Spatial Methods for Modeling Extreme and Rare Events

Samuel A. Morris

North Carolina State University

August 22, 2016

Overview

- Brief overview of theory for extremes
- Three principal contributions:
  1. A spatio-temporal skew-t model for threshold exceedances
  2. Modeling spatial rare binary events with a max-stable extension to the GEV link function
  3. Empirical basis functions to explore and model extremal spatial dependence

Motivation

- Average behavior is important to understand, but it does not paint the whole picture
  - e.g. When constructing river levees, engineers need to be able to estimate a 100-year or 1000-year flood levels
  - e.g. Probability of ambient air pollution exceeding a certain threshold level
- Estimating the probability of rare events is challenging because these events are, by definition, rare
- Spatial extremes is promising because it borrows information across space
- Spatial extremes are also useful for estimating probability of extremes at sites without data

Non-spatial analysis: Block maxima

Fisher-Tippett-Gnedenko theorem

- Let $X_1, \ldots, X_t$ be i.i.d.
- Consider the block maximum $M_t = \max(X_1, \ldots, X_t)$
- If there exist normalizing sequences $a_t > 0$ and $b_t \in \mathcal{R}$ such that

$$\frac{M_t - b_t}{a_t} \xrightarrow{\mathrm{d}} G(z)$$

  then $G(z)$ follows a generalized extreme value distribution (GEV) (Gnedenko, 1943)

- This motivates the use of the GEV for block maximum data

Non-spatial analysis: Block maxima

- GEV distribution

$$G(y) = \Pr(Y < y) = \begin{cases} \exp\left\{-\left[1+\xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} & \xi \neq 0 \\ \exp\left\{-\exp\left(-\frac{y-\mu}{\sigma}\right)\right\} & \xi = 0 \end{cases}$$

  where
  - $\mu \in \mathcal{R}$ is a location parameter
  - $\sigma > 0$ is a scale parameter
  - $\xi \in \mathcal{R}$ is a shape parameter
    - Unbounded above if $\xi \geq 0$
    - Bounded above by $(\mu - \sigma)/\xi$ when $\xi < 0$

- Challenges:
  - Lose information by only considering maximum of a block
  - Underlying data may not be i.i.d.

Other approaches: Threshold methods

- Perhaps there exists a threshold T beyond which values are extreme
- T can be selected using a mean residual life plot.
- Two general approaches:
  - Peaks over threshold: Model $Y > T$ using generalized Pareto distribution
  - Censored distribution function:

$$F(y) = \begin{cases} F(T), & y \leq T \\ F(y), & y > T \end{cases}$$

- Challenges:
  - Sensitive to threshold selection
  - Temporal dependence

Max-stable processes for spatial data

- Consider i.i.d. spatial processes $x_j(\mathbf{s})$, $j = 1, \ldots, t$
- Let $M_t(\mathbf{s}) = \bigvee_{j=1}^{t} x_j(\mathbf{s})$ be the block maximum at site $\mathbf{s}$
- If there exists normalizing sequences $a_t(\mathbf{s})$ and $b_t(\mathbf{s})$ such that for all sites, $\mathbf{s}$,

$$\frac{M_t(\mathbf{s}) - b_t(\mathbf{s})}{a_t(\mathbf{s})} \xrightarrow{\mathrm{d}} G(\mathbf{s})$$

  then $G(\mathbf{s})$ is a max-stable process (Smith, 1990)
- Therefore, max-stable processes are the standard model for block maxima

Multivariate representations

- Marginally at each site, observations follow a GEV distribution
- For a finite collection of d sites the distribution function for the multivariate GEV (mGEV) is

$$\Pr(\mathbf{Z} \leq \mathbf{z}) = G^*(\mathbf{z}) = \exp[-V(\mathbf{z})]$$

$$V(\mathbf{z}) = d \int_{\Delta_d} \bigvee_{i=1}^{d} \frac{w_i}{z_i} H(\mathrm{d}w)$$

where

- $V(\mathbf{z})$ is called the exponent measure (few closed-form expressions exist)
- $\Delta_d = \{\mathbf{w} \in \mathcal{R}_+^d \mid w_1 + \cdots + w_d = 1\}$
- H is a probability measure on Δ_d
- $\int_{\Delta_d} w_i H(\mathrm{d}w) = 1/d$ for $i = 1, \ldots, d$

Quantifying dependence

- Problem: Covariance and correlation focuses on deviations around the mean and not the extremes
- Want dependence measure to capture likelihood of seeing values that are jointly extreme
- Two common measures of dependence (bivariate):
  - Extremal coefficient $\vartheta \in (1, 2)$:

$$P[Z(\mathbf{s}_1) < c, Z(\mathbf{s}_2) < c] = P[Z(\mathbf{s}_1) < c]^{\vartheta(\mathbf{s_1}, \mathbf{s_2})}$$

  - $\chi \in (0, 1)$ (Coles, 1999):

$$\chi(\mathbf{s}_1, \mathbf{s}_2) = \lim_{c \to \infty} P[Z(\mathbf{s}_1) > c | Z(\mathbf{s}_2) > c]$$

Existing challenges

- Multivariate max-stable models have nice features, but they are
  - Computationally challenging (e.g, the asymmetric logistic has $2^{n-1}(n+2) - (2n+1)$ free parameters)
  - Joint density only available in low dimensions (Wadsworth and Tawn, 2014; Thibaud and Opitz, 2015)
- Some recent approaches
  - Bayesian hierarchical model (Reich and Shaby, 2012)
  - Pairwise likelihood (Padoan, 2010; Huser and Davison, 2014)
  - Trivariate likelihood (Genton et al, 2011)
- Many opportunities to explore new methods

Project 1:

A Space-time Skew-t Model for Threshold Exceedances

Under revision:
Biometrics

Motivation

Ozone compliance for Clean Air Act (EPA)

- Annual fourth-highest daily maximum 8-hour concentration, averaged over 3 years, not to exceed 75 ppb
- Annual fourth-highest is the 99th percentile for the year
- Common objectives are
  - To interpolate to unmonitored sites
  - Detect changes in extremes over time
  - Study meteorological conditions that lead to extreme events

Is max-stable the panacea?

- Max-stable process is an elegant approach, but does that mean it's correct?
  - It is only an approximation
  - There are less complicated approximations (e.g. we could model daily data as a Gaussian process (GP))
- If the goal is spatial interpolation, perhaps this is competitive

GP - Asymptotic Independence

- A GP leads to simple interpretation and computing, but asymptotic independence
  - If $Y(\mathbf{s}_1)$ and $Y(\mathbf{s}_2)$ are bivariate normal then $\chi(\mathbf{s}_1, \mathbf{s}_2) = 0$ (asymptotic independence)
- This suggests Kriging will not capture extremes
- So much is known for the Gaussian case: nonstationarity, multivariate, numerical approximations, . . .
- Rather than toss it out, can we patch it up?

Spatial skew-t process

A spatial skew-t process (Azzalini and Capitanio, 2014) resembles a GP but exhibits asymptotic dependence

$$
\begin{aligned}
Y_t(\mathbf{s}) &= \mathbf{X}(\mathbf{s})^\top \boldsymbol{\beta} + \lambda \sigma_t |z_t| + \sigma_t v_t(\mathbf{s}) \\
z_t &\sim \text{Normal}(0, 1) \\
\sigma_t^2 &\sim \text{InvGamma}(a/2, b/2) \\
v_t &\sim \text{Spatial GP}
\end{aligned}
$$

- Location: $\mathbf{X}(\mathbf{s})^\top \boldsymbol{\beta}$
- Scale: $b > 0$
- Skewness: $\lambda \in \mathcal{R}$
- Degrees of freedom: $a > 0$

Good properties

- Flexible t marginal distribution with four parameters including the degrees of freedom which allows for heavy tails ($a = 1$ gives a Cauchy)
- Computation on the order of a GP; the only extra steps are z_t and σ_t which have conjugate full conditionals
- Asymptotic dependence: $\chi(\mathbf{s}_1, \mathbf{s}_2) > 0$ for all $\mathbf{s}_1$ and $\mathbf{s}_2$

Bad properties and remedies

- Modeling all the data (bulk and extreme) can lead to poor tail probability estimates if the model is misspecified
- Long-range dependence: $\chi(\mathbf{s}_1, \mathbf{s}_2) > 0$ for all $\mathbf{s}_1$ and $\mathbf{s}_2$ even if $\mathbf{s}_1$ and $\mathbf{s}_2$ are far apart
- This occurs because all sites share z_t and σ_t
- Remedies:
  - We use a censored likelihood to focus on the tails
  - We propose a local skew-t process

Censored likelihood

- Censored likelihood: We censor the data

$$\tilde{Y}_t(\mathbf{s}) = \begin{cases} T & \text{for } Y_t(\mathbf{s}) \leq T \\ Y_t(\mathbf{s}) & \text{for } Y_t(\mathbf{s}) > T \end{cases}$$

- Censoring is handled using standard Bayesian imputation methods
- The threshold T is chosen by cross-validation
- If T is moderately extreme in the distribution (e.g. $q(0.75)$), set $\lambda = 0$

Local skew-t process

- Consider a set of spatial knots for day t
- Let the knots $\mathbf{v}_{t1}, ..., \mathbf{v}_{tK}$ follow a homogeneous Poisson process over the domain of interest (in practice we fix K)
- The knots partition the domain if we assign location $\mathbf{s}$ to subregion $k = \arg\min_l ||\mathbf{s} - \mathbf{v}_{tl}||$

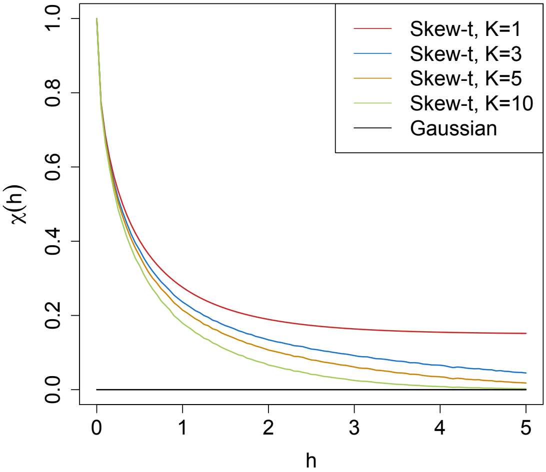
Local skew-t process

- Associated with each is
  - $z_{tk} \sim \text{Normal}(0, 1)$
  - $\sigma^2_{tk} \sim \text{InvGamma}(a/2, b/2)$
- If s is in subregion k then

$$Y_t(\mathbf{s}) = \mathbf{X}(\mathbf{s})^T \boldsymbol{\beta} + \lambda \sigma_{tk} |z_{tk}| + \sigma_{tk} v_t(\mathbf{s})$$

- The marginal distribution remains a t, but partitioning breaks long-range spatial dependence

χ-statistic by $h = ||s_1 - s_2||$



Temporal dependence

- It may not be reasonable to assume that observations are temporally independent (e.g. flooding, high temperatures)
- Temporal dependence is handled through the z_{tk}, σ_{tk} and $\mathbf{v}_{tl}$
- Method:
  - Use a copula to transform parameters to *nice* space (i.e. $\mathcal{R}$)
  - AR(1) structure imposed on parameters in transformed space
  - Transform back to original parameter space to preserve skew-t

Results of a simulation study

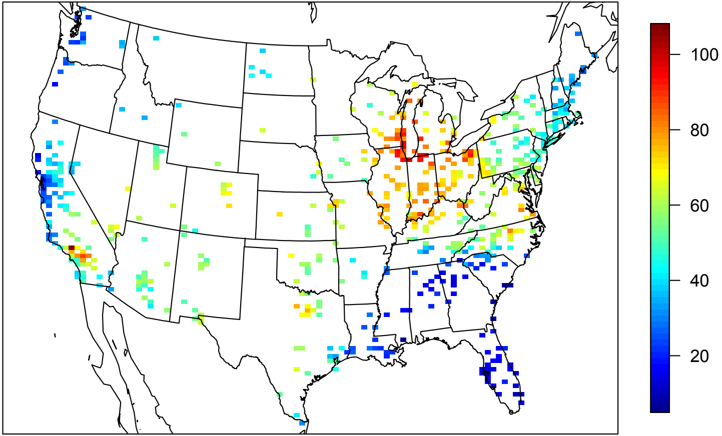
In terms of Brier scores for spatial prediction:

- Data generated as a GP:
    - skew-t is close to GP
    - max-stable is 15% – 30% worse than GP
- Data generated as a skew-t with multiple partitions:
    - skew-t is 15% better than GP
    - max-stable is 30% worse than GP
- Data generated as asymmetric logistic (max-stable):
    - skew-t is close to GP
    - max-stable performs 10% better than GP
- Data generated as Brown-Resnick (max-stable):
    - skew-t performs 40% – 60% better than GP
    - max-stable performs 40% – 60% better than GP

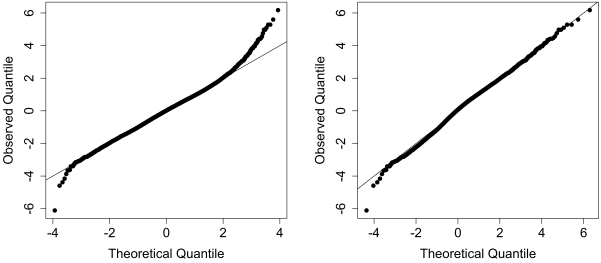
Application to ozone

- The USEPA has an extensive network of ozone monitors throughout the US
- We will analyze ozone for 31 days in July, 2005 at $n = 1,089$ stations
- Currently the EPA regulates the annual 99^{th} percentile
- Our objective is to map the probability of an extreme ozone event

Ozone on July 10



Q-Q plots

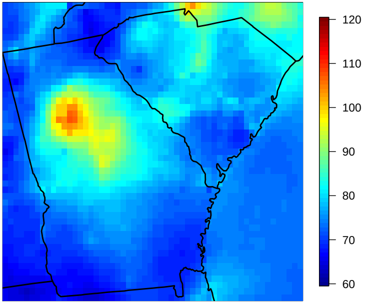


Gaussian Q-Q plot (left) and skew-t with $a = 10$ and $\lambda = 1$ Q-Q plot (right)

Cross-validation

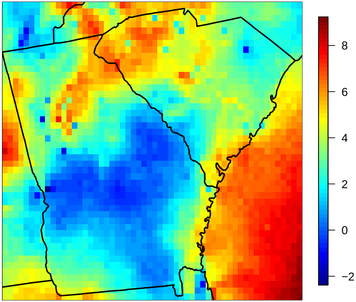
- We split the sites into training and testing
- We found that $K = 15$ knots and censoring at T equal to the median with no time series gave the best results
- Results were not sensitive to these tuning parameters
- This model was 5% more accurate (Brier score) than GP
- The max-stable model fit was 15% less accurate than GP

Fitted 99^{th} percentiles



Gaussian (left) Symmetric-t, 10 knots, T = 75, Time series (right)

Difference (Thresholded t - Gaussian)



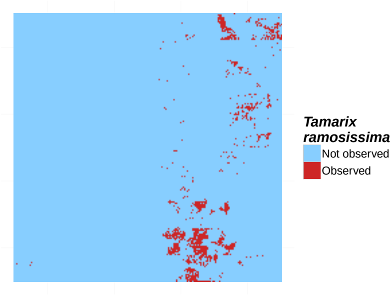
Difference between Symmetric-t, 10 knots, T = 75 and Gaussian

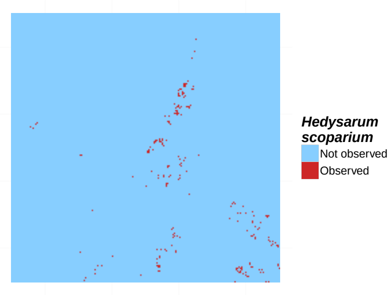
Project 2:

Rare Spatial Binary Regression

To be submitted:
Journal of Agricultural, Biological, and Environmental Statistics

Motivation





1-km^2 area in PR China.

- *Tamarix ramosissima* (left) Rareness: $\approx 6\%$
- *Hedysarum scoparium* (right) Rareness: $\approx 0.5\%$

Motivation

- For rare species, occupancy modeling can be challenging
  - Rare: to mean not occurring often (e.g. 5% or less)
- Binary regression is based on thresholding:

$$Y(\mathbf{s}) = I[Z(\mathbf{s}) > c]$$

- GP traditionally used for $Z(\mathbf{s})$, but theory suggests that GP will not capture dependence
  - c is in the tail of the distribution
  - Asymptotic dependence does not exist for GP
  - Therefore for large c this is virtually a non-spatial analysis
- Propose max-stable extension to GEV link (Wang and Dey, 2010)

Hierarchical max-stable representation

- Representation from Reich and Shaby (2012)
- Consider a set of L knots, $\mathbf{v}_1, \ldots, \mathbf{v}_L$
- Standardized Gaussian weights $w_l(\mathbf{s}_i)$

$$w_l(\mathbf{s}_i) = \frac{\exp\left[-0.5\left(\frac{||\mathbf{s}_i - \mathbf{v}_l||}{\rho}\right)^2\right]}{\sum_{l=1}^{L} \exp\left[-0.5\left(\frac{||\mathbf{s}_i - \mathbf{v}_l||}{\rho}\right)^2\right]}$$

- Is a valid a low-rank approximation if $L < n$, but in this application we place knots at all data points

Hierarchical max-stable representation

- Model the spatial dependence using

$$\theta(\mathbf{s}) = \left[\sum_{l=1}^{L} A_l w_l(\mathbf{s})^{1/\alpha}\right]^{\alpha}$$

where

- $A_l \overset{\text{iid}}{\sim} \text{PS}(\alpha)$ are positive stable random effects
- $w_l(\mathbf{s})$ are scaled kernel function so that $\sum_{l=1}^{L} w_l(\mathbf{s}) = 1$
- $\alpha \in (0, 1)$ is a parameter controlling strength of spatial dependence in residuals (0: high, 1: independent)

Hierarchical max-stable representation

- Let $Z(\mathbf{s})$ be a max-stable process
- Conditioned on the random effects $A_1, \ldots, A_L$, then

$$Z(\mathbf{s}_i) \mid A_l \overset{\text{ind}}{\sim} \text{GEV}[\mu^*(\mathbf{s}_i), \sigma^*(\mathbf{s}_i), \xi^*]$$
$$A_l \overset{\text{iid}}{\sim} \text{PS}(\alpha)$$

where

$$\mu^*(\mathbf{s}_i) = \mathbf{X}(\mathbf{s}_i)^\top \boldsymbol{\beta} + \frac{\theta(\mathbf{s}_i)^\xi - 1}{\xi}$$
$$\sigma^*(\mathbf{s}_i) = \alpha\theta(\mathbf{s}_i)^\xi$$
$$\xi^* = \alpha\xi$$

- The marginal distribution at site $\mathbf{s}_i$ is $\text{GEV}[\mathbf{X}(\mathbf{s}_i)^\top \boldsymbol{\beta}, 1, \xi]$

Hierarchical model

- Hierarchical model:

$$Y(\mathbf{s}_i)|A_l \overset{\text{ind}}{\sim} \text{Bern}\{\pi(\mathbf{s}_i)\}$$
$$\pi(\mathbf{s}_i) = 1 - \exp\left\{\sum_{l=1}^{L} A_l \left[\frac{w_l(\mathbf{s}_i)}{z(\mathbf{s}_i)}\right]^{1/\alpha}\right\}$$

where

$$z(\mathbf{s}_i) = \begin{cases} (1 - \xi \mathbf{X}(\mathbf{s}_i)\boldsymbol{\beta})^{-1/\xi} & \xi \neq 0 \\ \exp(-\mathbf{X}(\mathbf{s}_i)\boldsymbol{\beta}) & \xi = 0 \end{cases}$$

- Fix $\xi = 0$ when no covariates

Spatial dependence

- Cohen's kappa:

$$\kappa(\beta) = \frac{P_A - P_E}{1 - P_E}$$

  where

  - P_A: Joint probability of agreement
  - P_E: Joint probability of agreement under assumption of independence

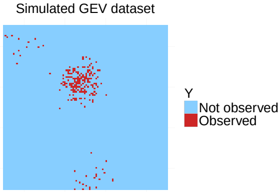
- Consider $Z(\mathbf{s}_1)$ and $Z(\mathbf{s}_2)$ both GEV$(\beta, 1, 1)$ then

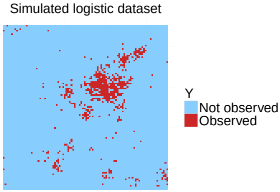
$$\kappa = \lim_{\beta \to \infty} \kappa(\beta) = 2 - \vartheta(\mathbf{s}_1, \mathbf{s}_2) = \chi$$

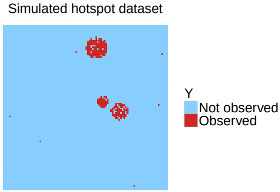
- When $Z(\mathbf{s})$ is Gaussian, $\kappa = 0$

Simulation study: Settings

- 50 simulated populations:
  - Link: GEV, Logistic, Hotspot
  - Data generated on 100×100 grid
- Sample datasets:







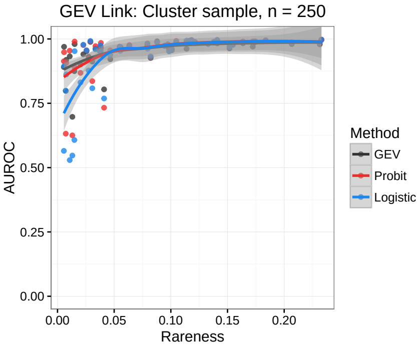
Simulation study: Settings

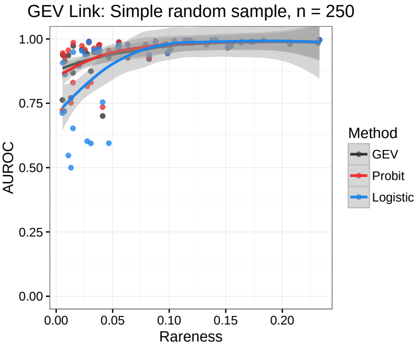
- Sampling strategies:
  - Sample type: Cluster, Simple Random Sample
  - Sample size: $n = 100, 250$ initial sites
- Models fit:
  - Spatial GEV ($\xi = 0$)
  - Spatial probit
  - Spatial logistic (`spBayes::spGLM`)
- Consider Brier score and area under receiver operating characteristic (AUROC) curve

Simulation study: Results

- Results:
  - GEV model shows small advantage over others for GEV link with $n = 250$ and cluster sampling
  - Probit wins in other settings, but GEV is similar
- Results provide some evidence of advantage for GEV method as rareness increases

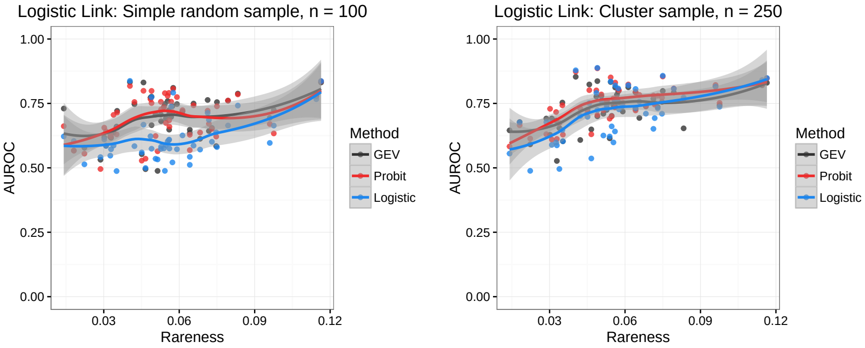
Simulation study: Results





Loess smooth of AUROC by rareness for GEV link functions

Simulation study: Results



Loess smooth of AUROC by rareness for logistic link functions

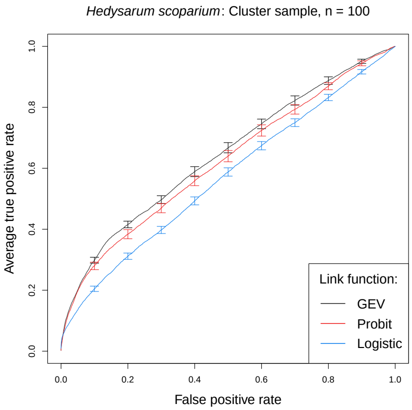
Data analysis

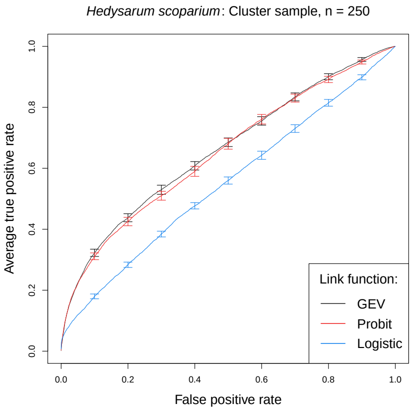
- Census data for *Tamarix ramosissima* ($\approx 6\%$) and *Hedysarum scoparium* ($\approx 0.5\%$) in 1-km^2 region of PR China
- Analysis similar to simulation study
  - Sample types: Cluster, Simple Random Sample
  - Sample sizes: $n = 100, 250$ initial sites
  - Models fit: Spatial GEV ($\xi = 0$), probit, logit
- 50 samples taken from each species for each sample type and sample size

Data analysis: Results

- Support simulation finding about GEV performance with respect to rareness
- For *Tamarix ramosissima*, probit generally performs better
- For *Hedysarum scoparium*, with cluster sampling and for smaller sample sizes, GEV model generally performs better

Data analysis: Results





ROC curves for *Hedysarum scoparium*

Project 3:

Empirical Basis Functions for Max-stable Dependence

To be submitted:
Annals of Applied Statistics

Motivation

- Want fit max-stable models on high dimensional data
- When n is large, computing is onerous
- Dimension reduction by placing $L < n$ spatial knots and using standardized Gaussian kernel functions as in Reich and Shaby (2012) and Project 2
- This works, but can we improve performance with a different set of basis functions
  - Exploratory data analysis like principal components (PC)
  - Useful for inference and predictions

Low-rank positive-stable representation

- Any max-stable process can be written as pointwise maximum of infinitely many processes (deHaan, 2006)
- Wang and Stoev (2011) truncate at L processes
- Unlikely that a realization is equal to the point-wise maximum of L processes, so we follow Reich and Shaby (2012) and set

$$Z_t(\mathbf{s}) = \theta_t(\mathbf{s})\varepsilon_t(\mathbf{s})$$

where $\theta_t(\mathbf{s})$ is a spatial process and $\varepsilon_t(\mathbf{s}) \stackrel{\text{iid}}{\sim} \text{GEV}(1, \alpha, \alpha)$

Low-rank positive-stable representation

- Spatial process is

$$\theta_t(\mathbf{s}) = \left(\sum_{l=1}^{L} A_{tl} B_l(\mathbf{s})^{1/\alpha} \right)^{\alpha}$$

  where $A_{tl} \sim \mathrm{PS}(\alpha)$

- Z_t is max-stable marginally over the random effects A_{tl}
- The joint distribution is mGEV with asymmetric logistic dependence function

Low-rank positive-stable representation

- Dependence is measured by the extremal coefficient ϑ, defined via

$$P[Z_t(\mathbf{s}_1) < c, Z_t(\mathbf{s}_2) < c] = P[Z_t(\mathbf{s}_1) < c]^{\vartheta(\mathbf{s}_1,\mathbf{s}_2)}$$

- For the low-rank PS model

$$\vartheta(\mathbf{s}_1, \mathbf{s}_2) = \sum_{l=1}^{L} \left[B_l(\mathbf{s}_1)^{1/\alpha} + B_l(\mathbf{s}_2)^{1/\alpha}\right]^{\alpha} \in [1, 2]$$

- Propose to use empirical basis functions for $B_l(\mathbf{s})$ (instead of $w_l(\mathbf{s})$ from rare binary)

Estimating the EBFs, $B_l(\mathbf{s})$

1. Use a rank transformation to standardize data for each $\mathbf{s}$
2. Estimate the extremal dependence between each pair of sites (using χ or madogram), $\hat{\vartheta}(\mathbf{s}_i, \mathbf{s}_j)$
3. Spatially (4D) smooth the sample dependence measures for $\tilde{\vartheta}(\mathbf{s}_i, \mathbf{s}_j)$
4. Constrained least squares (next slide) to minimize the distance between sample ($\tilde{\vartheta}$) and model (ϑ as a function of the B) spatial dependence
5. Order the terms by $v_l = \sum_{\mathbf{s}} B_l(\mathbf{s})$

Estimating the EBFs, $B_l(\mathbf{s})$

- The objective function to estimate the B_l is

$$\sum_{i<j} \left[\tilde{\vartheta}(\mathbf{s}_i, \mathbf{s}_j) - \vartheta(\mathbf{s}_i, \mathbf{s}_j) \right]^2$$

  where $\vartheta(\mathbf{s}_i, \mathbf{s}_j)$ is a function of B_l

- The EBFs must satisfy $B_l(\mathbf{s}) > 0$ and $\sum_l B_l(\mathbf{s}) = 1$ for all $\mathbf{s}$
- The solution is approximated by cycling through the sites and solving a series of constrained optimization problems
- Also gives estimate of $\hat{\alpha}$

Comparison with PCA

- Similarities to PCA:
  - Reduces dimension
  - Maps of $B_l(\mathbf{s})$ tell us about the most important spatial patterns
  - Captures a non-stationary spatial dependence structure
- Differences from PCA:
  - Basis functions are not orthonormal
  - Loadings are positive stable, not Gaussian
  - Loadings A_{lt} may not be independent
  - Computing A and B is not as simple as a few matrix operations

Bayesian implementation

- Given the basis function $B_l(\mathbf{s})$ and $\hat{\alpha}$ we can proceed with MCMC to estimate the remaining parameters

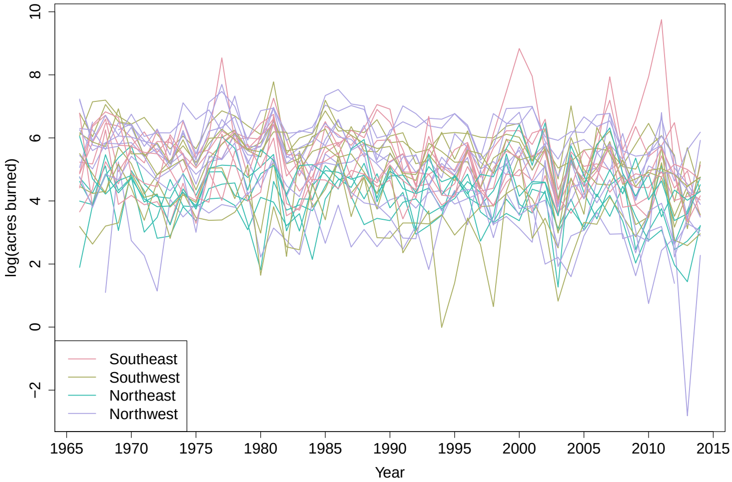
$$
\begin{aligned}
\mu_t(\mathbf{s}) &= \beta_{1,int}(\mathbf{s}) + \beta_{1,time}(\mathbf{s})t \\
\log[\sigma_t(\mathbf{s})] &= \beta_{2,int}(\mathbf{s}) + \beta_{2,time}(\mathbf{s})t \\
\xi_t(\mathbf{s}) &= \xi
\end{aligned}
$$

- Gaussian process priors on $\beta(\mathbf{s})$ terms
- We use cross-validation (quantile and Bries scores) to select L
- Alternative: select L so that $\sum_{l=1}^{L} v_l = 0.8$

Application 1: Wildfires in GA

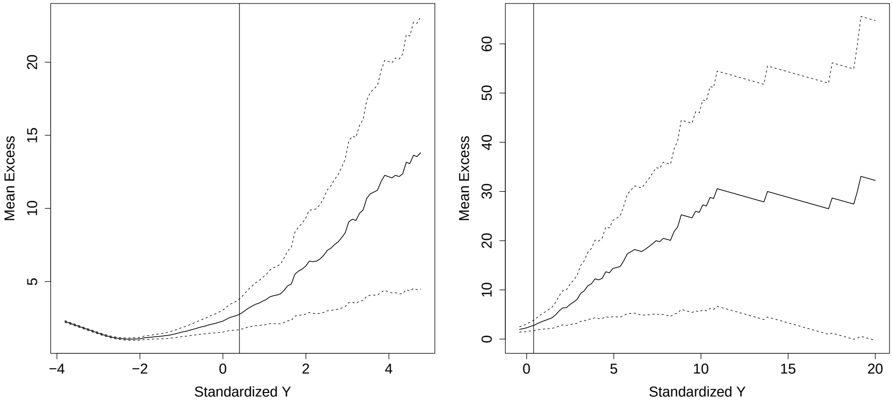
- The data are the number of acres burned by forest fires (wildfire) each year (1965–2014) in each county of Georgia
- We censor the data at the local 95^{th} percentile, $T(\mathbf{s})$
- The objectives are to map fire risk and determine if it is changing with time

Fire: Time series for each county



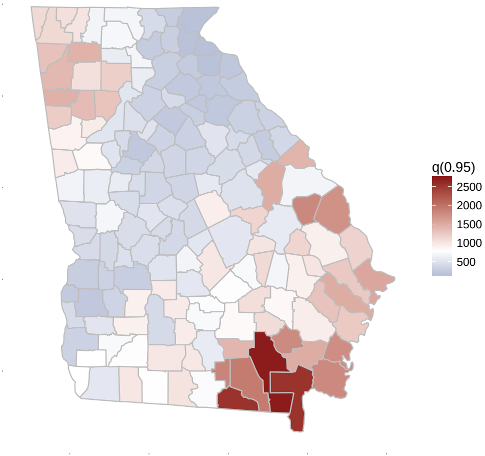
Time series of log(acres burned) color-coded by region.

Fire: Picking the threshold



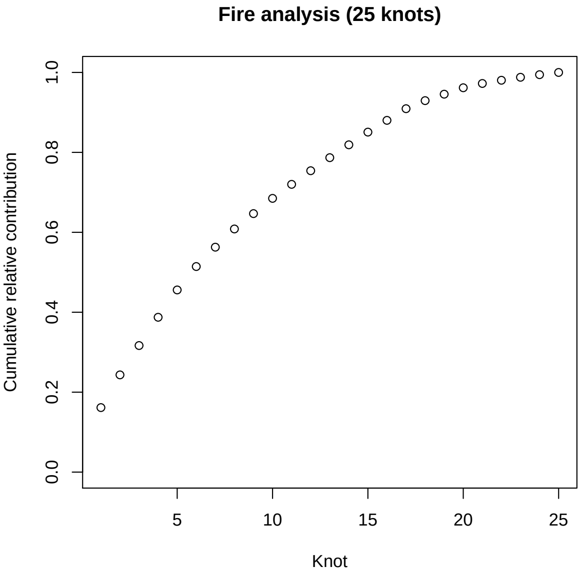
Mean residual life plot for fire data. Vertical line at sample 95th quantile.

Fire: 95th percentile by county, $T(s)$

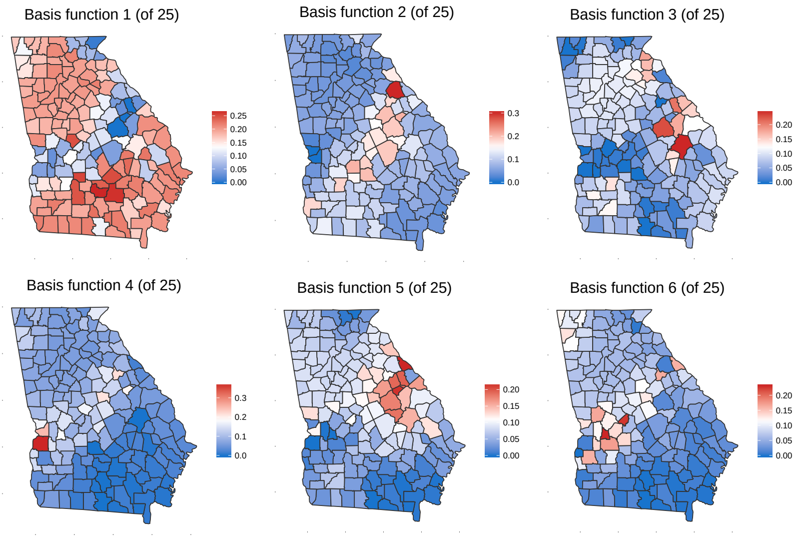


Spatially smoothed 95th percentile.

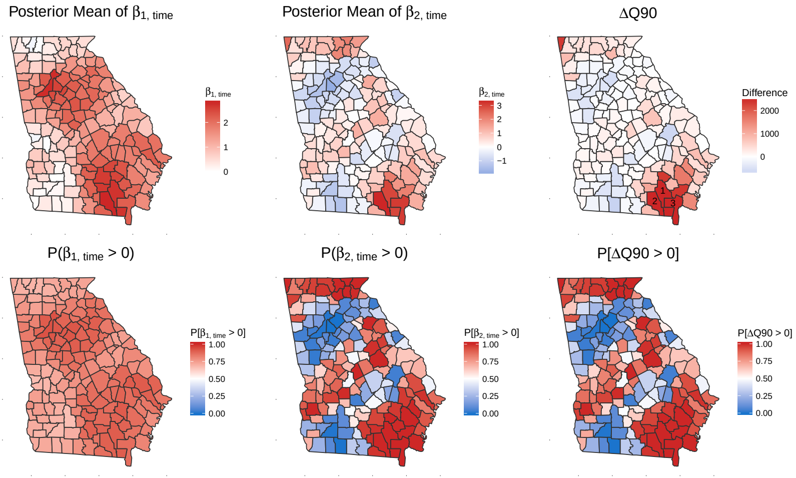
Fire: EBF weights, v_l



Fire: EBF's $B_l(s)$



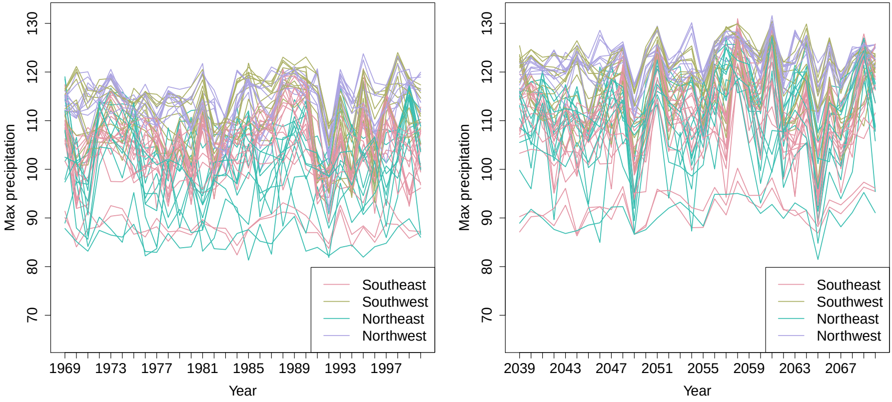
Fire: Posterior summaries ($L = 25$)



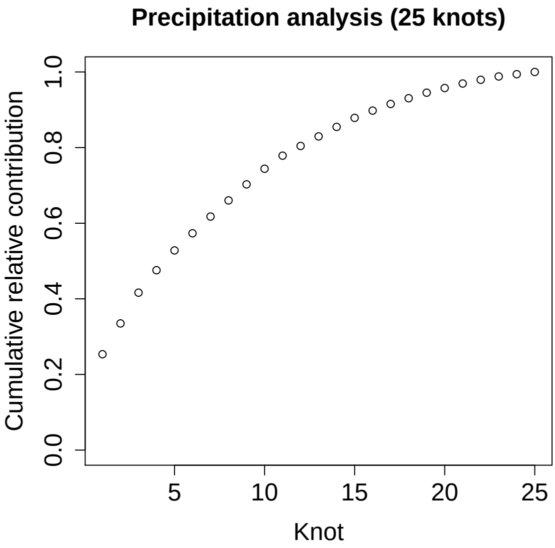
Application 2: NARCCAP climate model output

- Data consist of annual maximum precipitation at 697 grid cells in the Eastern US
- Model is run separately for 1969–2000 and 2039–2077
- The objective is to compare the extremes in the two climate periods
- We fit the same model as for the fire data except without censoring

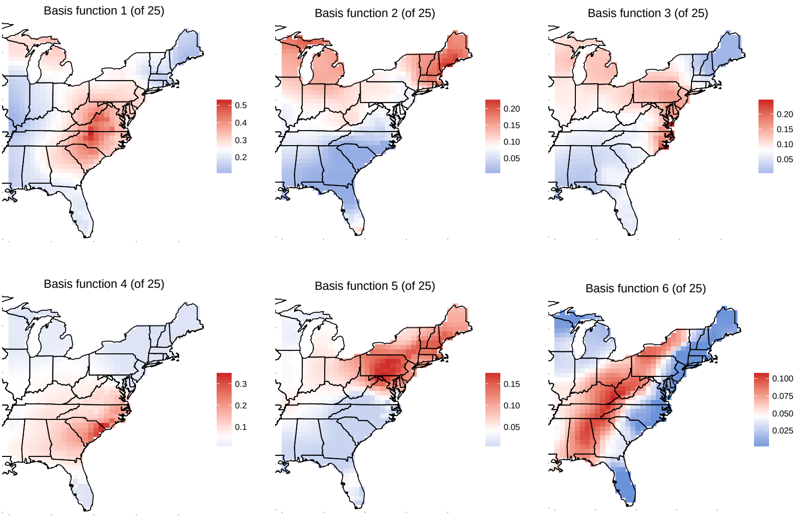
Climate model output for 1969



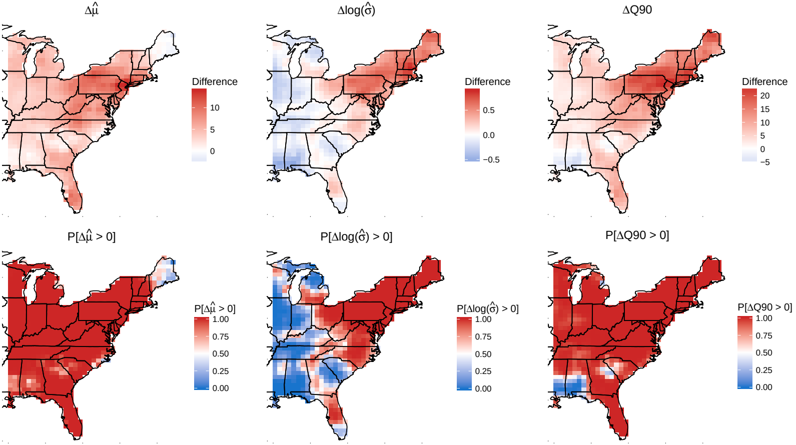
Precip: EBF weights, v_l



Precip: EBFs $B_l(s)$



Precip: Posterior summaries ($L = 25$)



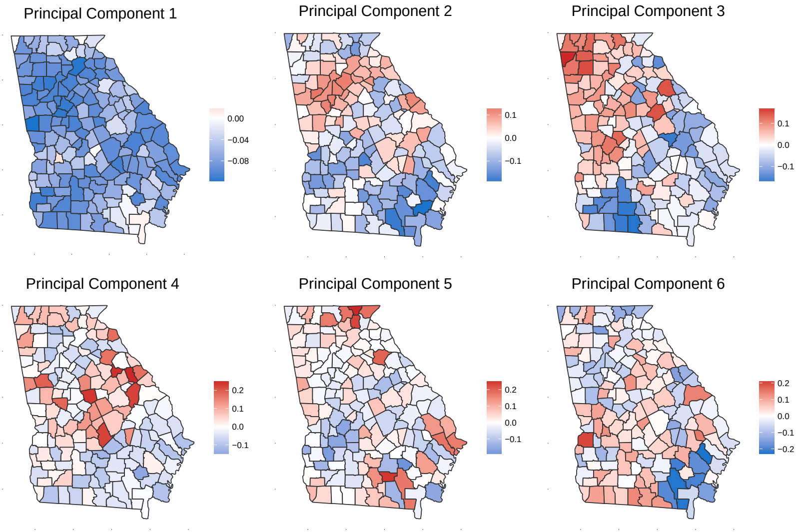
Results

- Comparison to principal components:
  - For fire data, EBFs and PCAs are different due to censoring
  - For precipitation data, EBFs and PCAs generally capture similar features
- EBF performs better than standardized Gaussian kernel functions when there is spatial dependence in the data
  - Fire: $\hat{\alpha} = 0.86$
  - Precipitation: $\hat{\alpha} = 0.28$

Questions

- Thank you for your attention.
- Questions?
- Acknowledgment: This work was funded by EPA STAR award R835228

Fire: PCs


Principal Component 1
0.00
−0.04
−0.08
Principal Component 2
0.1
0.0
−0.1
Principal Component 3
0.1
0.0
−0.1
Principal Component 4
0.2
0.1
0.0
−0.1
Principal Component 5
0.2
0.1
0.0
−0.1
Principal Component 6
0.2
0.1
0.0
−0.1
−0.2

Precipitation: PCs