SPATIAL EXTREMES
Richard L. Smith

Slides for class STOR 834: Extreme Value Theory

15 April 2025

THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL

CAS Survey

As you're all well aware, the College of Arts and Sciences (CAS) conducts a survey of every course at the end of semester.

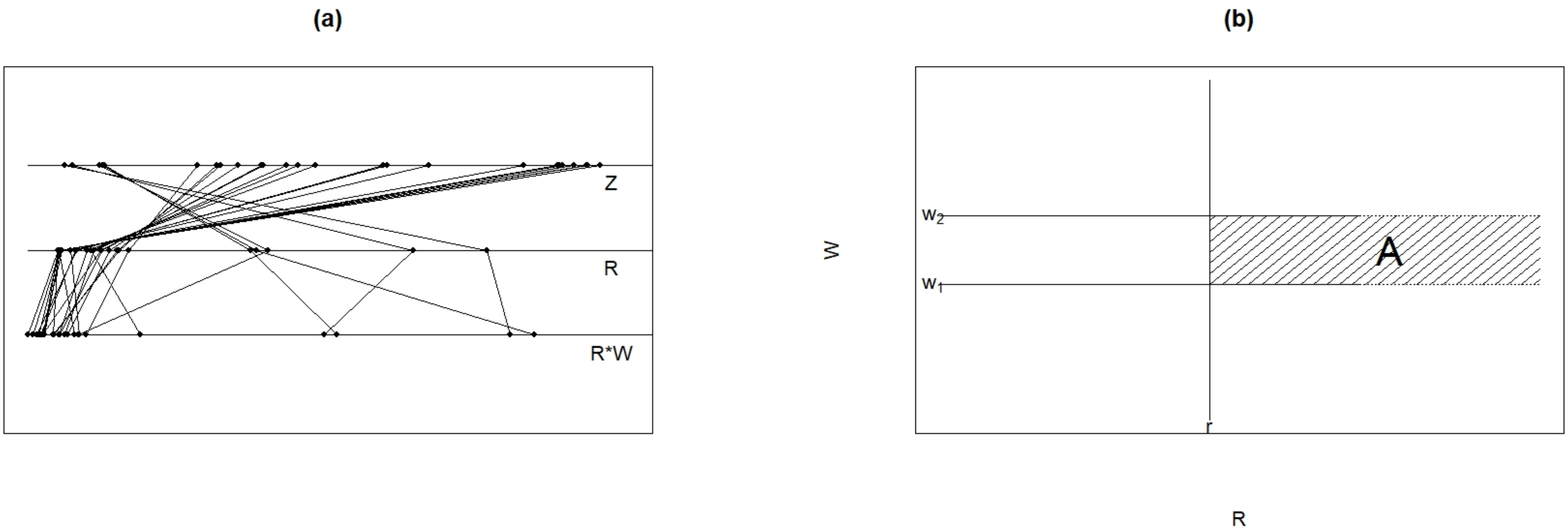
By now, you should have all received a link to the survey from this course.

Please respond! Even if you don't feel you have much to say, it's important that we record a response from as many students as possible, as the information is valuable to the CAS, to the STOR department, and to me personally. Your responses may determine whether this course is ever offered again!

Thank you for your cooperation.

The story so far ...

1. Multivariate Extremes
   (a) "Classical" theory — based on multivariate max-stable distributions — estimation theory exists in both block maxima and threshold exceedance settings (Tawn, Coles, Ledford etc. but see also the book by Beirlant *et al.* (2004) and many other references)
   (b) Extensions of classical theory — asymptotic dependence versus asymptotic independence — alternative models due to Ledford-Tawn, Heffernan-Tawn-Resnick and Ramos-Ledford but also other work that we haven't reviewed (e.g. many papers by Dan Cooley, Holger Rootzén, Johan Segers and others)
   (c) Attempts to unify the two approaches, e.g. Wadsworth *et al.* (2017) suggested a parametric model that included both AD and AI special cases, but still appears somewhat *ad hoc*.
2. Spatial Extremes
   (a) Latent process approach — not "fashionable" (i.e. most recent research other than mine has focused on other approaches) but still a viable technique for studying problems of spatial interpolation and spatial dependence
   (b) Introduction to max-stable process: reviewed basic theory of Poisson random measures (PRMs) and the mapping theorem
   (c) The "inverse Poisson process" (IPP), a PRM on $\mathbb{R}_+$ with intensity measure ν, where $\nu((z,\infty)) = \frac{1}{z}$, $0 < z < \infty$. Denoted $\{R_i,\ i = 1, 2, \ldots\}$ in the following.



(a) Illustration of Z, R and RW processes (W_i independent with mean 1)

$Z_i = E_1 + \ldots + E_i$, $\{E_i\}$ independent exponential (scaled)

$R_i = 1/Z_i$; multiply by $W_i = e^{X_i - 1/2}$, X_i IID $N(0,1)$; then $E\{W_i\} \equiv 1$; $\{R_i\}$ is an IPP

(b) Represent (R_i, W_i) as a Poisson process in $\mathbb{R}^2_+$; define set $A = (r, \infty) \times (w_1, w_2)$; then $\mu(A) = \frac{1}{r} \cdot \{F(w_2) - F(w_1)\}$

By the mapping theorem, $\{(R_i W_i), i = 1, 2, \ldots\}$ is a Poisson process in $\mathbb{R}_+$; the measure of the set $\{R_i W_i > z\}$ is $\int_0^\infty \int_0^\infty I(rw > z)\mu(drdw) = \int_0^\infty \left\{\frac{w}{z}\right\} F(dw) = \frac{1}{z}$. Hence $\{R_i W_i\}$ is an IPP.

The probability of the event $\max_i(R_i) < z$ is the probability that the set $R_i \in (z, \infty)$ is empty; this is $e^{-\nu((z,\infty))} = e^{-1/z}$; therefore, $\max_i(R_i)$ has unit Fréchet distribution

Hence $\max_i(R_i W_i)$ also has unit Fréchet distribution

Construction of a max-stable process

Assume $\{R_i\}$ is an IPP; $\{W_i(x),\ x \in \mathcal{X}\}$ are IID processes on $\mathcal{X}$; assume $E\{W_i(x)\} = 1$ for all i and $x \in \mathcal{X}$

Define $Z(x) = \max_i \{R_i W_i(x)\}$ for $x \in \mathcal{X}$.

Let $\{z(x),\ x \in \mathcal{X}\}$ be any non-negative function on $\mathcal{X}$; $z(x) = +\infty$ is allowed

$\Pr\{Z(x) \le z(x) \text{ for all } x \in \mathcal{X}\} = \Pr\{R_i W_i(x) \le z(x) \text{ for all } i > 0,\ x \in \mathcal{X}\}$
$\Pr\left\{R_i \le \inf_{\mathcal{X}} \frac{z(x)}{W_i(x)},\ i = 1, 2, \ldots\right\} = \exp\{-V(z(x),\ x \in \mathcal{X})\}$ where
$V(z(x),\ x \in \mathcal{X}) = E\left\{\sup_{x \in \mathcal{X}} \left(\frac{W(x)}{z(x)}\right)\right\}$, W any of $W_1,\ W_2,\ \ldots$.

1. For single $x \in \mathcal{X}$, $Z(x)$ has unit Fréchet distribution ($\Pr\{Z(x) \le z\} = e^{-1/z},\ 0 < z < \infty$)

2. V is *homogeneous of order -1*, meaning $V(a(z(x),\ x \in \mathcal{X})) = a^{-1} V((z(x),\ x \in \mathcal{X}))$

3. If $Z_1, Z_2, \ldots, Z_n$ are n independent copies of Z, then $\frac{1}{n}\max\{Z_1, Z_2, \ldots, Z_n\}$ has the same finite-dimensional distributions as Z.

4. If $\mathcal{D} \subset \mathcal{X}$ then $\Pr\{Z(x) \le z,\ \text{all } x \in \mathcal{D}\} = e^{-\theta_{\mathcal{D}}/z}$ where $\theta_{\mathcal{D}} = E\{\sup_{x \in \mathcal{D}} W(x)\}$ is the *extremal coefficient*.

The Brown-Resnick Process

Let $\{\epsilon(x),\ x \in \mathcal{X}\}$ be a Gaussian process with mean 0 and variance function $\mathrm{Var}\{\epsilon(x)\} = \sigma^2(x)$. Then $W(x) = \exp\left\{\epsilon(x) - \sigma^2(x)/2\right\}$ has mean 1 (by the well-known formula for the moment generating function for a normal distribution). Let $W_1, W_2, \ldots$ be independent copies of W and let $\{R_i,\ i = 1, 2, \ldots\}$ be an IPP. Then

$$Z(x) \;=\; \max_{i=1,2,\ldots} R_i W_i(x),\ x \in \mathcal{X}$$

is a max-stable process known as the *Brown-Resnick* process.

What to choose for $\epsilon(x)$? A process is *stationary* and *isotropic* if $\mathrm{Cor}\{\epsilon(x), \epsilon(x')\} = \rho\left(||x - x'||\right)$ for some function $\rho(\cdot)$ and $||x - x'||$ representing the (Euclidean) distance between x and x'. Examples: exponential, power law, Matérn...

However a better assumption is that the process be *intrinsically stationary* with a *variogram* $\mathrm{Var}\{\epsilon(x) - \epsilon(x')\} = \gamma\left(||x - x'||\right)$

If $\epsilon(\cdot)$ is stationary and $\sigma^2(\cdot)$ is constant then it is also intrinsically stationary with $\gamma(t) = 2\sigma^2(1 - \rho(t)),\ t > 0$. However there are also examples of intrinsically stationary processes that are not stationary, e.g. $\gamma(t) = c_0 + c_1 t^\lambda,\ c_0 \geq 0,\ c_1 \geq 0,\ 0 \leq \lambda < 2$. Special case: $c_0 = 0,\ \lambda = 1$ is Wiener process (a.k.a. Brownian motion)

Formula: $\Pr\left\{Z(x_1) \leq z_1, Z(x_2) \leq z_2\right\} = \exp\left\{-V(z_1, z_2)\right\}$ where
$V(z_1, z_2) = \frac{1}{z_1}\Phi\left(\frac{1}{a}\log\frac{z_2}{z_1} + \frac{a}{2}\right) + \frac{1}{z_2}\Phi\left(\frac{1}{a}\log\frac{z_1}{z_2} + \frac{a}{2}\right)$ where $a = \gamma(||x_1 - x_2||)$. This is the *Hüsler-Reiss* (1989) distribution.

Properties: as $a \to \infty$, $V(z_1, z_2) \to \frac{1}{z_1} + \frac{1}{z_2}$ (independent case). As $a \to 0$, $V(z_1, z_2) \to \frac{1}{\min(z_1, z_2)}$ (perfect dependence). Practical reason for preferring $a \to \infty$ as $||x_1 - x_2|| \to \infty$ (intrinsically stationary but not stationary)

Rainfall-Storms Interpretation

1. $\{R_i\}$ represent *magnitudes* of storms, normalized so that maxima are unit Fréchet
2. $\{W_i(x),\ x \in \mathcal{X}\}$ represent *shapes* of storms, showing how they are distributed over $\mathcal{X}$.
3. $Z(x)$ represents the largest storm in a given year at site x
4. However, like Wadsworth *et al.* (2017), we might get a wider class of processes by allowing more flexible choices

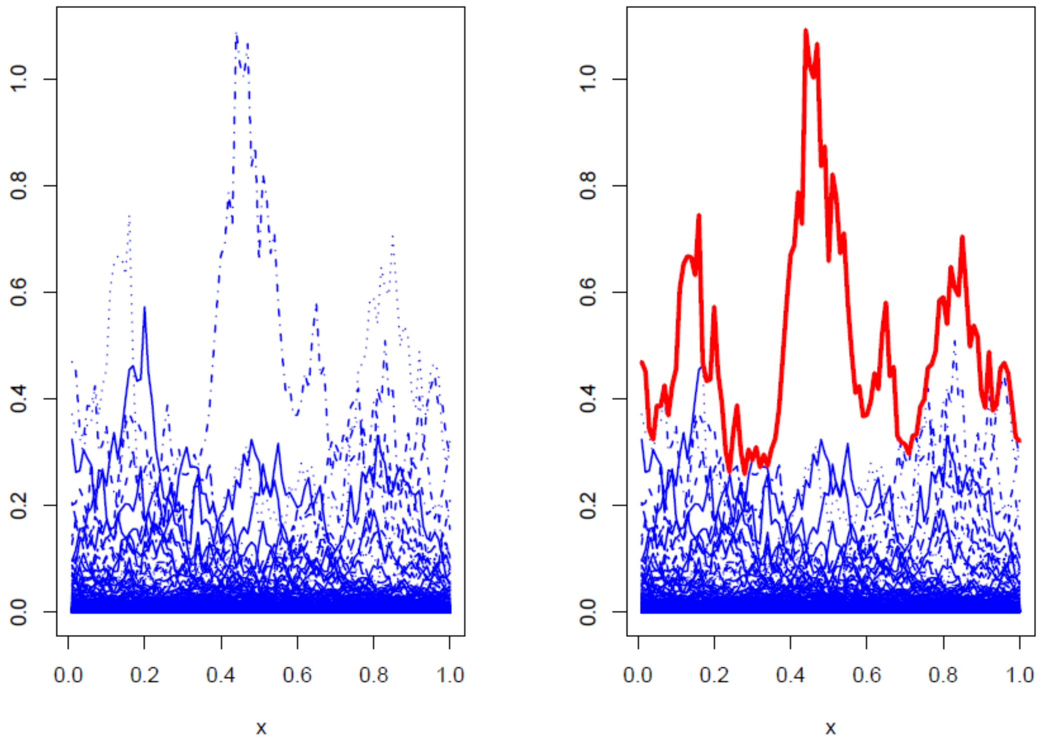


Figure 5.2 *Left hand plot: superimposed processes* $R_iW_i(x)$*, where* $\{R_i\}$ *is an IPP and* $W_i(x) = \exp\{\varepsilon_i(x) - 1/2\}$ *for independent Gaussian process* $\{\varepsilon(x)\}$*. Different indices i are indicated by different line types. Right hand plot: same, with the pointwise maximum process* $Z(x) = \max_i R_iW_i(x)$ *superimposed in red. Adapted from a figure in [50].*

Other Max-Stable Processes

1. Origins: characterization theorems by de Haan (1984), Giné, Hahn and Vatan (1990)

2. First attempt to apply the idea to real data: Smith (1990, unpub.)

3. Other developments: Coles (1993), Schlather (2002), Reich-Shaby (2012)

4. A different tack: *max-linear* models and extensions, e.g. Davis and Resnick (1989) paper on "max-ARMA" models; PhD thesis by Zhengjun Zhang (2002), see also Zhang and Smith (2010); recent papers by Falk and Zott (2017), Gissibl, Klüppelberg and Lauritzen (2021)

5. Extremal t process (Opitz 2013, Thibaud and Opitz 2015): $W(x) = m_\alpha \epsilon(x)_+^\alpha$ where $\epsilon(\cdot)$ is a stationary Gaussian process with mean 0 and variance 1, $\alpha > 0$ and $m_\alpha = \frac{\pi^{1/2} 2^{1-\alpha/2}}{\Gamma((\alpha+1)/2)}$.

   For this model, $V(z_1, z_2)$ can be calculated as a linear function of $T_{\alpha+1}$ distribution functions, limit $\alpha \to \infty$ may correspond to Brown-Resnick.

Estimation of Max-Stable Processes

1. Need to consider both marginal distributions and dependence structure

2. Marginal distributions: GEV parameters at site $\mathbf{s}$ are $\mu(\mathbf{s};\boldsymbol{\phi}), \psi(\mathbf{s};\boldsymbol{\phi}), \xi(\mathbf{s};\boldsymbol{\phi})$

3. Could have complicated structure, e.g. smoothing splines to allow for non-linear dependence on $\mathbf{s}$

4. Transform to unit Fréchet distributions (probability integral transform)

5. Dependence structure:a max-stable process with parameters $\boldsymbol{\theta}$ (e.g. Brown-Resnick with unknown σ^2 and variogram)

6. Two strategies: (a) estimate $\boldsymbol{\phi}$ first, then $\boldsymbol{\theta}$, (b) estimate jointly

7. Theory usually shows (b) better, but can be complicated to implement

8. Technical problem with estimating $\boldsymbol{\theta}$: although all the models we have considered have a closed-form solution for the bivariate joint distributions, they don't extend to joint distributions for 3 or more sites

9. Solution: *composite likelihood*

Method of Composite Likelihood

Source paper: Padoan, Ribatet and Sisson (2010, *JASA*)

Assume n replications on d sampling points, $\ell_{i,j,j'}(\boldsymbol{\theta}) = \log f(z_{ij}, z_{ij'}; \boldsymbol{\theta})$ where z_{ij} is ith observation at the jth site

Maximize $CL(\boldsymbol{\theta}) = \sum_{i=1}^{n} \sum_{j=2}^{d} \sum_{j'=1}^{j-1} w_{jj'} \ell_{i,j,j'}(\boldsymbol{\theta})$ where $w_{jj'}$ are fixed weights depending only on spatial locations

The MCLE $\widehat{\boldsymbol{\theta}}_C$ maximizes $CL(\boldsymbol{\theta})$

Properties of estimator: define $\widehat{K} = \sum_i \sum_{j'<j} \frac{\partial \ell_{i,j,j'}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \ell_{i,j,j'}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T}$, $\widehat{J} = -\sum_i \sum_{j'<j} \frac{\partial^2 \ell_{i,j,j'}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}$, both evaluated at $\widehat{\boldsymbol{\theta}}_C$.

"Sandwich estimator" $\widehat{J}^{-1} \widehat{K} \widehat{J}^{-1}$ is estimator of Cov$\{\widehat{\boldsymbol{\theta}}_C\}$

Theorem: under the right conditions

$$\left(\widehat{J}^{-1} \widehat{K} \widehat{J}^{-1}\right)^{-1/2} \left(\widehat{\boldsymbol{\theta}}_C - \boldsymbol{\theta}\right) \quad \stackrel{d}{\to} \quad \mathcal{N}_p(\mathbf{0}, I_p)$$

Model selection: use the *CLIC* (*C*omposite *L*ikelihood *I*nformation *C*riterion)

$$\text{CLIC} \;=\; 2\left\{\text{tr}(\widehat{J}^{-1}\widehat{K}) - \text{CL}(\widehat{\boldsymbol{\theta}}_C)\right\}$$

Progress Towards Exact Likelihood Methods

In some cases it is possible to compute an exact MLE but this is very computationally intensive

Stephenson-Tawn (2005) method; idea of breaking up time into short intervals so that $W_i(x)$ may be observed individually; however this creates risk of bias if misidentified (Wadsworth (2015))

Threshold methods are, in some respects, easier to apply than block maxima methods (papers by Wadsworth and Tawn; Thibaud and Opitz; Engelke *et al.*)

The "ABC method" (estimate likelihood by simulation; Erhardt and Smith 2012)

Neural network methods (many recent papers by Huser and co-authors)

However, overall the trend is moving away from max-stable models (see arxiv preprint, "Modeling of spatial extremes in environmental data science: Time to move away from max-stable processes" by Huser, Opitz and Wadsworth, 2024)

Back to Attribution of Extreme Weather Events

Source paper: Zhang, Risser, Wehner and O'Brien, *JABES* 2024.

Leveraging Extremal Dependence to Better Characterize the 2021 Pacific Northwest Heatwave

Likun ZHANG, Mark D. RISSER, Michael F. WEHNER, and Travis A. O'BRIEN

In late June, 2021, a devastating heatwave affected the US Pacific Northwest and western Canada, breaking numerous all-time temperature records by large margins and directly causing hundreds of fatalities. The observed 2021 daily maximum temperature across much of the U.S. Pacific Northwest exceeded upper bound estimates obtained from single-station temperature records even after accounting for anthropogenic climate change, meaning that the event could not have been predicted under standard univariate extreme value analysis assumptions. In this work, we utilize a flexible spatial extremes model that considers all stations across the Pacific Northwest domain and accounts for the fact that many stations simultaneously experience extreme temperatures. Our analysis incorporates the effects of anthropogenic forcing and natural climate variability in order to better characterize time-varying changes in the distribution of daily temperature extremes. We show that greenhouse gas forcing, drought conditions and large-scale atmospheric modes of variability all have significant impact on summertime maximum temperatures in this region. Our model represents a significant improvement over corresponding single-station analysis, and our posterior medians of the upper bounds are able to anticipate more than 96% of the observed 2021 high station temperatures after properly accounting for extremal dependence.

Supplementary materials accompanying this paper appear online.

Key Words: Extreme event attribution; Concurrent extremes; Extreme value theory; Gaussian scale mixtures; Granger causality; Spatial statistics.

Remark: I already discussed some of this paper, see lecture of March 18, about 16:30 in. However, I didn't get into the technical details of the spatial model.

Zhang, Risser, Wehner and O'Brien (2024), I

Temperature Data

- Global Historical Climatology Network-Daily (GHCN-D), USA only
- *Homogenized* records from 1950–2020, 487 stations within 116.5–125^{o}W, 40–49^{o}N.
- Summertime (JJA) annual maxima daily maxima (TXx); exclusion criteria reduce to 438 stations
- *Non-homogenized* data from 2021, 470 stations

Non-stationary elements (covariates to GEV model)

- Greenhouse gas forcings ($GHG_t,\ t = 1950, \ldots, 2021$)
- ENSO Longitude Index ($ELI_t,\ t = 1950, \ldots, 2021$)
- PDSI Drought Index for each location $\mathbf{s}$ ($PDSI(\mathbf{s}, t),\ t = 1950, \ldots, 2021$)
- Urbanization Binary Index for each location $\mathbf{s}$ ($UB(\mathbf{s}, t),\ t = 1950, \ldots, 2021$)
- Topographical covariates, e.g. elevation, distance to coast (fixed in time but not in space)

Zhang, Risser, Wehner and O'Brien (2024), II

Marginal Distributions

$$\Pr\{Y(\mathrm{s},t) \le y\} = \exp\left[-\left\{1+\xi(\mathrm{s},t)\left(\frac{y-\mu(\mathrm{s},t)}{\sigma(\mathrm{s},t)}\right)\right\}_{+}^{-1/\xi(\mathrm{s},t)}\right],$$

$$\begin{aligned}
\mu(\mathrm{s},t) &= \mu_0(\mathrm{s}) + \mu_1(\mathrm{s})GHG(t) + \mu_2(\mathrm{s})PDSI(\mathrm{s},t) + \mu_3(\mathrm{s})ELI(t) + \mu_4(\mathrm{s})UB(\mathrm{s},t),\\
\log\sigma(\mathrm{s},t) &= \sigma_0(\mathrm{s}) + \sigma_1(\mathrm{s})GHG(t) + \sigma_2(\mathrm{s})ELI(t),\\
\xi(\mathrm{s},t) &= \xi(\mathrm{s}),
\end{aligned}$$

where the functions $\mu_0(\mathrm{s}), \mu_1(\mathrm{s}), \ldots, \xi(\mathrm{s})$ are functions of s through *thin-plate splines* plus topographical covariates ($\gamma(\mathrm{s})$ is any of $\mu_0(\mathrm{s}), \mu_1(\mathrm{s}), \ldots, \xi(\mathrm{s})$)

$$\gamma(\boldsymbol{s}) \mid \beta_0^\gamma, \beta_1^\gamma, \ldots, \beta_B^\gamma = \beta_0^\gamma + \sum_{i=1}^{99} \beta_i^\gamma \underbrace{b_i(\boldsymbol{s})}_{\text{T.P.splines}} + \sum_{j=1}^{5} \alpha_j^\gamma \underbrace{g_j(\boldsymbol{s})}_{\text{topography}}, \qquad (5)$$

Zhang, Risser, Wehner and O'Brien (2024), III

Spatial Dependence Model

We assume that the *copula* of the spatial field $Y(\mathbf{s},t)$ (JJA max temperature at site $\mathbf{s}$ in year t) is that of

$$X(\mathbf{s},t) \;=\; R_t \cdot W(\mathbf{s},t) + \epsilon(\mathbf{s})$$

where $R_t \sim \text{Pareto}\{(1-\delta)/\delta\}$, $W(\mathbf{s},t)$ is a standard isotropic and stationary Gaussian process transformed to standard Pareto margins, and $\epsilon(\mathbf{s}) \sim \mathcal{N}(0,\tau^2)$ independent for each $\mathbf{s}$

Asymptotic independence when $\delta \in (0,1/2]$, asymptotic dependence when $\delta \in (1/2,1)$

[Precedents: Huser, Opitz and Thibaud (2017), "Bridging Asymptotic Independence and Dependence in Spatial Extremes Using Gaussian Scale Mixtures"; see also Huser and Wadsworth (2019); Zhang, Shaby and Wadsworth (2022)]

Parameters $\boldsymbol{\theta}_W$ reflect unknown spatial covariance parameters of $W(\mathbf{s},t)$ (presumably independent for each t)

Relate X to Y by equating marginal distributions, i.e.

$$F_Y\{Y(\mathbf{s},t)\} \;=\; F_X\{X(\mathbf{s},t)\} \text{ for each } t$$

Formulate as a Bayesian hierarchical model, fit by a big MCMC algorithm

Zhang, Risser, Wehner and O'Brien (2024), IV

Statistical Counterfactual to Quantify Human Influence

Framework of Granger (1969) causality

Directly based on observational data; doesn't use climate models (contrast "Pearl causality")

"Factual" predictions based on present-day GHG; "Counterfactual" predictions based on 1950 GHG

Calculate risk ratio for each $\mathbf{s}$:

$$RR(\mathbf{s}) \;=\; \frac{p_F(\mathbf{s})}{p_C(\mathbf{s})}$$

where p_F and p_C represent probabilities of observed event under both factual and counterfactual models

Results (GEV marginals)

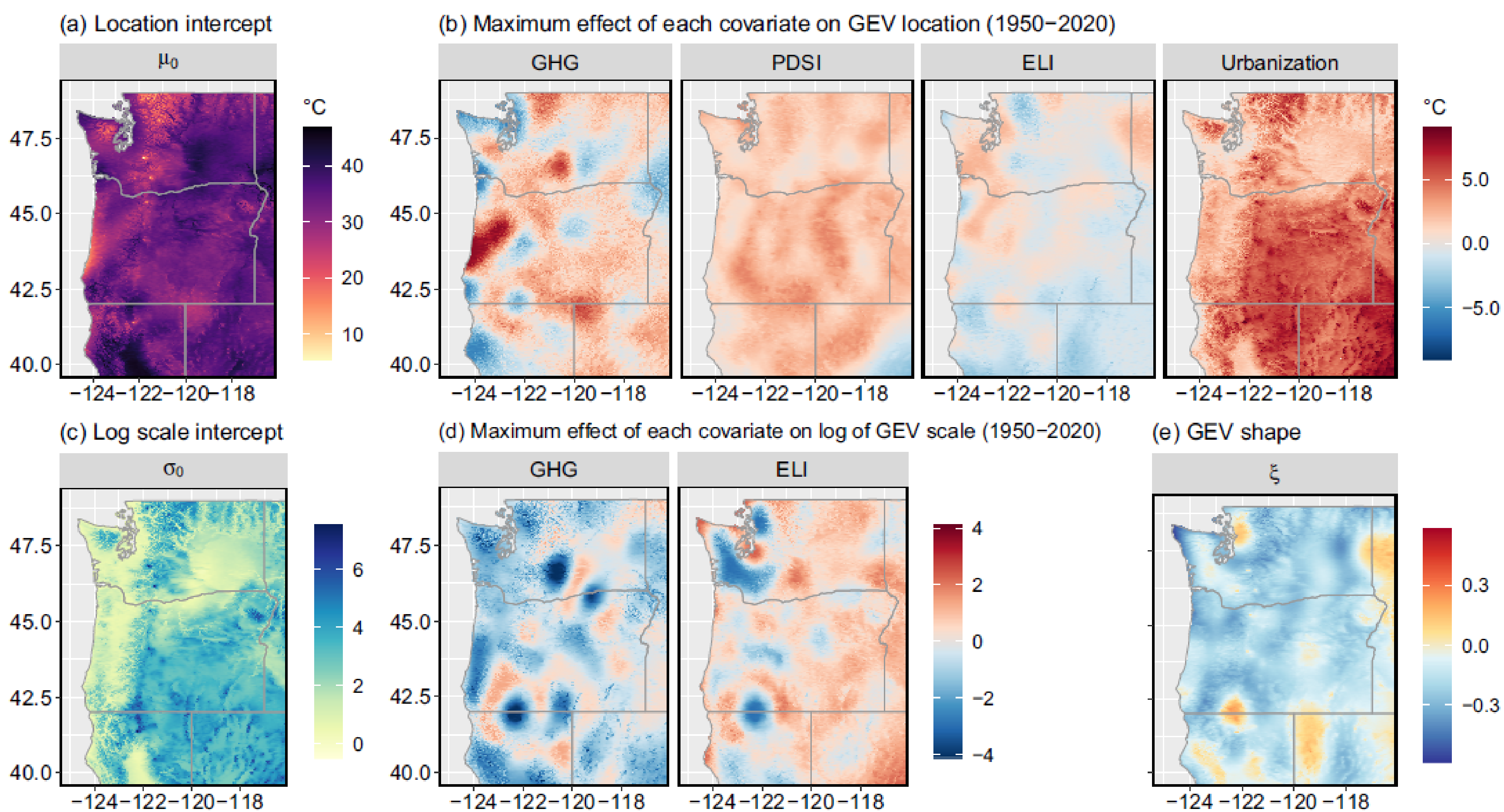


Figure 2. Posterior means of the maximum effects of GEV coefficients $\{\mu_0(s), \ldots, \mu_1(s), \sigma_0(s), \ldots, \sigma_2(s), \xi(s)\}$ under M4; see Sect. 3.3 for the definition of M4 and see Eq. (9) to learn how to calculate the maximum effects .

Zhang, Risser, Wehner and O'Brien (2024), VI

Results (Distribution of maximum temperatures)

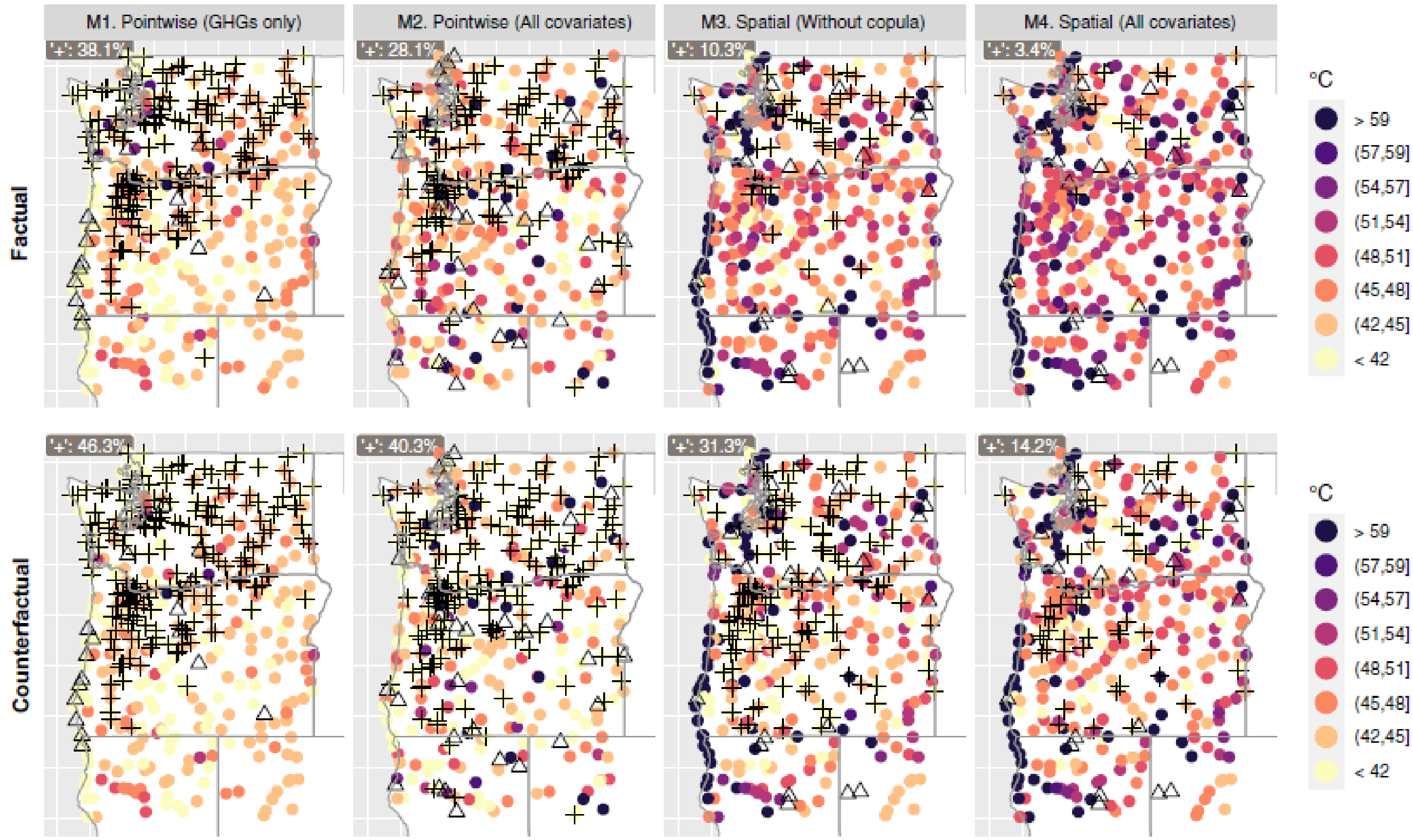


Figure 3. The posterior median of the GEV upper bound under four statistical models (M1–M4; see Sect. 3.3). The first and second rows show upper bound estimates under the factual and counterfactual scenarios, respectively. For these rows, the '△' signifies an infinite upper bound (corresponding to $\xi > 0$), and '+' signifies stations for which the observed 2021 TXx exceeded the posterior median of the upper bound. The inset text in each panel displays the fraction of stations where the upper bound is exceeded. See Supplemental Fig. B.8 for the differences in the upper bounds for each scenario.

Zhang, Risser, Wehner and O'Brien (2024), VII

Results (Risk Ratios)

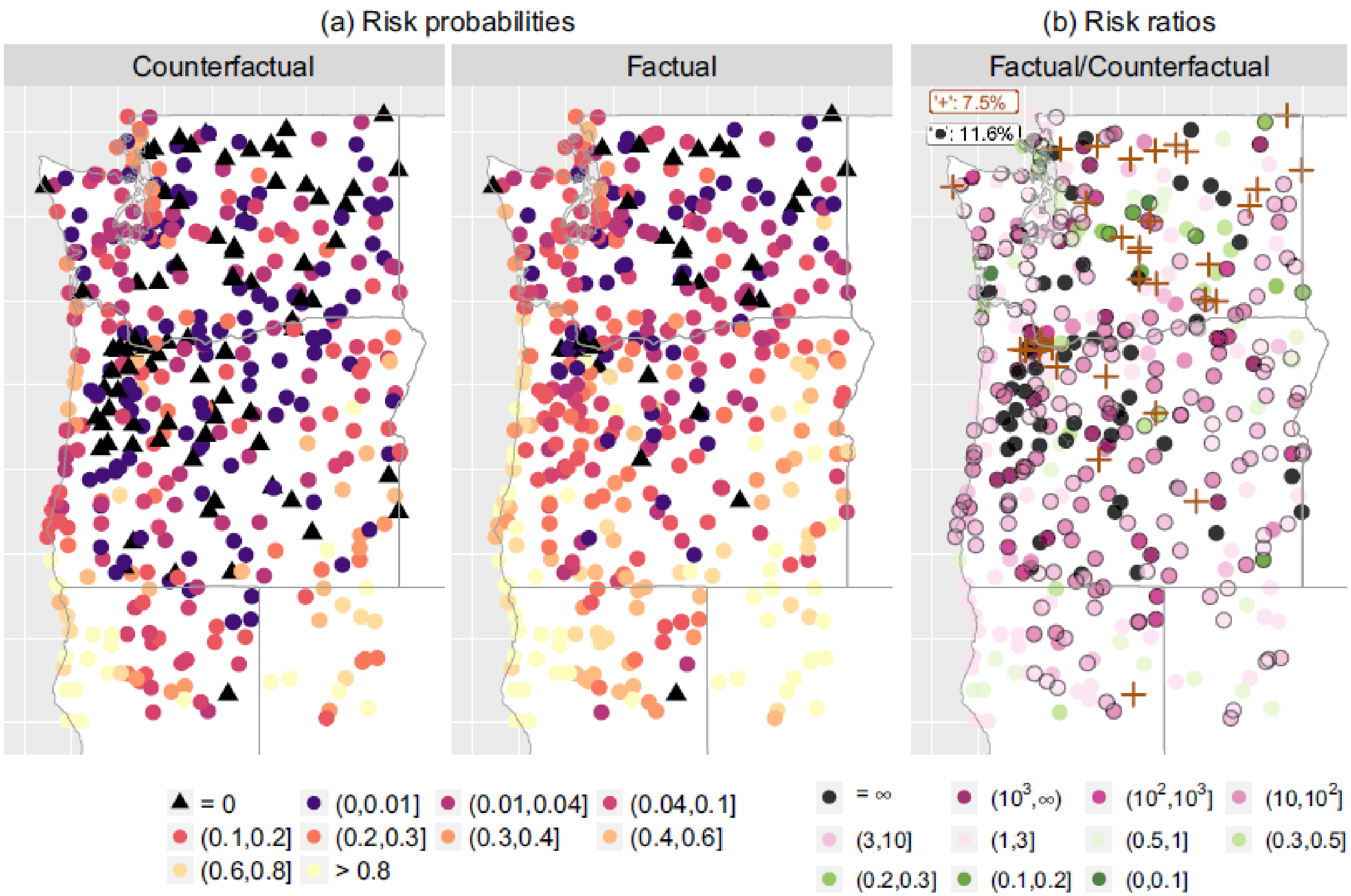


Figure 4. Subfigure **a** shows the station-specific posterior medians of the risk probabilities calculated from statistical model M4 for the Counterfactual (left) and Factual (right) climate scenarios, and subfigure **b** shows their ratio. For the risk probabilities, solid black triangles indicate gauged locations for which the risk probability best estimate is zero; for the risk ratios, solid black circles denote $RR(s) = \infty$ (wherein the counterfactual risk probability is zero but the factual risk probability is nonzero) and yellow '+' shows where the risk ratios are undefined (wherein both counterfactual and factual risk probabilities are zero). In the rightmost panel, points that are plotted with additional '$\circ$' sign indicates that risk ratio estimates are statistically significantly different from 1 .

Zhang, Risser, Wehner and O'Brien (2024), VIII

Summary and Conclusions

- Both covariate modeling and spatial dependence are needed for the final result but (according to the authors) the spatial dependence component is the more critical

- Analysis done only for US stations because Canada stations are not homogenized; need to repeat analysis for Canadian stations (worse for "heat dome" effect)

- Posterior median station maximum is greater than 2021 observed value in nearly all stations (authors' comment: literally "all stations" if posterior median is replaced by 95% quantile)

- Risk ratios in range 10–100 are much more interpretable than those in previous analyses, but there are still stations where the estimated risk ratio is infinite or undetermined (0/0), and some where it is < 1

- Properties of spatial model: need for one that exhibits "nonstationary tail dependence"??

Thank you for attending this course!!

Don't forget the survey

Now, over to you ....