Submitted to the Annals of Applied Statistics

EXPLORATION AND INFERENCE IN SPATIAL EXTREMES USING EMPIRICAL BASIS FUNCTIONS

BY SAMUEL A. MORRIS[*], BRIAN J. REICH[*] AND EMERIC THIBAUD[†]

North Carolina State University[*] *and Colorado State University*[†]

Abstract goes here

1. Introduction. The spatial Extreme Value Analysis (EVA) literature is expanding rapidly (Davison et al., 2012) to meet the demands of researchers to improve estimates of rare-event probabilities by borrowing information across space and to estimate the probability of extreme events occurring simultaneously at multiple locations. Environmental datasets commonly include observations from hundreds or thousands of locations, and advanced tools are required to explore and analyze these data. For Gaussian data, Principle Components Analysis (Everitt and Hothorn, 2008; PCA), also known as Empirically Orthogonal Functions (Hannachi et al., 2007; EOF), has proven to be a powerful tool to study correlation between spatial locations; understand the most important large-scale spatial features; and reduce the dimension of the problem to allow for simple computation even for massive datasets. Computation and exploration are arguably more difficult for EVA than Gaussian data, yet to our knowledge no tool analogous to spatial PCA has been developed for EVA.

In EVA, extremes are separated from the bulk of the distribution by either analyzing only points above a threshold or block maximums (Coles, 2001), e.g., the annual maximum of the daily precipitation. A natural spatial model for block maximum at several spatial locations is the max-stable process, which, under certain conditions, arises as the limit of the location-wise maximum of infinitely-many spatial processes (de Haan and Ferreira, 2006). Max-stable processes were also used to model spatial exceedances over a high threshold (Thibaud et al., 2013; Huser and Davison, 2014). De Haan (1984) showed that any max-stable process can be represented in terms of a countable number of spatial processes (e.g., stationary log Gaussian processes), and a finite truncation of this representation has been used for conditional simulation (Wang and Stoev, 2011). Fully-Bayesian analysis using max-stable processes is cumbersome for large data sets (Wadsworth and Tawn, 2014; Thibaud and Opitz, 2015). One option is to use non-max-stable models that retain extremal dependence such as the skew-t process in (Morris et al.,

Keywords and phrases: non-stationary data analysis, max-stable process, spectral representation, dimension reduction

under review). Alternatively, Reich and Shaby (2012) propose a low-rank method based on spatial kernel functions, and others have used pairwise (Padoan et al., 2010; Huser and Davison, 2014) and trivariate (Genton et al., 2011) likelihood methods for parameter estimation.

In this paper we propose an empirical basis function (EBF) approach that builds on a finite truncation of the spectral representation, and develops a method-of-moments estimator for the underlying spatial processes. Unlike PCA/EOFs, but similar to dictionary learning (Mairal et al., 2014) and non-negative matrix factorizations (Lee and Seung, 1999), the EBFs are not orthogonal. Nonetheless these spatial functions can be plotted for exploratory analysis to reveal important spatial trends. In addition to exploratory analysis, we show that the EBFs can be used for Bayesian inference on the marginal parameters at each location, modeling spatial dependence, and to test for covariate effects. By basing the spatial dependence on EBFs, the resulting spatial analysis does not require dubious assumptions such as stationarity. In addition, a Bayesian analysis for either block-maximum or point above a threshold is computationally feasible for large datsets because the entire spatial process is represented by a small number of basis functions.

The paper proceeds as follows. In Section 2 we present the low-rank model. Section 3 describes the algorithm used to estimate the spatial basis functions, and Section 4 describes the model fit using Markov chain Monte Carlo (MCMC) methods. In Section 5 we demonstrate the use of the EBFs for an analysis of precipitation data in the eastern U.S. Lastly in Section 6 we give some summary conclusions and a brief discussion of the findings.

2. Model. Let $Y_t(\mathbf{s})$ be the observation at spatial location $\mathbf{s}$ and time t. We temporarily drop the subscript t and describe the model for the process $Y(\mathbf{s})$ for a single time point, but return to the spatiotemporal notation in Section 3. To focus attention on the extreme values, we emphasize the statistical model for exceedances above a location-specific threshold $T(\mathbf{s})$. We begin by specifying a spatial model for the complete data $Y(\mathbf{s})$ and then use the censored likelihood defined by $T(\mathbf{s})$ for inference as described in Section 4. Although the model presented implements a censored likelihood, the model also can fit uncensored data (such as block-maxima) by setting $T(\mathbf{s}) = -\infty$.

Spatial dependence is captured by modeling $Y(\mathbf{s})$ as a max-stable process (de Haan and Ferreira, 2006). Max-stable processes have generalized extreme value (GEV; see Appendix A.1) marginal distribution. The GEV has three parameters: location $\mu(\mathbf{s})$; scale $\sigma(\mathbf{s})$; and shape $\xi(\mathbf{s})$. Spatial dependence is present both in the GEV parameters but also the standardized residual process

$$Z(\mathbf{s}) = \left\{1 + \frac{\xi(\mathbf{s})}{\sigma(\mathbf{s})}\left[Y(\mathbf{s}) - \mu(\mathbf{s})\right]\right\}^{1/\xi(\mathbf{s})}, \tag{1}$$

which has unit Fréchet (i.e., GEV with location, scale, and shape all equal one) marginal distribution for all $\mathbf{s}$.

Our objective is to identify a low-rank model for the spatial dependence of $Z(\mathbf{s})$. De Haan (1984; Chapter 9) show that any max-stable process can be written as

$$Z(\mathbf{s}) = \bigvee_{l=1}^{\infty} B(\mathbf{s}, \mathbf{k}_l) A_l \tag{2}$$

where the functions $B(\mathbf{s}, \mathbf{k}_l)$ satisfy $B(\mathbf{s}, \mathbf{k}_l) > 0$ for all $\mathbf{s}$ and $\int B(\mathbf{s}, \mathbf{k}_l)\, d\mathbf{k}_l = 1$ for all $\mathbf{s}$, and $(\mathbf{k}_l, A_l)$ for $l = 1, \ldots, \infty$ are a Poisson process with intensity measure $dA\ d\mathbf{k}/A^2$. In many representations of max-stable process, such as Smith (1990) and Reich and Shaby (2012), the $\mathbf{k}_l$ are spatial locations that represent the center of process l; however, in our proposed method the basis functions are not associated with one particular location and so to simplify notation we let $B_l(\mathbf{s}) = B(\mathbf{s}; \mathbf{k}_l)$

To arrive at a low-rank model, we assume there are a finite and known number of spatial basis functions $B_1(\mathbf{s}), \ldots, B_L(\mathbf{s})$ that explain the important spatial variation in the process. As in de Haan's expansion, the basis functions are restricted so that $B_l(\mathbf{s}) > 0$ and $\sum_{l=1}^{L} B_l(\mathbf{s}) = 1$ for all $\mathbf{s}$. Because it is unrealistic to assume that realizations of Z are exactly functions of L basis functions, we include independent error variables $\epsilon(\mathbf{s})$ to capture variation not explained by the $B_l(\mathbf{s})$. We follow Reich and Shaby (2012) and decompose $Z(\mathbf{s})$ as $Z(\mathbf{s}) = \theta(\mathbf{s})\varepsilon(\mathbf{s})$ where $\theta(\mathbf{s})$ is a spatial process and $\varepsilon(\mathbf{s}) \overset{\text{iid}}{\sim} \text{GEV}(1, \alpha, \alpha)$ is independent error. The spatial component is

$$\theta(\mathbf{s}) = \left(\sum_{l=1}^{L} B_l(\mathbf{s})^{1/\alpha} A_l \right)^{\alpha}. \tag{3}$$

If $B_l(\mathbf{s}) > 0$, $\sum_{l=1}^{L} B_l(\mathbf{s}) = 1$ for all $\mathbf{s}$, and the A_l have positive stable (PS; Appendix A.2) distribution $A_l \overset{\text{iid}}{\sim} \text{PS}(\alpha)$, then $Z(\mathbf{s})$ is max-stable and has unit Fréchet marginal distributions.

Extremal spatial dependence for max-stable processes can be summarized by the extremal coefficient (Schlather and Tawn, 2003; EC) $\vartheta(\mathbf{s}_1, \mathbf{s}_2) \in [1, 2]$, where

$$\text{Prob}[Z(\mathbf{s}_1) < c, Z(\mathbf{s}_2) < c] = \text{Prob}[Z(\mathbf{s}_1) < c]^{\vartheta(\mathbf{s}_1, \mathbf{s}_2)}. \tag{4}$$

For the PS random effects model the EC has the form

$$\vartheta(\mathbf{s}_1, \mathbf{s}_2) = \sum_{l=1}^{L} \left[B_l(\mathbf{s}_1)^{1/\alpha} + B_l(\mathbf{s}_2)^{1/\alpha} \right]^{\alpha}. \tag{5}$$

In particular, $\vartheta(\mathbf{s},\mathbf{s}) = 2^\alpha$ for all $\mathbf{s}$.

3. Estimating the basis functions. To estimate the extremal coefficient function, we consider the process at n_s spatial locations $\mathbf{s}_1,\dots,\mathbf{s}_{n_s}$ and n_t times $t = 1,\dots,n_t$. The basis functions are fixed over time, but the random effects and errors are independent over time. That is

$$Z_t(\mathbf{s}) = \theta_t(\mathbf{s})\epsilon_t(\mathbf{s}) \quad \text{where} \quad \theta_t(\mathbf{s}) = \left(\sum_{l=1}^{L} B_l(\mathbf{s})^{1/\alpha} A_{lt}\right)^{\alpha}, \tag{6}$$

$A_{lt} \overset{\text{iid}}{\sim} \text{PS}(\alpha)$, and $\epsilon_t(\mathbf{s}) \overset{\text{iid}}{\sim} \text{GEV}(1,\alpha,\alpha)$. Denote $Y_t(\mathbf{s}_i) = Y_{it}$, $B_l(\mathbf{s}_i) = B_{il}$, $T(\mathbf{s}_i) = T_i$, and $\vartheta(\mathbf{s}_i,\mathbf{s}_j) = \vartheta_{ij}$.

In this section we develop an algorithm to estimate the spatial dependence parameter α and the $n_s \times L$ matrix $\mathbf{B} = \{B_{il}\}$. Our algorithm has the following steps:

(1) Obtain an initial estimate of the extremal coefficient for each pair of locations, $\hat{\vartheta}_{ij}$.
(2) Spatially smooth these initial estimates $\hat{\vartheta}_{ij}$ using kernel smoothing to obtain $\tilde{\vartheta}_{ij}$.
(3) Estimate the spatial dependence parameters by minimizing the difference between model-based coefficients, ϑ_{ij}, and smoothed coefficients, $\tilde{\vartheta}_{ij}$.

The first-stage estimates are obtained from the estimator of Schlather and Tawn (2003) using the `fmadogram` function in the `SpatialExtremes` (Ribatet, 2015) package of `R` (R Core Team, 2016). Assuming the true EC is smooth over space, the initial estimates $\hat{\vartheta}_{ij}$ can be improved by smoothing. Let

$$\tilde{\vartheta}_{ij} = \frac{\sum_{u=1}^{n_s}\sum_{v=1}^{n_s} w_{iu}w_{jv}\hat{\vartheta}_{uv}}{\sum_{u=1}^{n_s}\sum_{v=1}^{n_s} w_{iu}w_{jv}}, \tag{7}$$

where $w_{iu} = \exp[-(||\mathbf{s}_i - \mathbf{s}'_u||/\phi)^2])$ is the Gaussian kernel function with bandwidth ϕ. The elements $\hat{\vartheta}_{ii}$ do not contribute any information as $\hat{\vartheta}_{ii} = 1$ for all i by construction. To eliminate the influence of these estimates we set $w_{ii} = 0$. However, this approach does give imputed values $\tilde{\vartheta}_{ii}$, which provide information about small-scale spatial variability.

The dependence parameters B_{lt} and α are estimated by comparing estimates $\tilde{\vartheta}_{ij}$ with the model-based values ϑ_{ij}. For all i, $\vartheta_{ii} = 2^\alpha$, and therefore we set α to $\hat{\alpha} = \log_2\left(\sum_{i=1}^{n_s}\tilde{\vartheta}_{ii}/n_s\right)$. Given $\alpha = \hat{\alpha}$, it remains to estimate $\mathbf{B}$. Similarly to

Smith (1990) for a stationary max-stable process, we use squared-error loss, so the estimate $\hat{\mathbf{B}}$ is the minimizer of

$$\sum_{i<j}\left(\tilde{\vartheta}_{ij}-\vartheta_{ij}\right)^2=\sum_{i<j}\left(\tilde{\vartheta}_{ji}-\sum_{l=1}^{L}\left[B_{il}^{1/\hat{\alpha}}+B_{jl}^{1/\hat{\alpha}}\right]^{\hat{\alpha}}\right)^2 \tag{8}$$

under the restrictions that $B_{il} \geq 0$ for all i and l and $\sum_{l=1}^{L} B_{il} = 1$ for all i. Since the minimizer of (8) does not have a closed form, we use block coordinate descent to obtain $\hat{\mathbf{B}}$. We cycle through spatial locations and update the vectors $\left(\hat{B}_{i1},\dots,\hat{B}_{iL}\right)$ conditioned on the values for the other location and repeat until convergence. At each step, we use the restricted optimization routine in the `R` function `optim`. This algorithm gives estimates of the B_{il} at the n_s data locations, but is easily extended to all s for spatial prediction. The kernel smoothing step ensures that the estimates for $\hat{B}_{il}$ are spatially smooth, and thus interpolation of the $\hat{B}_{il}$ gives spatial functions $\hat{B}_l(\mathbf{s})$.

These functions provide useful exploratory data analysis techniques. Maps of $\hat{B}_l(\mathbf{s})$ show important spatial features in the extremal dependence. Furthermore, they allow for a non-stationary spatial dependence structure. The relative contribution of each term can be measured by

$$v_l = \frac{1}{n_s}\sum_{i=1}^{n_s}\hat{B}_{il}. \tag{9}$$

Since $\sum_{l=1}^{L}\hat{B}_{il} = 1$ for all i, we have $\sum_{l=1}^{L} v_l = 1$. Therefore, terms with large v_l are the most important. The order of the terms is arbitrary, and so we reorder the terms so that $v_1 \geq \cdots \geq v_L$.

4. Bayesian implementation details. For our data analysis in Section 5 we allow the GEV location and scale parameters, denoted $\mu_t(\mathbf{s})$ and scale $\sigma_t(\mathbf{s})$ respectively, to vary with space and time. The model we choose is as follows

$$\mu_t(\mathbf{s}) = \beta_{1,\text{int}}(\mathbf{s}) + \beta_{1,\text{time}}(\mathbf{s})t \tag{10}$$

$$\log[\sigma_t(\mathbf{s})] = \beta_{2,\text{int}}(\mathbf{s}) + \beta_{2,\text{time}}(\mathbf{s})t \tag{11}$$

where

$$\begin{aligned}
&\beta_{1,\text{int}}(\mathbf{s}) \sim \text{N}(\mu_{1,\text{int}}\mathbf{1}, \sigma^2_{1,\text{int}}\boldsymbol{\Sigma}) &\qquad &\beta_{1,\text{time}}(\mathbf{s}) \sim \text{N}(\mu_{1,\text{time}}\mathbf{1}, \sigma^2_{1,\text{time}}\boldsymbol{\Sigma})\\
&\beta_{2,\text{int}}(\mathbf{s}) \sim \text{N}(\mu_{2,\text{int}}\mathbf{1}, \sigma^2_{2,\text{int}}\boldsymbol{\Sigma}) &\qquad &\beta_{2,\text{time}}(\mathbf{s}) \sim \text{N}(\mu_{2,\text{time}}\mathbf{1}, \sigma^2_{2,\text{time}}\boldsymbol{\Sigma})
\end{aligned} \tag{12}$$

are Gaussian process priors and $\mathbf{\Sigma}$ is an exponential spatial correlation matrix obtained from $\rho(h) = \exp\left\{-\frac{h}{\phi}\right\}$ where $h = ||\mathbf{s}_1 - \mathbf{s}_2||$ is the Euclidean distance between sites $\mathbf{s}_1$ and $\mathbf{s}_2$. The GEV shape parameter ξ is held constant over space and time because this parameter is challenging to estimate. Collectively, let the marginal GEV parameters at location i and time t be $\Theta_{it} = \{\mu_{it}, \sigma_{it}, \xi\}$ where $\mu_{it} = \mu_t(\mathbf{s}_i)$ and $\sigma_{it} = \sigma_t(\mathbf{s}_i)$.

As shown in Reich and Shaby (2012), the uncensored responses $Y_t(\mathbf{s})$ are conditionally independent given the spatial random effects, with conditional distribution

$$Y_{it}|\theta_{it}, \Theta_{it} \overset{\text{ind}}{\sim} \text{GEV}(\mu^*_{it}, \sigma^*_{it}, \xi^*), \tag{13}$$

where $\mu^*_{it} = \mu_{it} + \frac{\sigma_{it}}{\xi}(\theta^{\xi}_{it} - 1)$, $\sigma^*_{it} = \alpha\sigma_{it}\theta^{\xi}_{it}$, and $\xi^* = \alpha\xi$. Therefore, the conditional likelihood conveniently factors across observations; marginalizing over the random effect θ_{it} induces extremal spatial dependence. To focus on the extreme values above the local threshold T_i, we use the censored likelihood

$$d(y; \theta_{it}, \Theta_{it}, T_i) = \begin{cases} F(y; \mu^*_{it}, \sigma^*_{it}, \xi^*) & y \le T_i \\ f(y; \mu^*_{it}, \sigma^*_{it}, \xi^*) & y > T_i, \end{cases} \tag{14}$$

where F and f are the GEV distribution and density functions, respectively, defined in Appendix A.1.

In summary, given the estimates of α and $\mathbf{B}$, the hierarchical model is

$$\begin{aligned} Y_{it}|\theta_{ij} &\overset{\text{indep}}{\sim} d(y; \theta_{it}, \Theta_{it}, T_i) \\ \theta_{it} &= \left(\sum_{l=1}^{L} \hat{B}^{1/\hat{\alpha}}_{il} A_{lt}\right)^{\hat{\alpha}} \quad \text{where} \quad A_{lt} \overset{\text{iid}}{\sim} PS(\hat{\alpha}) \\ \mu_{it} &= \beta_{1,\text{int}}(\mathbf{s}_i) + \beta_{1,\text{time}}(\mathbf{s}_i)t \\ \log(\sigma_{it}) &= \beta_{2,\text{int}}(\mathbf{s}) + \beta_{2,\text{time}}(\mathbf{s})t. \end{aligned} \tag{15}$$

We estimate parameters $\Theta = \{A_{lt}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \xi\}$ using Markov chain Monte Carlo methods. We use a Metropolis-Hastings algorithm to update the model parameters with random walk candidate distributions for all parameters. The PS density is challenging to evaluate as it does not have a closed form. One technique to avoid this complication is to incorporate auxiliary random variables (Stephenson, 2009), but we opt for a numerical approximation to the integral as described in Appendix A.2. The hyperparameters $\mu_{1,\text{int}}$, $\mu_{1,\text{time}}$, $\mu_{2,\text{int}}$, $\mu_{2,\text{time}}$ and $\sigma^2_{1,\text{int}}$, $\sigma^2_{1,\text{time}}$, $\sigma^2_{2,\text{int}}$, $\sigma^2_{2,\text{time}}$ are updated using Gibbs sampling since their prior distributions are conjugate.

The first-stage estimate of the extremal coefficients has two tuning parameters: the kernel bandwidth ϕ, and the number of terms L. In Section 5 we explore a few possibilities for L and discuss sensitivity to this choice. The second-stage Bayesian analysis requires selecting thresholds $T_i, \dots, T_{n_s}$. For this we use spatially smoothed sample quantiles. That is, we set T_i to the 0.95 quantile of the Y_{it} and its five nearest neighbors.

5. Data analysis.

In this section, we illustrate our method with an analysis of block maxima precipitation data in the eastern U.S. We compare our method with results from a more naïve approach that uses standardized Gaussian kernels for the spatial basis functions.

5.1. *Gaussian kernel basis functions.*

To provide a comparison of our model with another approach, we also fit a model that uses standardized Gaussian kernels for the spatial basis functions (Reich and Shaby, 2012). In this method, Reich and Shaby introduce a set of $\mathbf{k}_1, \dots, \mathbf{k}_L$ spatial knots and use standardized Gaussian kernel functions (GSK; see Appendix A.3) instead of using EBFs for the $\hat{B}_l(\mathbf{s})$. For the comparison between EBF and GSK methods, we use the same number of basis functions. We obtain estimates of the kernel bandwidth $\hat{\rho}$ and spatial dependence $\hat{\alpha}$, using the same least squares minimization as with the EBF method, and treat these as fixed in the MCMC.

5.2. *Analysis of annual precipitation.*

We also conduct an analysis of the precipitation data presented in Reich and Shaby (2012). The data are climate model output from the North American Regional Climate Change Assessment Program (NARCCAP). This data consists of $n_s = 697$ grid cells at a 50km resolution in the eastern US, and includes historical data (1969 – 2000) as well as future conditions (2039 – 2070). Because the data are block maxima, we set $T = -\infty$.

For this dataset, to estimate the EBFs, we use the combined current and future data. The first six EBFs for the combined data along with the cumulative sum of the contributions for $v_1, \dots, v_{25}$ are given in Figure 2. As a comparison, we provide the first six principal components of the fire data along with the cumulative sum of the first 25 eigenvalues in Appendix B. For the precipitation data, we run the MCMC for 25,000 iterations using a burnin period of 15,000 iterations. We consider models fit with both EBF and GSK, and fit the model using $L = 5, 10, \dots, 40$. Timing for each setting of L is given in Table 2 for 1,000. These timings come from a single core of an Intel Core i7-5820K Haswell-E processor, using the OpenBLAS optimized BLAS library (`http://www.openblas.net`).

5.3. *Precipitation results.*

We use 5-fold cross-validation to assess the predictive performance of a model. For each method, we randomly select 80% of the observations across counties and years to be used as a training set to fit the model. The remaining 20% of sites and years are withheld for testing model predictions.

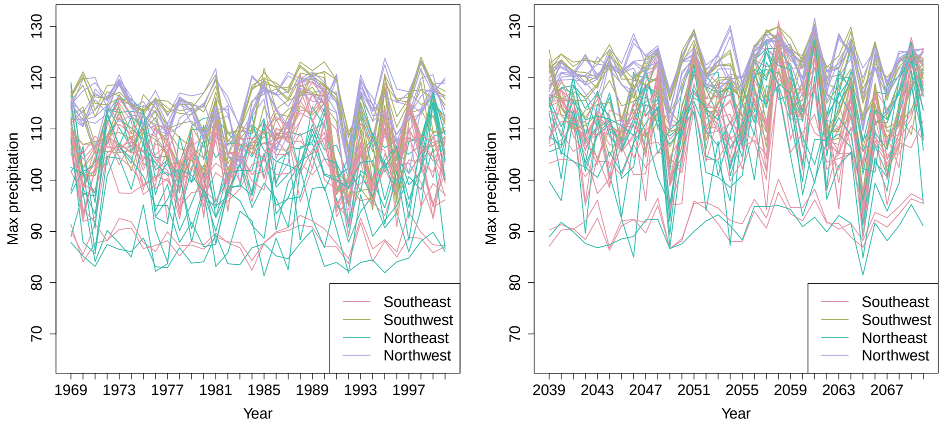

FIG 1. *Time series of yearly max precipitation for current (1969 – 2000) (left). Time series of yearly max precipitation for future (2039 – 2070) (right).*

To assess the predictions for the test set, we use quantile scores and Brier scores (Gneiting and Raftery, 2007). The quantile score (QS) for quantile level q^* is given by $2\{I[Y(\mathbf{s}) > \hat{q}(\mathbf{s})] - q^*]\}\{\hat{q}(\mathbf{s}) - Y(\mathbf{s})\}$ where $\hat{q}(\mathbf{s})$ is the estimated q^*th quantile at site $\mathbf{s}$, and $I[\cdot]$ is an indicator function. The Brier score (BS) for predicting an exceedance of a level c at site $\mathbf{s}$ is given by $\{I[Y(\mathbf{s}) > c] - \hat{P}[Y(\mathbf{s}) > c]\}^2$. For both of these methods, a lower score indicates a better fit. We also present the Continuous Rank Probability Score (CRPS) (Gneiting and Raftery, 2007; see equation (21)) and mean absolute deviation (MAD). The Brier and quantile scores for the current and future precipitation data analysis are given in Table 2. For these data, we observe more variation in the scores across the number of basis functions and generally an advantage in using EBF over GSK. When using the EBFs, the estimate of residual dependence for the precipitation data is $\hat{\alpha} = 0.280$ ($\alpha = 1$ is residual independence).

Based on the cross-validation results, we run a full analysis using all of the data with $L = 25$ and EBF. Figure 3 gives posterior summaries for three quantities of interest. Because we have two separate time periods, current and future, we look at the differences between the estimates for $\hat{\mu}$, $\log(\hat{\sigma})$, and $Q90$ between $t_1 = 2000$ and $t_2 = 2070$.

We construct the posterior distribution of the the estimated 90th quantile $Q90_{i,t}$ using the GEV parameters as follows. Let $Q90_{i,t}^{(j)}$ be the estimated 90th quantile at

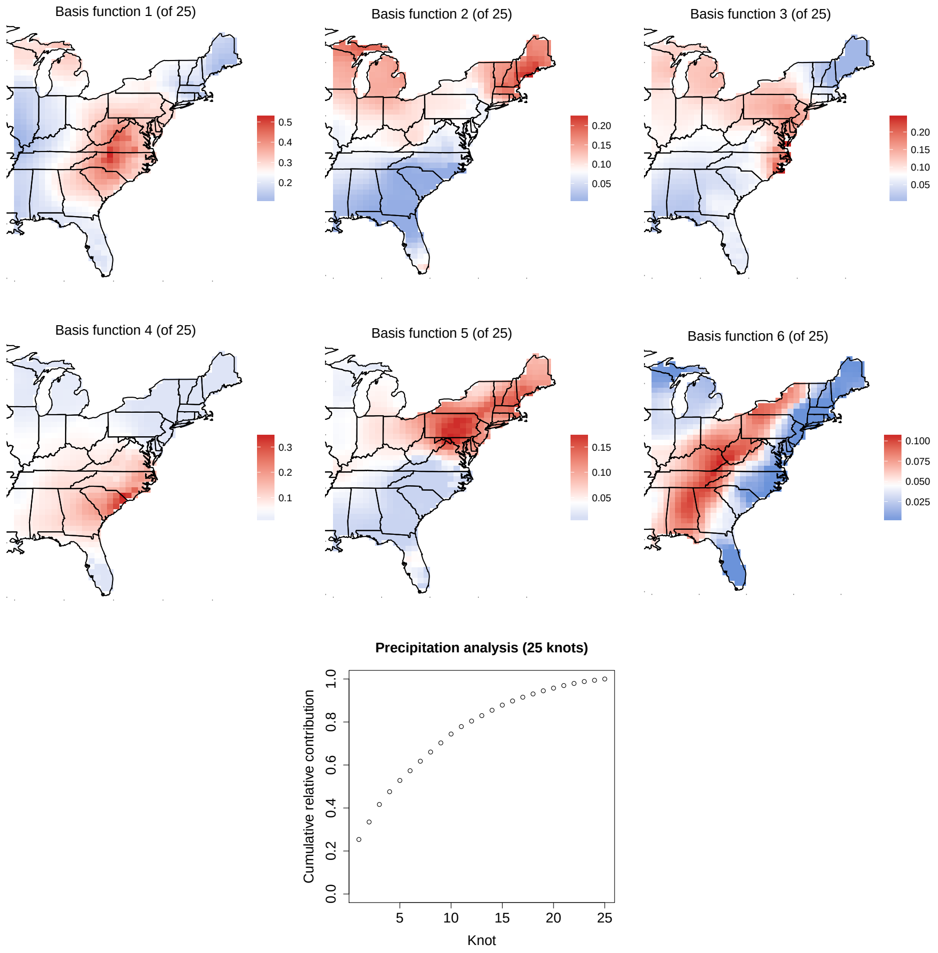

FIG 2. *First six EBFs for the combined precipitation data and the cumulative sum of contributions* $v_1, \ldots, v_{25}$.

site i for time t at iteration j. We first compute

$$\begin{aligned} \mu_{i,t}^{(j)} &= \beta_{1,\text{int},i}^{(j)} + \beta_{1,\text{time},i}^{(j)} t \\ \log\left(\sigma_{i,t}^{(j)}\right) &= \beta_{2,\text{int},i}^{(j)} + \beta_{2,\text{time},i}^{(j)} t. \end{aligned} \tag{16}$$

Let $Q90_{i,t}^{(j)} = \mu_{i,t}^{(j)} + \sigma_{i,t}^{(j)} F^{-1}\left(0.90, \xi^{(j)}\right)$ where $F^{-1}(q, \xi)$ is the inverse distribution function of the $\text{GEV}(1, 1, \xi)$ distribution evaluated at the qth quantile. We summarize $hat\mu_{i,t}$, $\log(\hat{\sigma})_{i,t}$, and $Q90_{i,t}$ with the respective posterior means. To obtain the posterior probability of seeing an increase over time, we take the posterior distributions of each parameter of interest at the two time points. Consider two time points $t_1 < t_2$. Let $\varphi_t^{(j)}$ be the parameter of interest at iteration j and time t. We then take the posterior mean of $I\left[\varphi_{t_2}^{(j)} > \varphi_{t_1}^{(j)}\right]$, for the posterior probability of seeing an increase in φ from time t_1 to t_2. We plot $\Delta\mu = \hat{\mu}_{2070} - \hat{\mu}_{2000}$, $\Delta\log(\hat{\sigma}) = \log(\hat{\sigma})_{2070} - \log(\hat{\sigma})_{2000}$, and $\Delta Q90 = \hat{q}(0.90)_{2070} - \hat{q}(0.90)_{2000}$, and the estimated probabilites that each are positive.

The results seem to suggest that the strength of extreme rain events will increases between 2000 and 2070 as well as greater variability in the northeast region of the U.S. as well as Ohio and parts of the south. There very strong evidence to suggest that most of the eastern U.S. should expect to see an increase in the 10-year return level between 2000 and 2070. Exceptions to this trend appear in southern parts of Alabama and Mississippi and regions on the border between South Carolina and Georgia which will likely experience a decrease in the 10-year return level.

6. Discussion. In this paper we have proposed new empirical basis functions for a data-driven low-rank approximation to a max-stable process. The basis functions provide researchers with an exploratory data analysis tool to explore maps of extremal dependence over space. The functions can also be used as inputs to an MCMC algorithm for inference and predicitons over space. The results from the data analysis provide evidence to suggest that in the presence of strong spatial dependence as with the precipitation data, the empirical basis functions show an improvement in quantile scores over using knots and standardized Gaussian kernel functions without an increase in the amount of time for computing.

We have used the EBF for exploratory analysis and Bayesian inference. Another possibility is to use the methods to reduce the data under consideration from the actual responses to loadings A_{kt}. That is, given the EBF, one could obtain estimates of the A_{kt} using a separate maximum likelihood estimation for each time point. Time series of the estimated A_{kt} may be used as a fast and simple method to study large-scale spatiotemporal trends.

TABLE 1
Average Brier scores ($\times 100$), average quantile scores for $q(0.95)$ and $q(0.99)$, CRPS, MAD, and time (in minutes) for 1,000 iterations for current precipitation analysis.

		BS ($\times$100)		QS				
L	Process	$q(0.95)$	$q(0.99)$	$q(0.95)$	$q(0.99)$	CRPS	MAD	Time
2	EBF	4.189	1.194	0.842	0.244	2.297	3.176	4.85
	GSK	4.144	1.193	0.885	0.250	2.388	3.312	4.88
3	EBF	3.973	1.105	0.793	0.222	2.108	2.901	4.96
	GSK	4.015	1.070	0.804	0.225	2.142	2.952	4.97
5	EBF	3.623	1.022	0.737	0.207	1.947	2.666	5.18
	GSK	3.765	1.016	0.756	0.210	1.941	2.658	5.19
10	EBF	3.691	1.192	0.727	0.212	1.818	2.455	5.69
	GSK	3.813	1.491	0.803	0.256	1.771	2.366	5.71
15	EBF	3.815	1.529	0.781	0.246	1.791	2.394	6.27
	GSK	3.898	1.652	0.832	0.272	1.722	2.267	6.30
20	EBF	3.858	1.651	0.788	0.252	1.740	2.296	6.86
	GSK	3.908	1.715	0.848	0.281	1.689	2.204	6.87
25	EBF	4.080	1.776	0.802	0.255	1.754	2.298	7.43
	GSK	3.906	1.760	0.852	0.284	1.676	2.177	7.46
30	EBF	4.029	1.827	0.794	0.257	1.716	2.235	8.00
	GSK	3.965	1.799	0.859	0.288	1.670	2.160	8.00
35	EBF	4.107	1.850	0.801	0.257	1.730	2.246	8.56
	GSK	3.959	1.799	0.860	0.290	1.656	2.139	8.61
40	EBF	4.063	1.889	0.808	0.261	1.748	2.259	9.14
	GSK	4.012	1.841	0.868	0.293	1.673	2.160	9.17

TABLE 2

Average Brier scores ($\times 100$), average quantile scores for $q(0.95)$ and $q(0.99)$, CRPS, MAD, and time (in minutes) for 1,000 iterations for future precipitation analysis.

		BS ($\times 100$)		QS				
L	Process	$q(0.95)$	$q(0.99)$	$q(0.95)$	$q(0.99)$	CRPS	MAD	Time
2	EBF	3.796	1.148	0.840	0.237	2.156	2.947	4.82
	GSK	3.470	1.142	0.850	0.237	2.145	2.925	4.82
3	EBF	3.477	1.112	0.785	0.220	2.008	2.747	4.96
	GSK	3.075	1.039	0.781	0.225	1.963	2.663	4.96
5	EBF	3.350	1.035	0.724	0.204	1.852	2.526	5.16
	GSK	3.219	1.046	0.736	0.209	1.846	2.500	5.15
10	EBF	3.136	1.088	0.692	0.203	1.684	2.257	5.69
	GSK	3.020	1.118	0.710	0.208	1.668	2.221	5.72
15	EBF	3.219	1.141	0.697	0.205	1.659	2.207	6.28
	GSK	3.078	1.222	0.739	0.226	1.641	2.162	6.32
20	EBF	3.159	1.280	0.703	0.211	1.619	2.133	6.87
	GSK	3.175	1.369	0.775	0.242	1.634	2.135	6.91
25	EBF	3.273	1.336	0.723	0.221	1.649	2.159	7.45
	GSK	3.198	1.467	0.802	0.257	1.625	2.103	7.47
30	EBF	3.399	1.555	0.771	0.243	1.645	2.125	8.03
	GSK	3.320	1.609	0.849	0.278	1.648	2.113	8.04
35	EBF	3.433	1.592	0.755	0.239	1.647	2.121	8.62
	GSK	3.378	1.668	0.866	0.285	1.644	2.097	8.60
40	EBF	3.529	1.679	0.791	0.252	1.681	2.145	9.16
	GSK	3.418	1.690	0.878	0.290	1.656	2.111	9.19

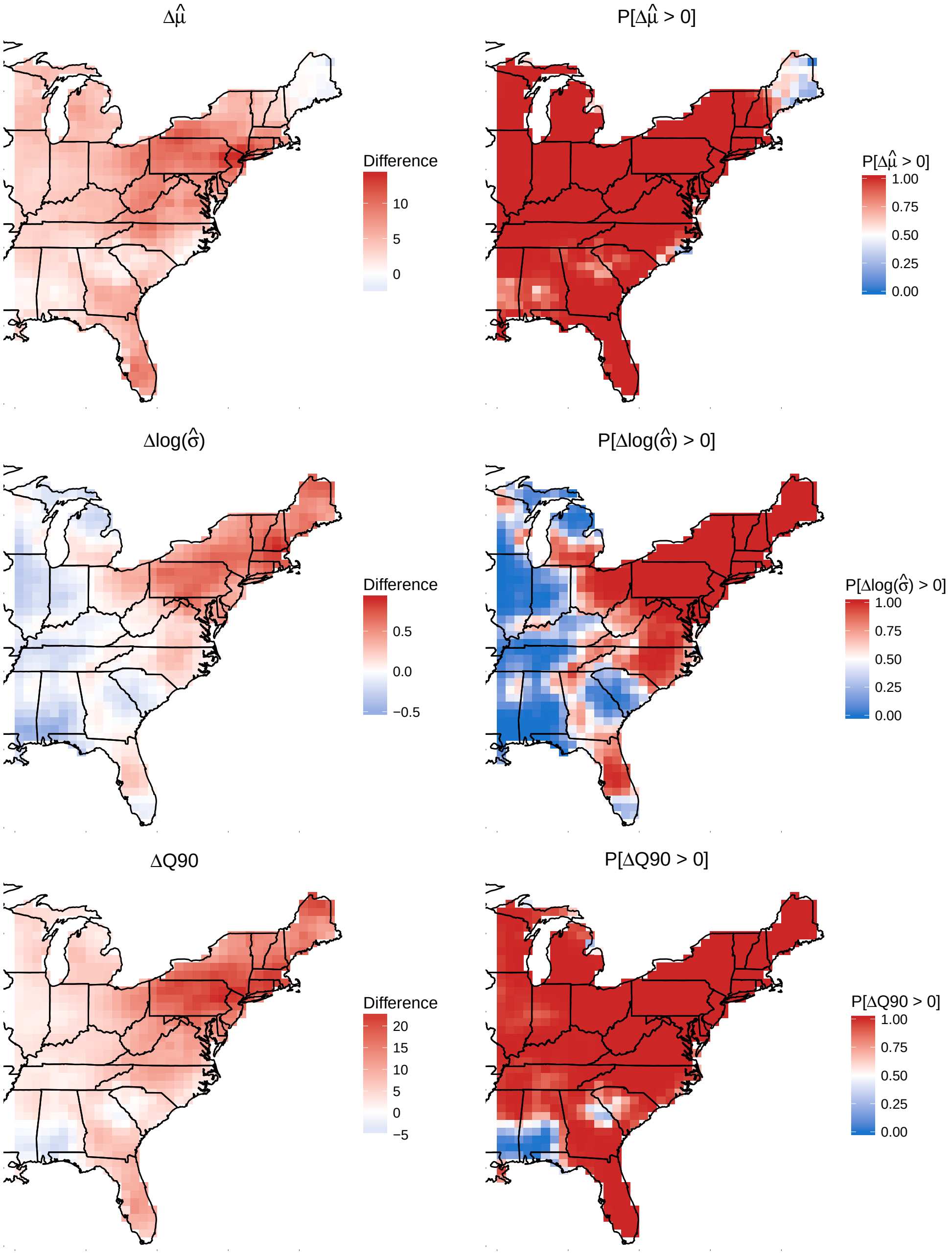

FIG 3. *Posterior mean of $\Delta\mu$ (top left), posterior mean of $\Delta\log(\sigma)$ (middle left), estimate of $\Delta Q90$ (bottom left), $P[\Delta\mu > 0]$ (top right), $P[\Delta\log(\sigma) > 0]$ (middle right), and $P[\Delta Q90 > 0]$ between 2000 and 2070 for precipitation data using EBF.*

Acknowledgements. The authors would like to acknowledge Dan Cooley for his helpful suggestions on the manuscript. The authors' work was partially supported by grants from the Department of the Interior (14-1-04-9), National Institutes of Health (R21ES022795-01A1), the US Environmental Protection Agency (R835228), the National Science Foundation (1107046).

APPENDIX A: APPENDICES

A.1. Extreme value distributions.

The cumulative distribution function for the GEV is $F(y) = \exp\{-t(y)\}$ where

$$t(y) = \begin{cases} \left[1 + \xi \dfrac{y-\mu}{\sigma}\right]^{-1/\xi}, & \xi \neq 0 \\ \exp\left\{-\dfrac{y-\mu}{\sigma}\right\}, & \xi = 0. \end{cases} \tag{17}$$

The probability density function for the GEV is given by $f(y) = \dfrac{1}{\sigma} t(x)^{\xi+1} \exp\{-t(y)\}$ where $t(y)$ is defined in (17).

A.2. Grid approximation to PS density.

The PS(α) density can be challenging to use because it does not have a closed form. From Section 2 of (Stephenson, 2009), the density can be expressed as

$$g_1(A) = \int_0^1 g_1(A, B)\ \mathrm{d}B, \tag{18}$$

where

$$g_1(A, B) = \frac{\alpha}{1-\alpha}\left(\frac{1}{A}\right)^{1/1-\alpha} c(\pi B) \exp\left\{-\left(\frac{1}{A}\right)^{\alpha/(1-\alpha)} c(\pi B)\right\}, \tag{19}$$

with

$$c(\psi) = \left[\frac{\sin(\alpha\psi)}{\sin(\psi)}\right]^{1/(1-\alpha)} \frac{\sin[(1-\alpha)\psi]}{\sin(\alpha\psi)}. \tag{20}$$

Stephenson (2009) presents an auxiliary variable technique to deal with the integral in the density function, but we opt to numerically evaluate the integral because it is only one-dimensional. To evaluate the integral, we use 50 evenly spaced quantiles of a $\text{Beta}(0.5, 0.5)$ distribution as the midpoints $B_1, \ldots, B_{50}$, and then use the midpoint rule to evaluate $\int_0^1 g_1(A, B)\ \mathrm{d}B$.

A.3. Standardized Gaussian kernel functions. Reich and Shaby (2012) use standardized Gaussian kernel functions as their spatial basis functions in the low-rank max-stable model. Consider a set of $\mathbf{k}_1, \ldots, \mathbf{k}_L$ spatial knot locations in $\mathcal{D}^2$, the region of interest. Then

$$\hat{B}_{il} = \frac{\exp\left\{-\dfrac{||\mathbf{s}_i - \mathbf{k}_l||^2}{2\rho^2}\right\}}{\sum_{j=1}^{L} \exp\left\{-\dfrac{||\mathbf{s}_i - \mathbf{k}_j||^2}{2\rho^2}\right\}} \tag{21}$$

where $||\cdot||$ is the Euclidean distance between a site and a knot location.

APPENDIX B: PRINCIPAL COMPONENTS

As a comparison to the EBFs, Figure 4 gives the first six principal components for the precipitation data. These figures show that the EBFs resemble the EOFs for the precipitation data.

REFERENCES

Coles, S. (2001) *An Introduction to Statistical Modeling of Extreme Values*. Lecture Notes in Control and Information Sciences. London: Springer.

Davison, A. C., Padoan, S. A. and Ribatet, M. (2012) Statistical modeling of spatial extremes. *Statistical Science*, **27**, 161–186.

Everitt, B. and Hothorn, T. (2008) Principal components analysis. In *An Introduction to Applied Multivariate Analysis with R*, 21–54. New York, NY: Springer New York.

Genton, M. G., Ma, Y. and Sang, H. (2011) On the likelihood function of Gaussian max-stable processes. *Biometrika*, **98**, 481–488.

Gneiting, T. and Raftery, A. E. (2007) Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, **102**, 359–378.

de Haan, L. (1984) A Spectral representation for max-stable processes. *The Annals of Probability*, **12**, 1194–1204.

de Haan, L. and Ferreira, A. (2006) *Extreme Value Theory: An Introduction*. Springer Series in Operations Research and Financial Engineering. Springer.

Hannachi, A., Jolliffe, I. T. and Stephenson, D. B. (2007) Empirical orthogonal functions and related techniques in atmospheric science: A review. *International Journal of Climatology*, **27**, 1119–1152.

Huser, R. and Davison, A. C. (2014) Space-time modelling of extreme events. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **76**, 439–461, arXiv:1201.3245.

Lee, D. D. and Seung, S. H. (1999) Learning the parts of objects by non-negative matrix factorizations. *Nature*, **401**, 788 – 791.

Mairal, J., Bach, F. and Ponce, J. (2014) Sparse modeling for image and vision processing. *Foundations and Trends in Computer Graphics and Vision*, **8**, 85 – 283.

Morris, S. A., Reich, B. J., Thibaud, E. and Cooley, D. (under review) A space-time skew-*t* model for threshold exceedances. *Biometrics*.

Padoan, S. A., Ribatet, M. and Sisson, S. A. (2010) Likelihood-based inference for max-stable processes. *Journal of the American Statistical Association*, **105**, 263–277.

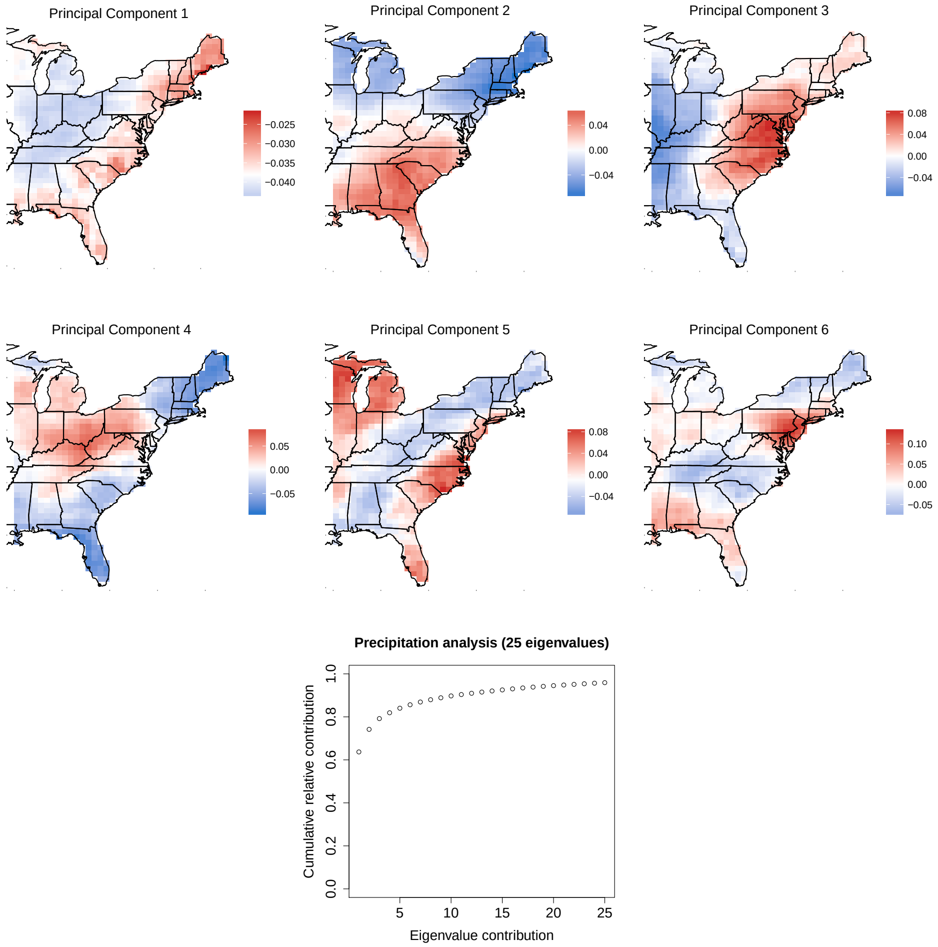

FIG 4. *First six principal components and the cumulative sum of the first 25 eigenvalues for the precipitation data.*

R Core Team (2016) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Reich, B. J. and Shaby, B. A. (2012) A hierarchical max-stable spatial model for extreme precipitation. *The Annals of Applied Statistics*, **6**, 1430–1451.

Ribatet, M. (2015) *SpatialExtremes: Modelling Spatial Extremes*. https://CRAN.R-project.org/package=SpatialExtremes. R package version 2.0-2.

Schlather, M. and Tawn, J. A. (2003) A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika*, **90**, 139–156.

Smith, R. L. (1990) Max-stable processes and spatial extremes. Unpublished manuscript.

Stephenson, A. G. (2009) High-dimensional parametric modelling of multivariate extreme events. *Australian & New Zealand Journal of Statistics*, **51**, 77–88.

Thibaud, E., Mutzner, R. and Davison, A. C. (2013) Threshold modeling of extreme spatial rainfall. *Water Resources Research*, **49**, 4633–4644.

Thibaud, E. and Opitz, T. (2015) Efficient inference and simulation for elliptical Pareto processes. *Biometrika*, **102**, 855–870, arXiv:1401.0168v2.

Wadsworth, J. L. and Tawn, J. A. (2014) Efficient inference for spatial extreme value processes associated to log-Gaussian random functions. *Biometrika*, **101**, 1–15.

Wang, Y. and Stoev, S. A. (2011) Conditional sampling for spectrally discrete max-stable random fields. *Advances in Applied Probability*, **43**, 461–483, arXiv:1005.0312v2.

E-MAIL: samorris@ncsu.edu