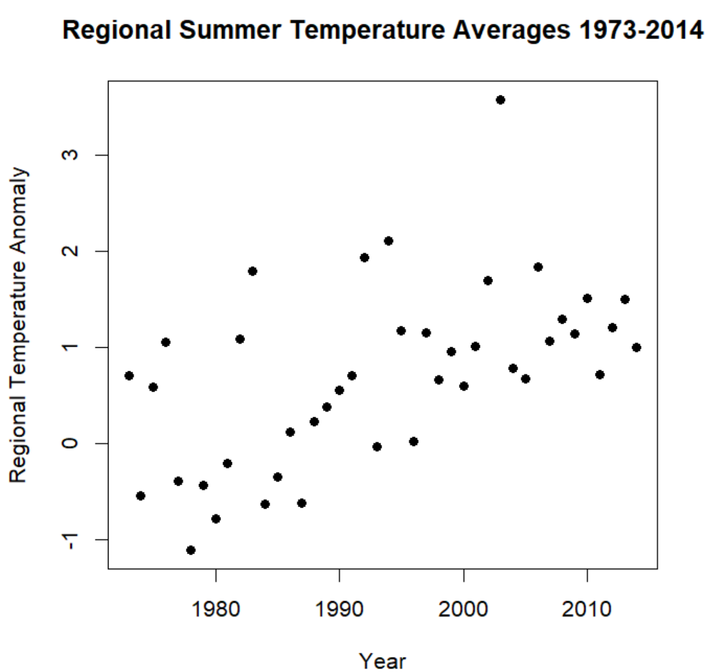
- Downloaded data from Heathrow back to the 1973; calculated annual temperature maxima (where available)
- Computed summer average temperatures across a region of southern UK and nearby parts of Europe (regional summer means)
- Downloaded data from 19 climate models (CMIP6) and calculated the equivalent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



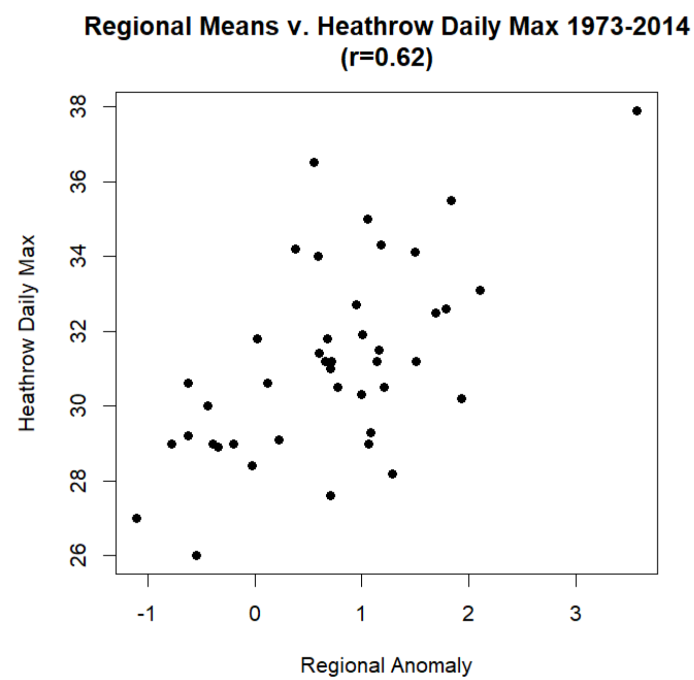
Daily Maximum Summer Temperatures at Heathrow, 1973-2022
Daily Max Temperature (C)
25
30
35
40
1980
1990
2000
2010
2020
Year
(Not all years have complete data)



Boundaries for "Northern Europe" Region



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

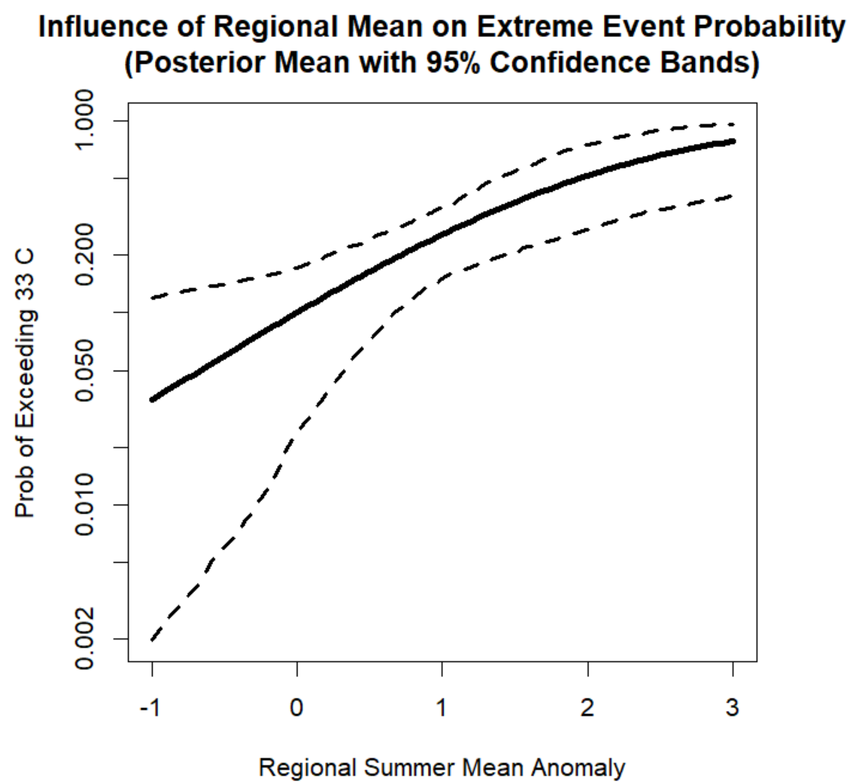


Regional Means v. Heathrow Daily Max 1973-2014
(r=0.62)
Heathrow Daily Max
26
28
30
32
34
36
38
-1
0
1
2
3
Regional Anomaly

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

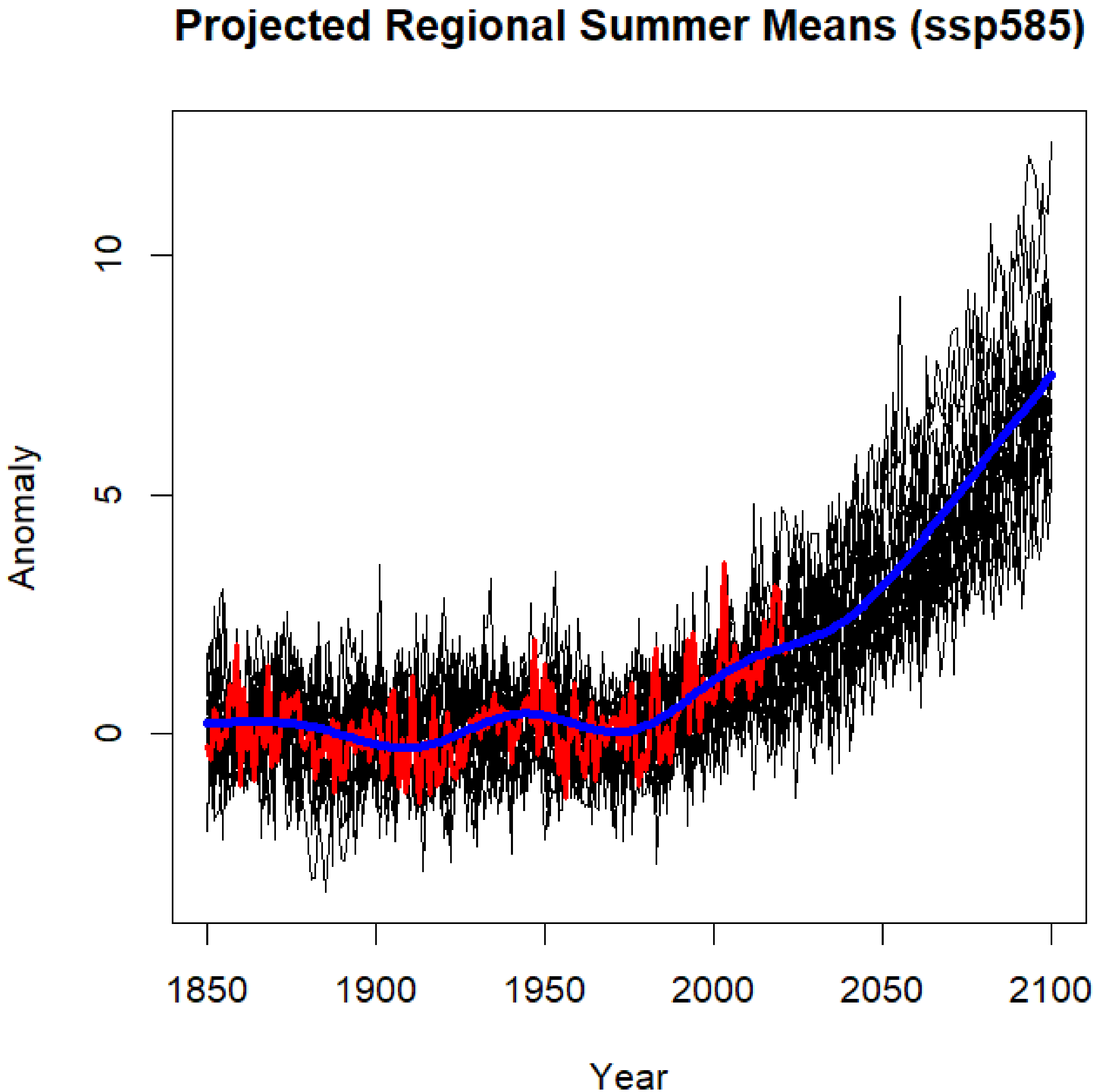
  - ssp585: no systematic reduction in greenhouse gas emissions

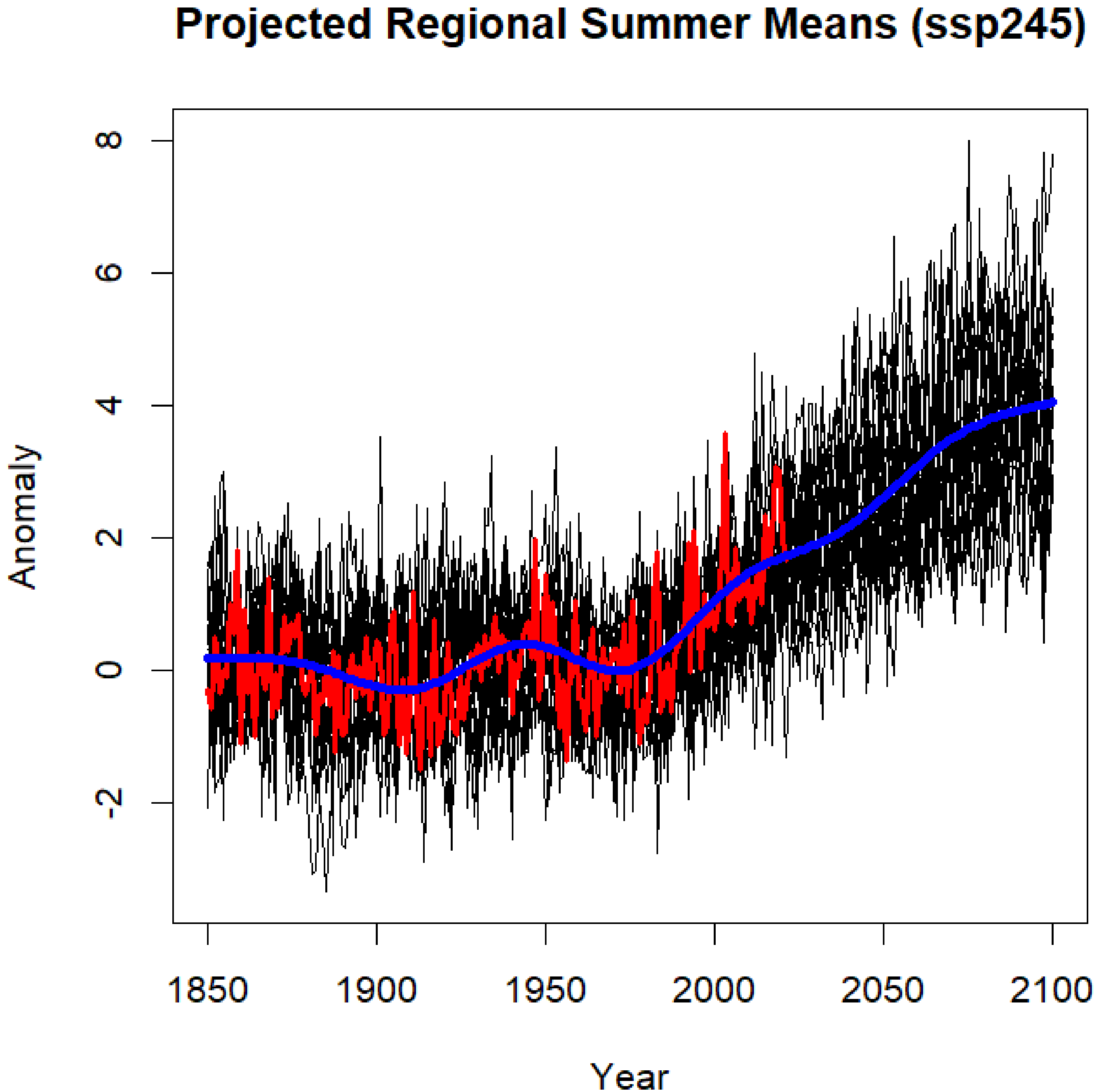
  - ssp126: sharp reductions in GHGs: temperatures will increase for a while and then level off

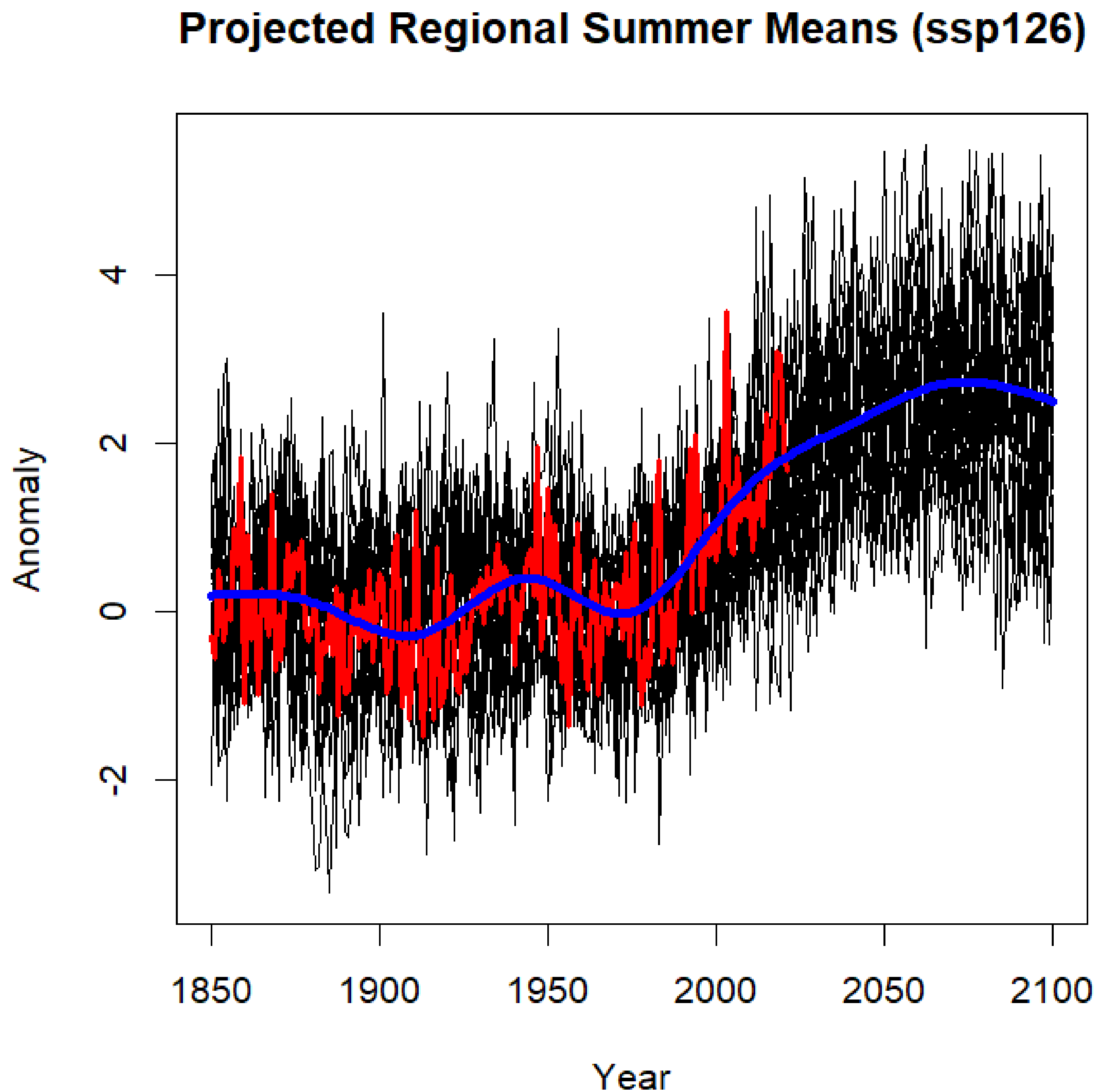
  - ssp245: intermediate, may be achievable with a concerted effort of many nations

- 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)

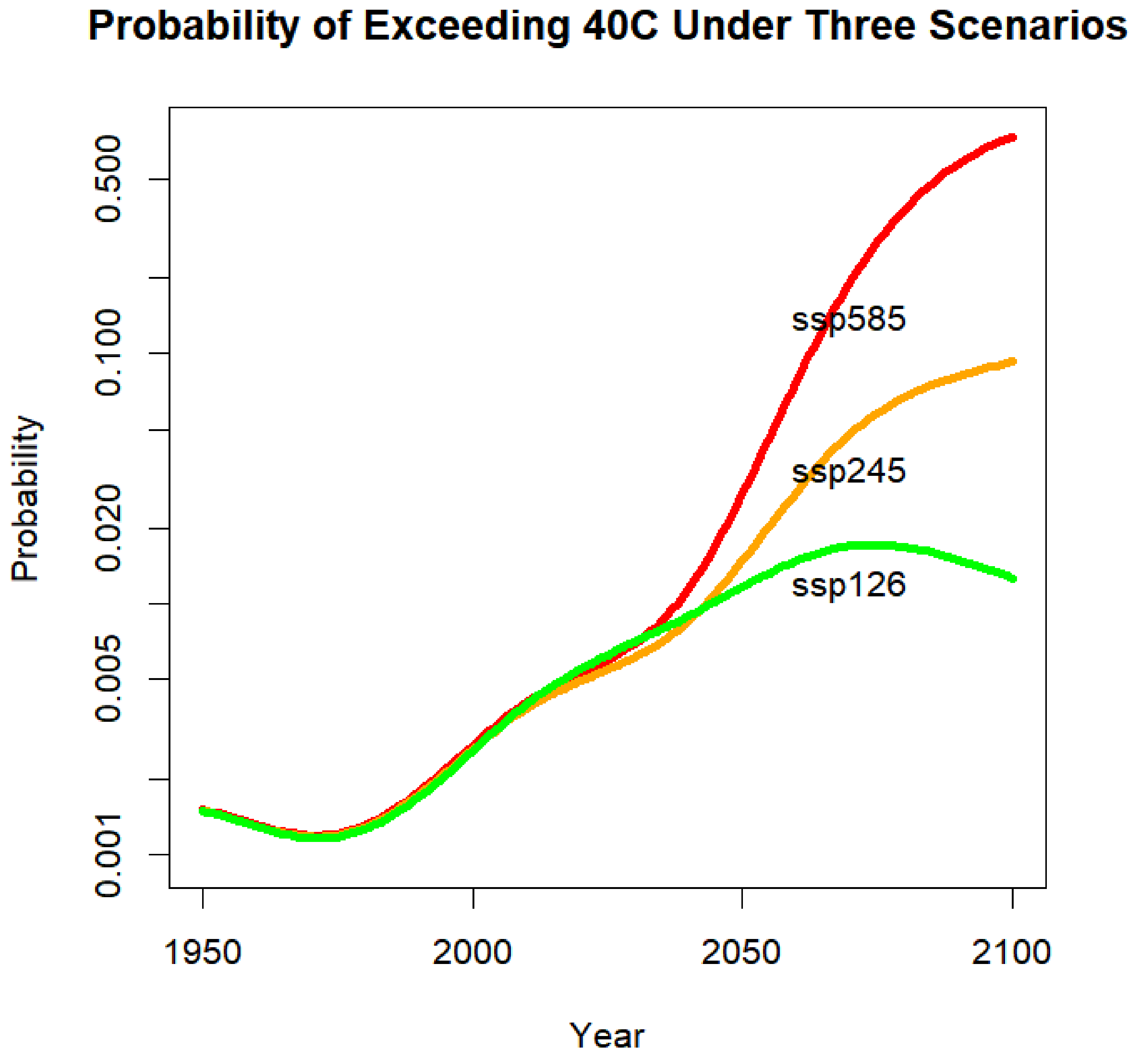
- In deference to the possibility of long-tailedness in Z_t, the model for that variable assumed t-distributed errors (though there is no evidence to contradict normality)


Projected Regional Summer Means (ssp585)
Anomaly
10
5
0
1850
1900
1950
2000
2050
2100
Year

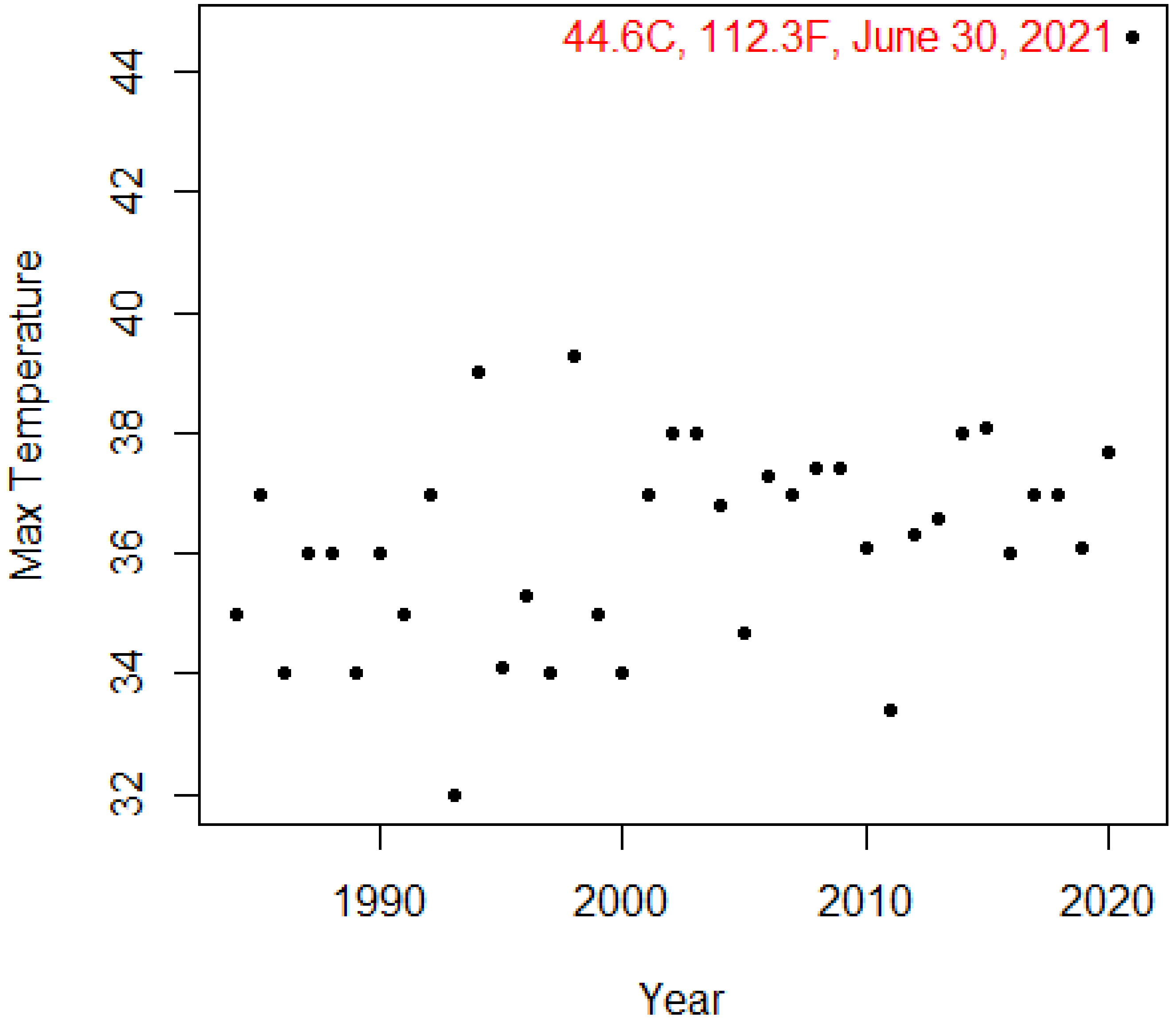

Projected Regional Summer Means (ssp245)
Anomaly
8
6
4
2
0
-2
1850
1900
1950
2000
2050
2100
Year



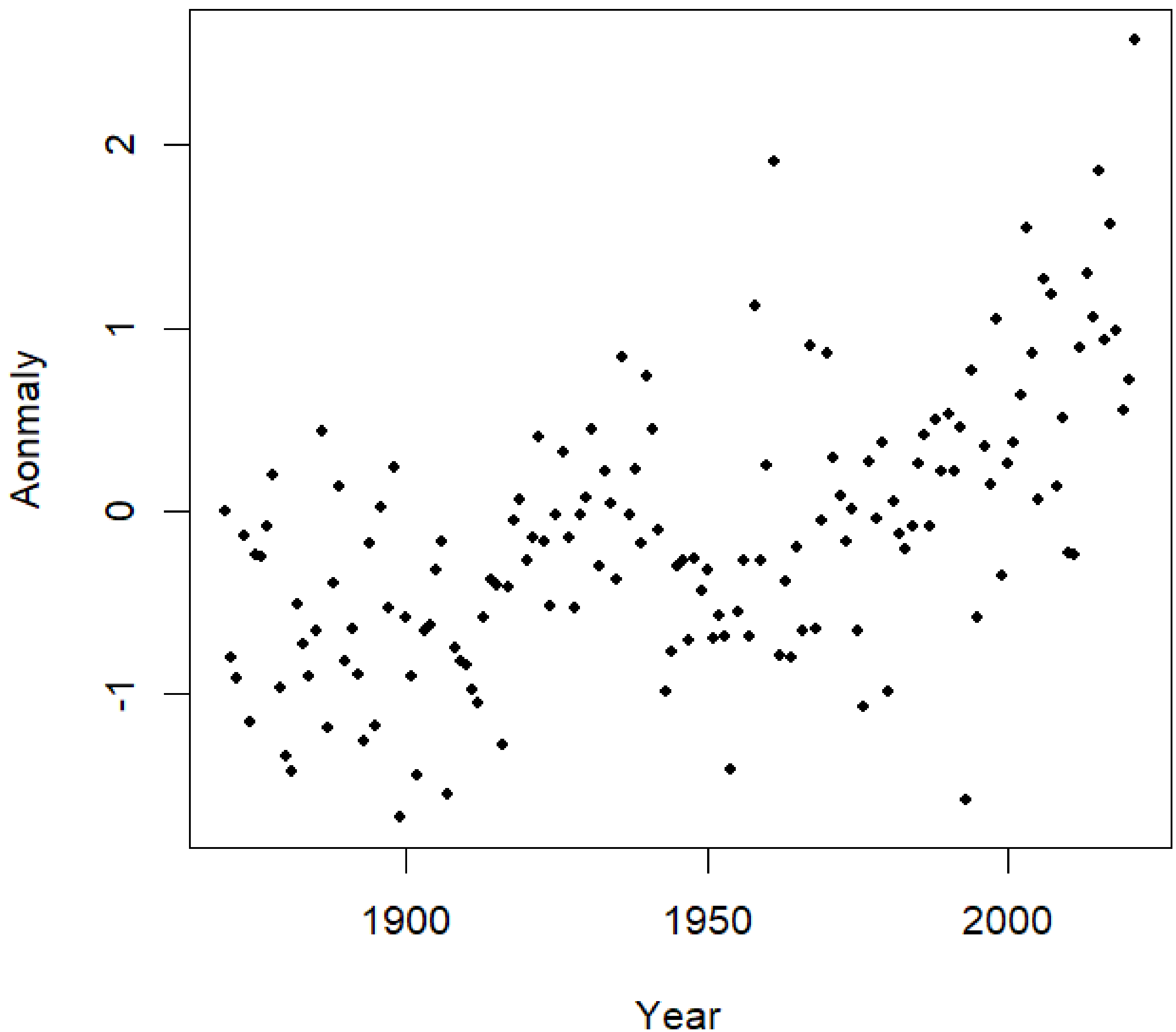
Projected Regional Summer Means (ssp126)
Anomaly
Year
-2
0
2
4
1850
1900
1950
2000
2050
2100


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

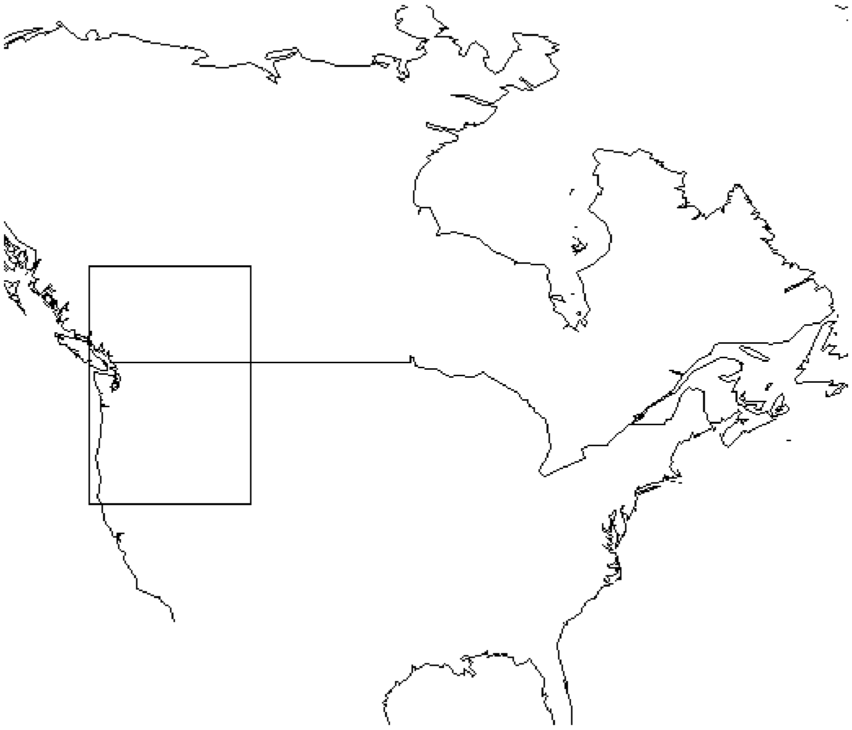
Annual Maximum Daily Maximum Temperatures in Kelowna, BC, 1984-2021

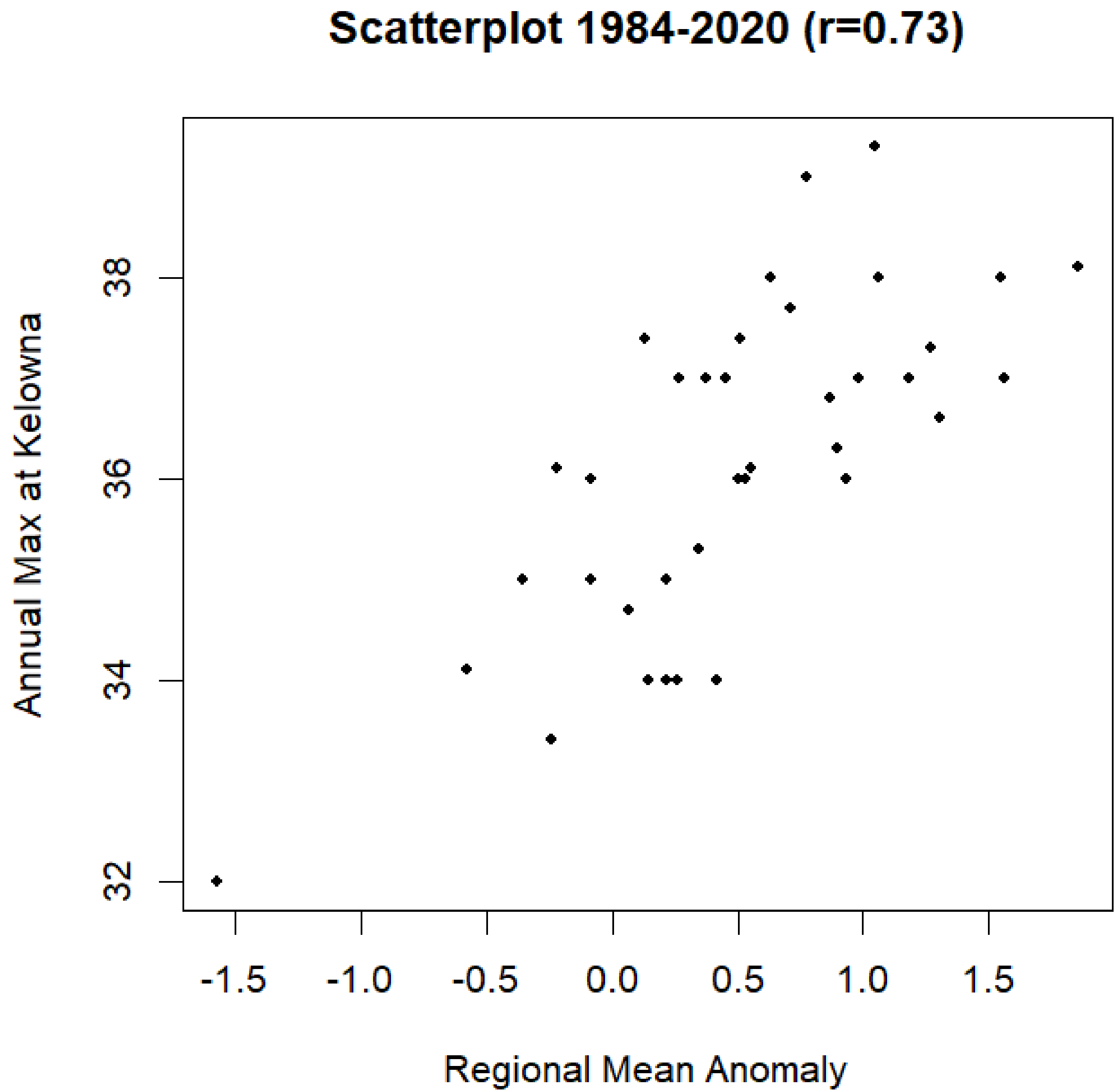


JJA Mean Temperature Anomalies 1870-2021 over 40-55 N, 110-125 W



Area Represented




Scatterplot 1984-2020 (r=0.73)
Annual Max at Kelowna
32
34
36
38
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
Regional Mean Anomaly

GEV Analysis for Kelowna

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

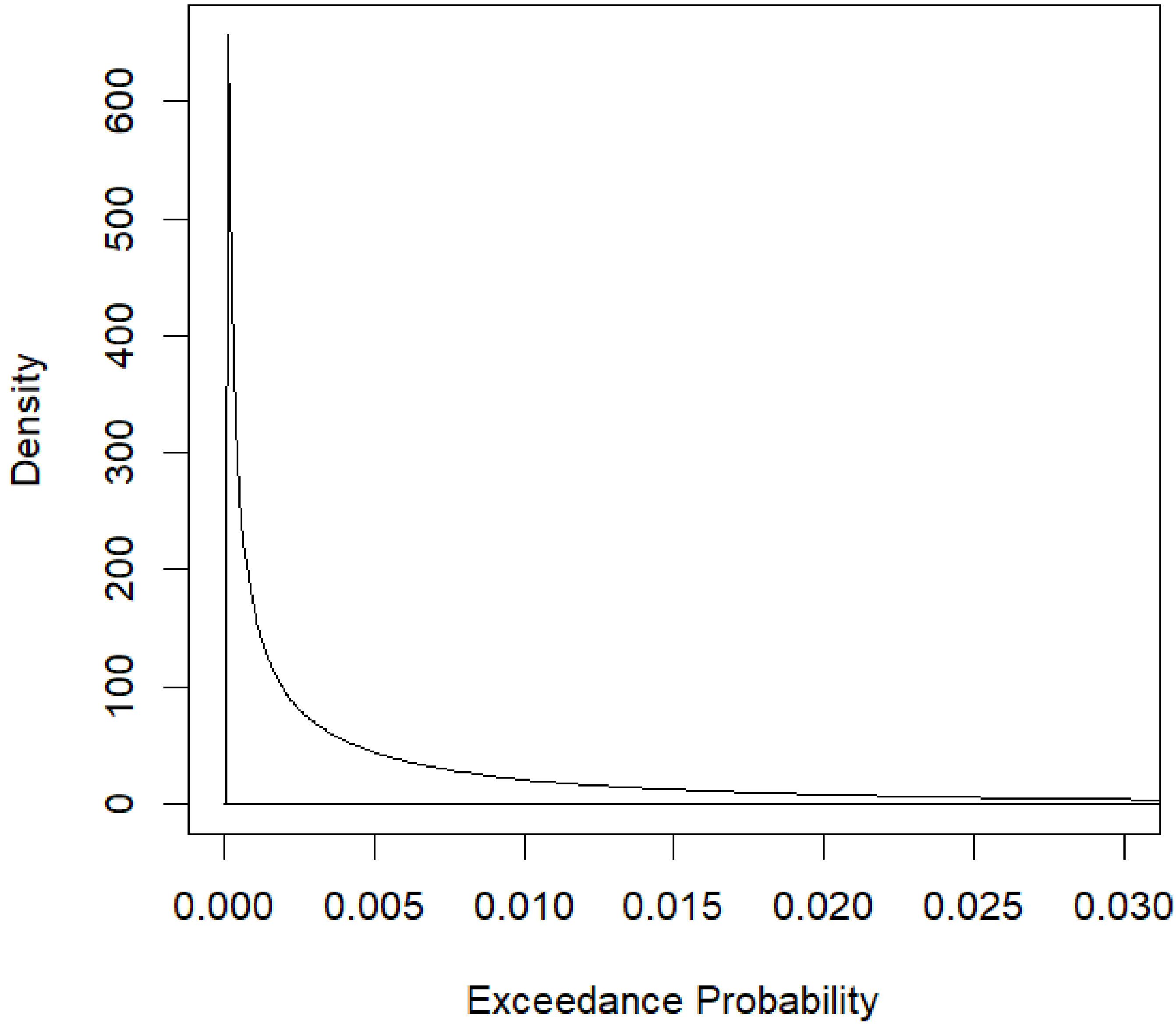
Parameter	Estimate	S.E.	t-value	p-value
β_0	34.8307	0.2384	146.0908	0.0000
β_1	1.8442	0.2656	6.9439	10^{-7}
$\log\sigma$	0.0365	0.1386	0.2633	0.7923
ξ	-0.2090	0.1466	-1.4257	0.1539

Estimated Annual Max Temperature in 2021: 44.5°C

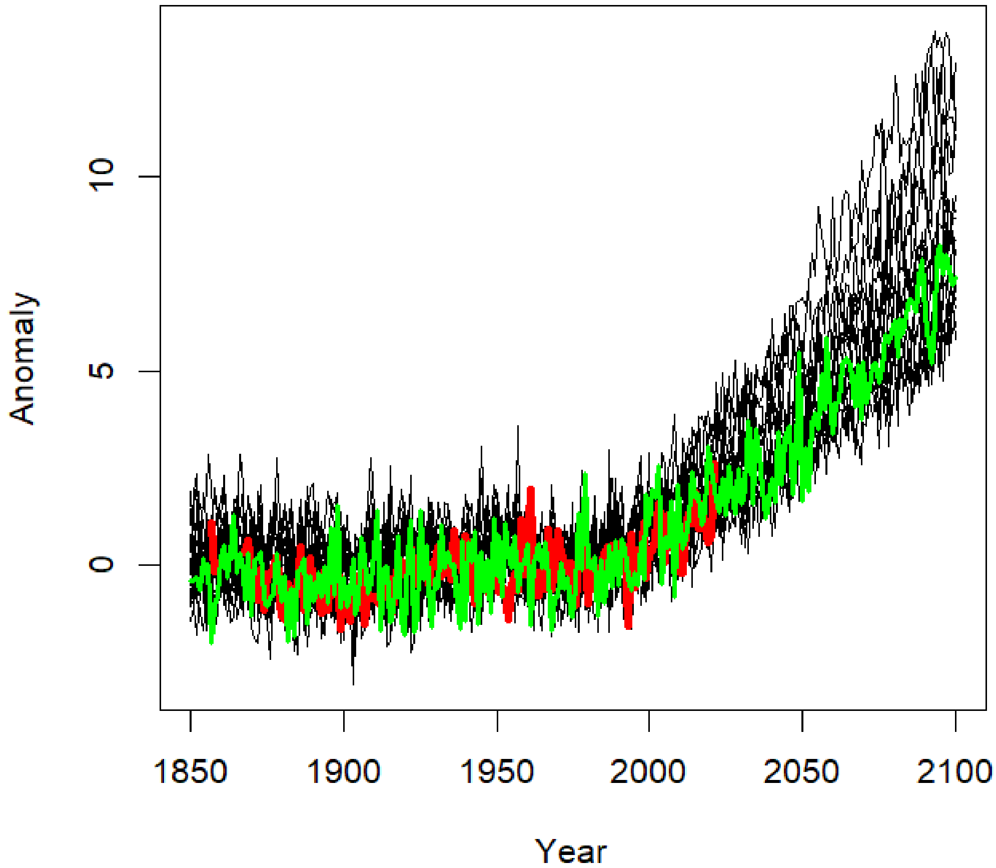
If P is probability of exceeding observed 2021 Value, then the posterior probability that $P > 0$ is 0.57, and the posterior mean of P is 0.0037 (MCMC analysis with flat prior)

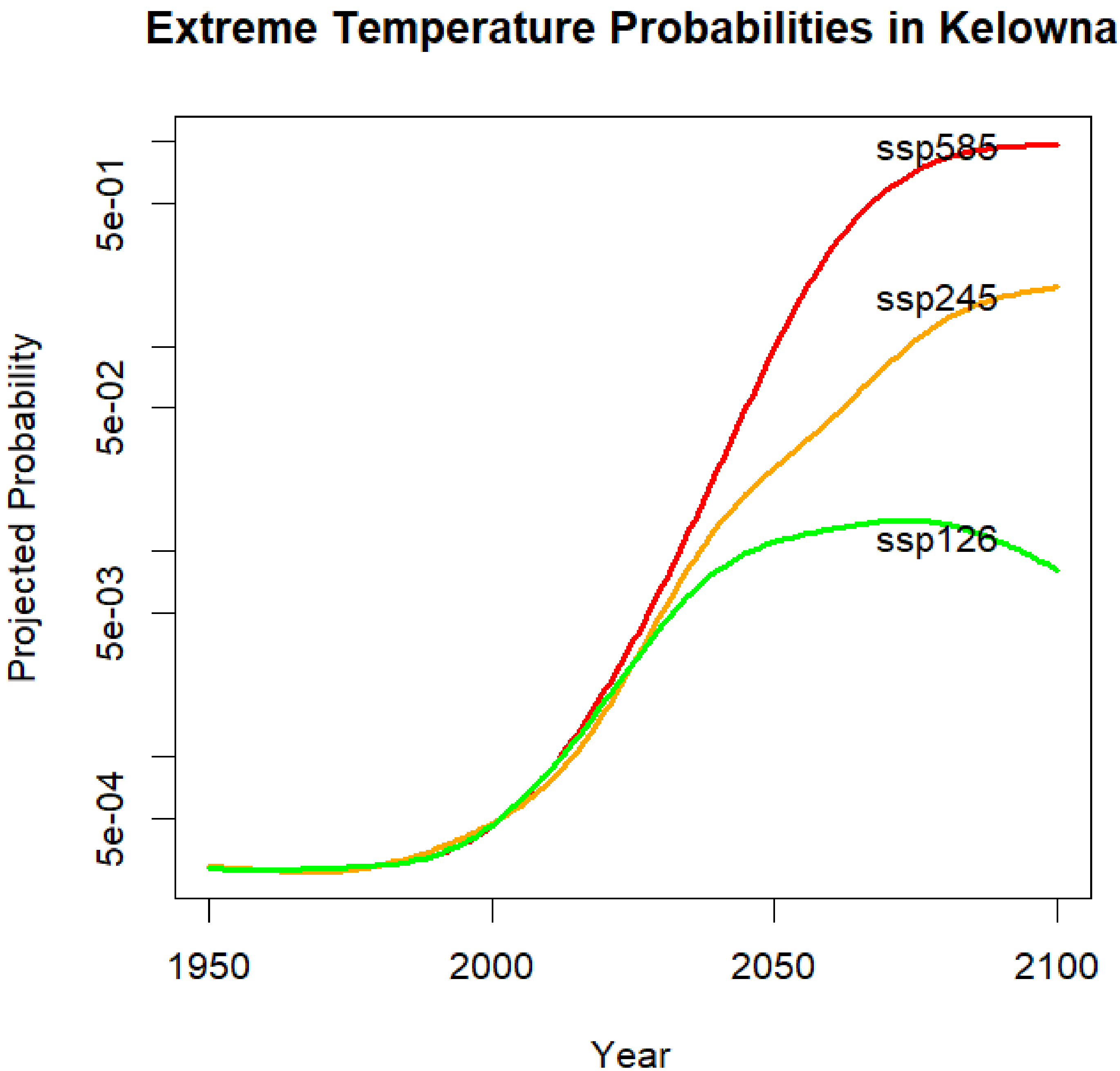
Side Note: Clarkson et al. (2022) claim the zero-probabiity problem is avoided by switching to a threshold exceedances approach. I have not verified that for myself.

Posterior PDF of the probability P of exceeding 44.6 C, conditional on P>0.0001



One Forward Projection Under ssp585




Extreme Temperature Probabilities in Kelowna
ssp585
ssp245
ssp126
5e-01
5e-02
5e-03
5e-04
Projected Probability
1950
2000
2050
2100
Year

Future Directions

- Improve the GEV analysis by spatial integration across stations (Casson and Coles 1999, Russell et al 2020, Risser 2020, Clarkson et al 2022, L. Zhang et al 2022)
- Further work needed on methodology for future projections (analogy with paleoclimatology, or with TLS method in conventional D&A, see D. Stephenson's talk at this meeting)
- Further examples, e.g. Hurricane Harvey precipitation extreme, conditioning on GoM SST means (Russell et al 2020, Hammerling et al. 2019)
- The objective is to obtain a fully automated method that can be applied using only public data sources
- Thank you for your attention!