AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2016WR018768

Spatial Bayesian hierarchical modeling of precipitation extremes over a large domain

C. Bracken[1,2], B. Rajagopalan[1,3], L. Cheng[3,4], W. Kleiber[5], and S. Gangopadhyay[2]

[1]Department of Civil, Environmental and Architectural Engineering, University of Colorado at Boulder, Boulder, Colorado, USA, [2]Bureau of Reclamation, Technical Service Center, Denver, Colorado, USA, [3]Cooperative Institute for Research in Environmental Sciences, University of Colorado at Boulder, Boulder, Colorado, USA, [4]NOAA Earth System Research Laboratory, Boulder, Colorado, USA, [5]Department of Applied Mathematics, University of Colorado at Boulder, Boulder, Colorado, USA

Key Points:
- A Bayesian spatial model for precipitation extremes is developed for large geographic regions
- Composite likelihood makes computation feasible for thousands of observation locations
- The model produces gridded return levels and associated uncertainty

Correspondence to:
C. Bracken,
cameron.bracken@colorado.edu

Citation:
Bracken, C., B. Rajagopalan, L. Cheng, W. Kleiber, and S. Gangopadhyay (2016), Spatial Bayesian hierarchical modeling of precipitation extremes over a large domain, *Water Resour. Res.*, *52*, 6643–6655, doi:10.1002/2016WR018768.

Received 9 FEB 2016
Accepted 31 JUL 2016
Accepted article online 5 AUG 2016
Published online 27 AUG 2016

Abstract We propose a Bayesian hierarchical model for spatial extremes on a large domain. In the data layer a Gaussian elliptical copula having generalized extreme value (GEV) marginals is applied. Spatial dependence in the GEV parameters is captured with a latent spatial regression with spatially varying coefficients. Using a composite likelihood approach, we are able to efficiently incorporate a large precipitation data set, which includes stations with missing data. The model is demonstrated by application to fall precipitation extremes at approximately 2600 stations covering the western United States, −125°E to −100°E longitude and 30°N–50°N latitude. The hierarchical model provides GEV parameters on a 1/8° grid and, consequently, maps of return levels and associated uncertainty. The model results indicate that return levels and their associated uncertainty have a well-defined spatial structure. Maps of return levels provide information about the spatial variations of the risk of extreme precipitation in the western US and is expected to be useful for infrastructure planning.

1. Introduction

Engineering design of infrastructure, such as flood protection, dams, and management of water supply, and flood control, requires robust estimates of return levels and associated errors of precipitation extremes. Spatial modeling of precipitation extremes can not only capture spatial dependence between stations but also reduce the overall uncertainty in at-site return level estimates by borrowing strength across spatial locations [*Cooley et al*., 2007]. A Bayesian hierarchical model (BHM) of extremes precipitation was first introduced by *Cooley et al*. [2007] and since has been widely discussed in the literature [*Cooley and Sain*, 2010; *Aryal et al*., 2010; *Atyeo and Walshaw*, 2012; *Davison et al*., 2012; *Ghosh and Mallick*, 2011; *Reich and Shaby*, 2012; *Sang and Gelfand*, 2010, 2009; *Apputhurai and Stephenson*, 2013; *Dyrrdal et al*., 2014]. BHMs have also been applied successfully to runoff extremes [*Najafi and Moradkhani*, 2013, 2014]. Recently, BHMs have emerged as a regional frequency analysis (RFA) approach which improves upon many aspects of traditional index-flood-based RFA models [*Hosking and Wallis*, 1993; *Bradley*, 1998; *Wang et al*., 2014; *Yan and Moradkhani*, 2015], such as eliminating the need for delineation of homogeneous regions and providing full uncertainty distributions at ungaged locations [*Renard*, 2011].

While they have seen an increase in popularity in recent years, BHMs for spatial extremes have typically been limited to small geographic regions that include on the order of 100 stations covering areas on the order of 100,000 km^2. Large geographic regions with thousands of stations and diverse climatologies present a computational challenge for BHMs, specifically when computing the likelihood of underlying Gaussian processes (GPs), which for n data points requires solving a linear system of n equations, an $O(n^3)$ operation. Some attempts have been made to model extremes in large regions and with large data sets in a Bayesian hierarchical context. *Reich and Shaby* [2012] use a hierarchical max-stable model with climate model output in the east coast to examine spatially varying GEV parameters, with a max-stable process for the data dependence level. *Ghosh and Mallick* [2011] model gridded precipitation data over the entire US, for annual maxima at a 5° × 5° resolution (43 grid cells) and copula for data dependence, incorporating spatial dependence directly in a spatial model on the data, not parameters. *Cooley and Sain* [2010] and *Sang and Gelfand* [2009] model over 1000 grid cells of climate model output using spatial autoregressive models. The

spatial autoregressive models are parameterized in terms of the precision (inverse covariance) matrix so that no matrix inversion (or Cholesky factorization) is needed. Furthermore, the structure of the generated precision matrix is sparse and utilizing sparse matrix algorithms leads to substantial performance improvements.

When only point data are available, the computational tricks which apply to gridded data cannot be used, though other approximation methods may be employed. One such class of methods for speeding up GP likelihood computations are low-rank approximations [*Banerjee et al*., 2008], where the likelihood is evaluated at only a specific set of knots placed throughout the domain, effectively reducing the size of the covariance matrix. While attractive, low-rank methods can produce large uncertainties between knot locations. Composite likelihood (CL) methods [*Lindsay*, 1988; *Heagerty and Lele*, 1998; *Caragea and Smith*, 2007; *Varin et al*., 2011] approximate the likelihood function itself by breaking stations into groups and evaluating the likelihood for each group. Spectral methods such as *Fuentes* [2007] imagine that irregularly spaced data are on a regular lattice containing missing points and then apply likelihood approximations based on Fourier transforms. Restricted likelihood methods [*Stein et al*., 2004] are similar to CL methods but approximate the full likelihood via conditional distributions. The so-called integrated nested Laplace approximation (INLA) method uses approximations to the posterior marginal distributions to efficiently compute approximate Bayesian posterior samples [*Rue et al*., 2009].

Among all of these approximation methods, the CL method stands out for its flexibility (groupings of stations can be any size or spatial distribution) and ease of implementation inside of arbitrary Bayesian models. Unlike low-rank methods, parameters can be fit for each station, improving the accuracy of spatial predictions. CL methods also have attractive asymptotic properties; as more data are included in each group, the likelihood approximation approaches the true likelihood [*Lindsay*, 1988]. This suggests that if we can strike a balance between group size and computation time, we can obtain accurate parameter estimates and greatly reduced computational burden. In some cases, CL has been shown to diverge greatly from the truth when small group sizes are used [*Ribatet et al*., 2012; *Wang et al*., 2014]. This problem is largely remedied when larger group sizes are used [*Castruccio et al*., 2014].

With the CL method in mind, we turn our attention to the appropriate choice of model structure for large spatial extremes data sets in a Bayesian framework. Bayesian hierarchical spatial extremes models are typically composed of three layers: (1) a data layer consisting of a specification of a joint distribution for the data; (2) a process layer capturing spatial dependencies among the at-site distribution parameters using Gaussian processes; and (3) priors. In the literature, three main methodologies exist for specifying a data layer joint distribution. Conditional independence [*Cooley et al*., 2007] assumes stations are independent given their distribution parameters where all spatial dependence is captured in the process layer. While computationally attractive, conditional independence models preclude the simulation of realistic fields of extremes. Max-stable processes [*Schlather*, 2002; *Cooley et al*., 2006; *Shang et al*., 2011; *Ribatet et al*., 2012; *Padoan et al*., 2010; *Sang*, 2015] refer to the specification of a joint distribution that is specifically formulated for spatial extreme data. While theoretically appropriate, these methods have serious computational limitations for large data sets, which necessitate very small CL group sizes, in turn leading to inaccurate results [*Castruccio et al*., 2014]. Finally, elliptical copulas can be used to specify a joint distribution for spatial data with arbitrary marginal distributions. While they require that some underlying assumptions be met, elliptical copulas are attractive for their flexibility and ease of implementation.

Given a lack of such models in the literature, we propose a Bayesian hierarchical spatial extremes models which can handle thousands of observation locations and arbitrarily large geographic regions. This model pulls together several techniques and approximation approaches to produce gridded return levels and uncertainty distributions at any desired resolution. In the first layer of the hierarchy, an elliptical copula is used represent the joint distribution of the data. In the second layer, a spatial regression is used to model the spatial dependence of the marginal distribution parameters. The spatial regression parameters are allowed to vary in space, allowing the model to adapt to the varied climatologies of large geographic regions. Any remaining spatial dependence is captured using latent GPs. A CL approximation approach is employed to reduce computation time for the elliptical copula and GP likelihoods. In addition, the model is capable of incorporating stations with missing data with little additional computational overhead. The model is applied to observe 3-day fall precipitation extreme in the western US, providing estimated return levels and uncertainty estimates on a 1/8° grid for the entire domain.

In section 2 the general model structure is described. Section 3 describes the composite likelihood procedure. Section 4 describes details of the application to seasonal extreme precipitation in the western US. Results are discussed in section 5 and discussion and conclusions are given in section 6.

2. Model Structure

The joint distribution of the m stations in each year is modeled as a realization from a Gaussian elliptical copula with generalized extreme value (GEV) distribution marginals. The copula is characterized by pairwise dependence matrix Σ. Spatial dependence is further captured through spatial processes on the location $\mu(\mathbf{s})$, scale $\sigma(\mathbf{s})$, and $\xi(\mathbf{s})$ parameters. We assume the parameters can be described through a latent spatial regression where the residual component $w_\gamma(\mathbf{s})$ follows a mean 0, stationary, isotropic Gaussian process (GP) with covariance function $C_\gamma(\mathbf{s}, \mathbf{s}')$ where γ represents any GEV parameter (μ, σ, ξ). The corresponding covariance matrix is $C_\gamma(\boldsymbol{\theta}_\gamma)=[C_\gamma(\mathbf{s}_i, \mathbf{s}_j; \boldsymbol{\theta}_\gamma)]_{i,j=1}^m$ where $\boldsymbol{\theta}_\gamma$ represents the covariance parameters. The first layer of the hierarchical model structure is

$$(Y(\mathbf{s}_1, t), \ldots, Y(\mathbf{s}_m, t)) \sim C_m[\Sigma; \{\mu(\mathbf{s}), \sigma(\mathbf{s}), \xi(\mathbf{s})\}], \tag{1}$$

$$Y(\mathbf{s}, t) \sim \text{GEV}[\mu(\mathbf{s}), \sigma(\mathbf{s}), \xi(\mathbf{s})], \tag{2}$$

where $Y(\mathbf{s}, t)$ is the response at site $\mathbf{s}$ and time t and C_m stands for "m-dimensional Gaussian elliptical copula" with dependence matrix Σ. The spatial data layer processes in each year are assumed independent and identically distributed. Marginally, observations have a generalized extreme value (GEV) distribution, the theoretical distribution of block maxima data.

The second layer of the hierarchy, also known as the process layer, involves spatial models for the GEV parameters

$$\mu(\mathbf{s})=\beta_{\mu 0}+\mathbf{x}_\mu^T(\mathbf{s})\boldsymbol{\beta}_\mu(\mathbf{s})+w_\mu(\mathbf{s}), \tag{3}$$

$$\sigma(\mathbf{s})=\beta_{\sigma 0}+\mathbf{x}_\sigma^T(\mathbf{s})\boldsymbol{\beta}_\sigma(\mathbf{s})+w_\sigma(\mathbf{s}), \tag{4}$$

$$\xi(\mathbf{s})=\beta_{\xi 0}+\mathbf{x}_\xi^T(\mathbf{s})\boldsymbol{\beta}_\xi(\mathbf{s})+w_\xi(\mathbf{s}), \tag{5}$$

where $\beta_{\gamma 0}$ are spatially constant intercept terms, $\mathbf{x}_\gamma^T(\mathbf{s}_i)$ is a vector of p spatially varying predictors, and $\boldsymbol{\beta}_\gamma(\mathbf{s})=[\beta_{\gamma 1}(\mathbf{s}), \ldots, \beta_{\gamma p}(\mathbf{s})]^T$ is a vector of p spatially varying regression coefficients. Covariates will be discussed in section 4.2.

The shape parameter ξ is notoriously difficult to estimate its value determining the support of the GEV distribution and the heaviness of the tail. Positive values of ξ indicate a lower bound to the distribution and a heavy upper tail which are typically seen with precipitation data. Negative values of indicate an upper bound with a tail that may be heavy or light. A zero value of ξ indicates an unbounded distribution which is light tailed. In many studies, ξ is modeled as a single value per study area or per region within the study area [*Cooley et al.*, 2007; *Renard*, 2011; *Atyeo and Walshaw*, 2012; *Apputhurai and Stephenson*, 2013]. As in *Cooley and Sain* [2010], we cannot assume that this parameter is constant over the large study region and so it is modeled spatially along with the other GEV parameters.

For large regions we cannot assume that a constant spatial regression holds for the entire domain and thus must introduce spatial variation in the regression coefficients. The second layer of the hierarchy also involves a spatial model for these regression coefficients

$$\beta_{\gamma i}(\mathbf{s})=\sum_{j=1}^{k} c_{\gamma ij}\eta_{\gamma ij}(\mathbf{s}; a_{\gamma ij}) \quad i=1, \ldots, p, \tag{6}$$

where the γ can represent any GEV parameter, $\beta_{\gamma i}(\mathbf{s})$ is the ith (spatially varying) regression coefficient, $c_{\gamma ij}$'s are weights for k radial basis functions, the $\eta_{\gamma ij}$'s, which are distributed throughout the domain. Additional details are provided in section 2.2. The hierarchical relationship between parameters is summarized in Figure 1.

2.1. Elliptical Copula for Data Dependence

Elliptical copulas are a flexible tool for modeling multivariate data [*Renard*, 2011; *Sang and Gelfand*, 2010; *Ghosh and Mallick*, 2011; *Renard and Lang*, 2007]. This class of copulas can represent spatial data with any

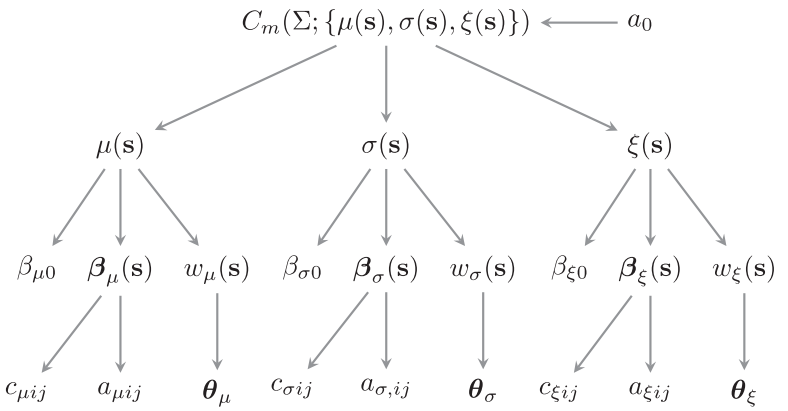

Figure 1. Hierarchical relationship between model parameters. Note that while four layers are shown for illustrative purposes, the rows in the diagram are not layers in the hierarchical model since only components which contribute to the likelihood are considered layers.

marginal distribution, a particularly attractive feature for extremal data. The Gaussian copula constructs the joint pdf of a random vector $(Y_1, \ldots Y_m)$ as

$$F_{Cop}(y_1, \ldots, y_m) = \Phi_{\Sigma}(u_1, \ldots .u_m), \tag{7}$$

where $\Phi_{\Sigma}(u_1, \ldots .u_m)$ is the joint cdf of an m-dimensional multivariate normal distribution with dependence matrix Σ, $u_i = \phi^{-1}(F_i[y_i])$, ϕ is the cdf of the standard normal distribution, and F_i is the marginal GEV cdf at site i. The corresponding joint pdf is

$$f_{Cop}(y_1, \ldots, y_m) = \frac{\prod_{i=1}^{m} f_i[y_i]}{\prod_{i=1}^{m} \psi[u_i]} \Psi_{\Sigma}(u_1, \ldots .u_m), \tag{8}$$

where f_i is the marginal GEV pdf at site i, ψ is the standard normal pdf, and Φ_{Σ} is the joint pdf of an m-dimensional multivariate normal distribution.

The dependence between sites is assumed to be a function of distance [*Renard*, 2011]. The dependence matrix is constructed with a simple exponential model

$$\Sigma(i,j) = \exp\left(-||\mathbf{s}_i - \mathbf{s}_j||/a_0\right), \tag{9}$$

where a_0 is the copula range parameter. Note that the values in this dependence matrix are not covariances since they are not scaled by a marginal variance parameter, though the dependence matrix is a valid covariance matrix. By analogy with the variogram, the dependence model is termed the dependogram [*Renard*, 2011].

2.2. Spatial Regression Model

For large regions, spatial regression relationships may not hold constant for the entire domain due to varying regional climatologies and topographies. In these situations a single regression coefficient for each covariate is not appropriate and it is necessary to allow for spatial variation in the spatial regressions for each GEV parameter. Each regression coefficient is represented as a weighted sum of radial basis functions basis functions (equation (6)). The form of these radial basis functions is

$$\eta_{\gamma ij}(\mathbf{s}; a_{\gamma ij}) = \exp\left(-||\mathbf{s} - \mathbf{s}_i||^2 / a_{\gamma ij}^2\right), \tag{10}$$

where $a_{\gamma ij}^2$ is a range parameter determining the spatial extent of the basis function. These radial basis functions, also known as Gaussian kernels, are placed at points throughout the domain, known as knots. The sum of the radial basis functions creates a smoothly varying surface for each regression coefficient.

The knots are placed according to a space-filling design [*Johnson et al.*, 1990; *Nychka and Saltzman*, 1998]. For each GEV parameter, we use 10 knots (Figure 2) since based on the author's experience, regression

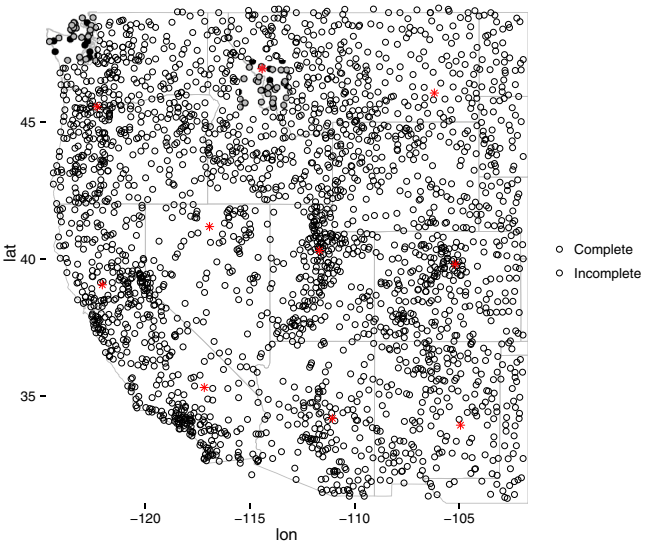

Figure 2. Station locations with complete data (black solid dots) and station locations with incomplete data (gray filled dots). Red asterisks are knot locations for the spatially varying regression coefficients.

relationships in the western US region tend to hold for regions of a few square degrees. For simplicity, the same knot locations were used for each GEV parameter though each could use different configurations of knots.

2.3. Missing Data

Stations with missing data can be easily incorporated in the model. When the GEV likelihood is computed, years with missing data are simply skipped. With at least 30 years of data at each station, the GEV parameters can be estimated adequately based on only the available data. For simplicity, the copula likelihood was only computed for stations with complete data. This did not strongly affect the estimates of the copula range parameter during testing. An alternative approach to incorporating missing data when computing the likelihood for each year of data is to use a covariance matrix of a size corresponding to the number of available stations. While potentially more accurate, managing multiple matrices of varying sizes leads to additional computational costs and coding challenges.

2.4. Likelihood and Priors

The marginal distribution of $Y(\mathbf{s}_i, t)$ is $\text{GEV}(y(\mathbf{s}_i, t)|\mu(\mathbf{s}_i), \sigma(\mathbf{s}_i), \xi(\mathbf{s}_i))$ where the log-likelihood for some data point y is

$$\log \text{GEV}(y|\mu, \sigma, \xi) = -\log(\sigma) - (1+1/\xi)\log(b) - b^{-1/\xi}, \tag{11}$$

where $b = 1+\xi(y-\mu)/\sigma$.

Let γ represent any of the GEV parameters (μ, σ, ξ). The residual Gaussian processes likelihood $p(\mathbf{w}_\gamma|\boldsymbol{\theta}_\gamma)$ is obtained from the multivariate normal density function $\mathbf{w}_\gamma|\boldsymbol{\theta}_\gamma \sim \text{MVN}(\mathbf{0}, \Sigma_\gamma)$, where $\Sigma_\gamma = C(\boldsymbol{\theta}_\gamma)$. We use an exponential covariance function with parameters δ_γ^2 (the partial sill or marginal variance), a_γ (the range), and τ_γ^2 (the nugget), so $\boldsymbol{\theta}_\gamma = (\delta_\gamma^2, a_\gamma, \tau_\gamma^2)$. The parametric form of the covariance function is

$$C(\mathbf{s}_i, \mathbf{s}_j; \boldsymbol{\theta}_\gamma) = \begin{cases} \delta_\gamma^2 \exp(-||\mathbf{s}_i - \mathbf{s}_j||/a_\gamma) & i \neq j \\ \delta_\gamma^2 + \tau_\gamma^2 & i = j \end{cases}.$$

We use weakly informative normal priors centered at 0, with a standard deviations as follows: 0.1 $(\delta_\xi^2, \tau_\xi^2)$, 1 $(\delta_\mu^2, \delta_\sigma^2, \tau_\mu^2, \tau_\sigma^2, \beta_{\xi 0}, c_{\mu ij}, c_{\sigma ij}, c_{\mu ij}; i = 1, \ldots, m)$, 10 $(\beta_{\mu 0}, \beta_{\sigma 0})$, 1000 $(a_\mu, a_\sigma, a_\xi, a_0,)$, 5000 $(a_{\mu ij}, a_{\sigma ij}, a_{\xi ij};\ i = 1, \ldots, p,\ j = 1, \ldots, k)$. For ξ we restrict values to the range $[-0.5, 0.5]$, motivated by the typical ranges seen in precipitation data [*Cooley and Sain*, 2010].

3. Estimation

3.1. Composite Likelihood

Composite likelihood for spatial data are a method in which the full likelihood is approximated by a set of conditional or marginal likelihoods (see *Varin et al.* [2011] for a recent review). Conditional approaches construct the composite likelihood as a product of conditional likelihoods for each observation given neighboring observations [*Vecchia*, 1988; *Stein et al.*, 2004]. Marginal approaches construct the conditional likelihood as a product of joint densities of groups of observations of two or more. The case when a group consists of one observation is known as the independence likelihood, which precludes the computation of spatial

dependence parameters [*Varin et al*., 2011]. Composite likelihood methods have also been applied to max-stable spatial processes (see *Sang* [2015] for a recent review).

In our approach, the stations are broken up into G groups each with n_g stations. The marginal composite likelihood estimator (L_c) is constructed as a product of the group likelihoods

$$L_c(\theta|\mathbf{y}_1,\ldots,\mathbf{y}_G)=\prod_{g=1}^{G} L_g(\theta|\mathbf{y}_g), \tag{12}$$

where θ contains covariance parameters and $\mathbf{y}_g$ contains observations from group g. This approach is similar to the "small blocks" approach from *Caragea and Smith* [2006, 2007]. Approximating the likelihood in this way requires $O(Gn_g^3)$ computations as opposed to $O(n^3)$. An assumption in this approach is that each group is independent, which is expected to introduce some loss of statistical efficiency. As n_g increases (and G decreases) the composite likelihood estimator approaches the true likelihood at the cost of increased computation time [*Caragea and Smith*, 2007]. The choice of n_g must be a balance between computation time and accuracy. Along these lines, *Caragea and Smith* [2007] suggest that computational efficiency is maximized when n_g is between $m^{1/2}$ and $m^{2/3}$, where m is the total number of stations. In this application, the composite likelihood approximation is applied to compute the copula likelihood as well as each of the latent GEV parameter residuals.

3.2. Composite Likelihood Group Size and Distribution

In order to use a composite likelihood approach, we must decide how many stations to use in each group (n_g). The number of stations in each group should be small enough so as not to incur substantial computational cost but large enough so that the covariance parameters can be adequately estimated. We used 30 stations per group or approximately 1% of the total number of stations. The consequences of this choice are explored in section 5.2.

We must also choose how stations are to be grouped. Several approaches come to mind such as selecting groups based on climatological regions, elevation bands, or a course grid. We chose to group stations randomly, expecting that groups will have a mixture of stations with a range of spatial proximities, allowing for estimation of both small and large scale behavior.

What remains in the model are a few application specific details: selection of the knot locations and the selection of covariates. These are described in the next sections.

4. Application to the Western US

4.1. Precipitation Data

Daily fall (SON) precipitation data were obtained from the Global Historical Climatology Network (GHCN). We use all available stations in the western US which contain more than 30 years of data from 1950 to 2013. Maximum 3 day total precipitation was computed for each year during the fall period (SON). For a data point to be considered nonmissing, we required no more than 25% of the days to be missing in the corresponding 3 month period. The number of stations included (with the number of complete stations in parentheses) was 2618 (848). Figure 2 shows the station locations, with solid black points indicating stations with complete data and filled gray points indicating stations with incomplete data. Red asterisks indicate the centers (knots) for the radial basis functions.

4.2. Covariates

For all GEV parameters the same covariates are used, i.e., $\mathbf{x}_\mu(\mathbf{s})=\mathbf{x}_\sigma(\mathbf{s})=\mathbf{x}_\xi(\mathbf{s})=\mathbf{x}(\mathbf{s})$. The covariates are elevation and mean seasonal precipitation. Typically, latitude and longitude are used as well, but the spatially variation of the regression coefficients precludes this. Covariates were obtained at knot locations, station locations, and at a 1/8° grid throughout the study area. Elevation data were obtained from the NASA Land Data Assimilation Systems (NLDAS) (http://ldas.gsfc.nasa.gov/nldas/NLDASelevation.php) [*Xia et al*., 2012a, 2012b]. Mean seasonal precipitation was computed from the Maurer data set [*Maurer et al*., 2002].

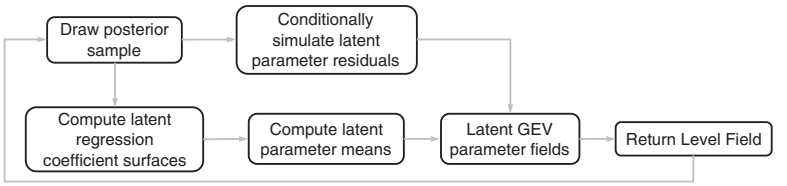

Figure 3. Schematic of the procedure for generating gridded return levels from posterior samples.

4.3. Implementation and Model Fitting

The model was implemented in the Stan modeling language [*Stan Development Team*, 2015a] using the RStan interface [*Stan Development Team*, 2015b]. Stan uses the No-U-Turn Sampler (NUTS), an implementation of Hamiltonian Monte Carlo (HMC) [*Hoffman and Gelman*, 2014]. The Stan/Rstan interface to the NUTS algorithm allows for Bayesian. The NUTS sampler uses information about the gradient of the posterior to avoid Metropolis-Hastings (MH) random walk behavior. NUTS is a multivariate sampler that deals well with high-dimensional problems, highly correlated parameters, tends to need very few warmup iterations, and typically produces nearly uncorrelated samples. For these reasons, very long chains are usually not needed, nor is thinning. The tradeoff we observed in this application was that the NUTS sampler required longer computation time per sample compared to a traditional MH sampler.

Three chains of length 3000 were run, with the first 1000 iterations discarded as warmup, resulting in 6000 samples for each parameter. Chains took 2–3 days to run on an 8-core 2.6 GHz Intel Core i7 processor. To assess convergence, we computed the $\hat{R}$ statistic to ensure it is below 1.1, as well as visually inspected trace plots.

4.4. Computation of Gridded Return Levels

After model fitting is complete, distributions of each GEV parameter are obtained at each 1/8° grid cell via conditional simulation. The gridded parameter values are used to compute return levels at each grid cell using the GEV return level formula

$$z_i(r) = \mu_i + \sigma_i((-\log(1-1/r))^{-\xi_i} - 1)/\xi_i, \quad i = 1, .., m,$$

where r is the return period in years (100 years for example). The detailed steps for this procedure are shown in Figure 3 and are described as follows:

1. Select a single posterior sample of all model parameters.
2. Conditionally simulate latent GEV parameter residuals, $w_\mu(\mathbf{s})$, $w_\sigma(\mathbf{s})$, and $w_\xi(\mathbf{s})$.
3. Compute latent regression coefficient surfaces, $\boldsymbol{\beta}_\mu(\mathbf{s})$, $\boldsymbol{\beta}_\sigma(\mathbf{s})$, and $\boldsymbol{\beta}_\xi(\mathbf{s})$ by combining the radial basis functions at knot locations.
4. Compute GEV parameter mean surfaces, $\beta_{\mu 0} + \mathbf{x}_\mu^T(\mathbf{s})\boldsymbol{\beta}_\mu(\mathbf{s})$, $\beta_{\sigma 0} + \mathbf{x}_\sigma^T(\mathbf{s})\boldsymbol{\beta}_\sigma(\mathbf{s})$, and $\beta_{\xi 0} + \mathbf{x}_\xi^T(\mathbf{s})\boldsymbol{\beta}_\xi(\mathbf{s})$.
5. Combine the results of steps 2 and 4 to create GEV parameter surfaces, $\mu(\mathbf{s})$, $\sigma(\mathbf{s})$, $\xi(\mathbf{s})$.
6. Combine the GEV parameter surfaces at each grid cell using equation (4) to create return level surfaces.
7. Repeat steps 1–6 for each posterior sample.

5. Results

5.1. Testing the Validity of the Gaussian Copula

An implication of the Gaussian copula is that marginal distributions are asymptotically independent, or $P(F_x(X) > p | F_y(Y) > p) \to 0$ as $p \to 1$ [*Renard and Lang*, 2007]. To test this we conducted asymptotic independence tests [*Reiss and Thomas*, 2007] for all pairs of stations. The null hypothesis of this test is dependence, so setting a significance level of 99% ensures that stations passing the test exhibit strong asymptotic dependence. At the 99% significance level, 0.30% of pairwise stations exhibited dependence, less than 1% expected from chance (Figure 4). In addition, we examined plots of the station locations when dependence was indicted by the test. These plots did not show any discernible spatial pattern of dependence; for example, dependent stations did not tend to fall near each other.

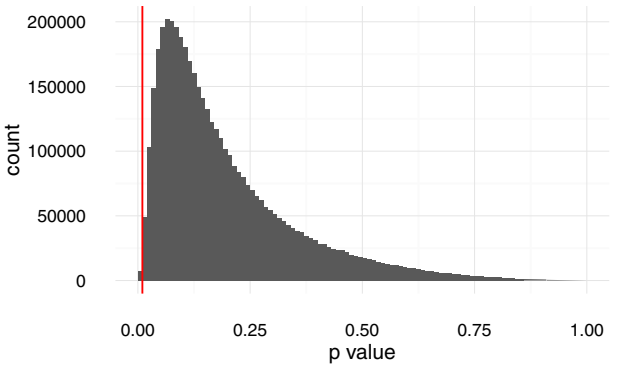

Figure 4. Histogram of *P* values from the asymptotic independence test between all pairwise stations. Values below the red line indicate asymptotic dependence at the 99% level, which is the case for only 0.3% of pairwise stations, less than 1% expected from chance.

5.2. Group Size Selection

To demonstrate that the selection of group size has little effect on return levels, a small experiment is conducted. We run the model for a region encompassing most of the state of Oregon, using 4 knots. The group size is set to be 2, 5, 10, 15, 20, and 30 stations representing approximately 1, 2, 4, 6, 8, and 13% of the total number of stations, respectively. The same 240 stations (60 complete, 180 incomplete) are used in each model run.

Figure 5 shows the median return level for each model run. The results are nearly identical for this range of group sizes, indicating that median return levels are not sensitive to the choice of group size. Credible intervals of return levels (not shown) were quite similar as well, with credible intervals decreasing as group size is increased indicating that a larger group size yields more accurate results, as expected. In light of this we chose a group size of 30 for the large domain which provides both a diversity in the distribution of stations within a group but is small enough to not significantly hider computation.

5.3. Gridded Return Levels

Figure 6 shows the median of the GEV parameters after interpolation by conditional simulation as well as the average extreme precipitation for the fall period. The location and scale fields are highly correlated; locations with higher average extreme precipitation tend to have more variability in these extremes. As expected, the location parameter field corresponds quite well with the location parameter field. *Cooley and Sain* [2010] who model extreme precipitation from a climate model over the same study region, though for winter.

Values of ξ are always positive, indicating a heavy upper tail for precipitation throughout the western US. The southwestern coastal region exhibits the highest shape parameters reflecting the arid climate of the

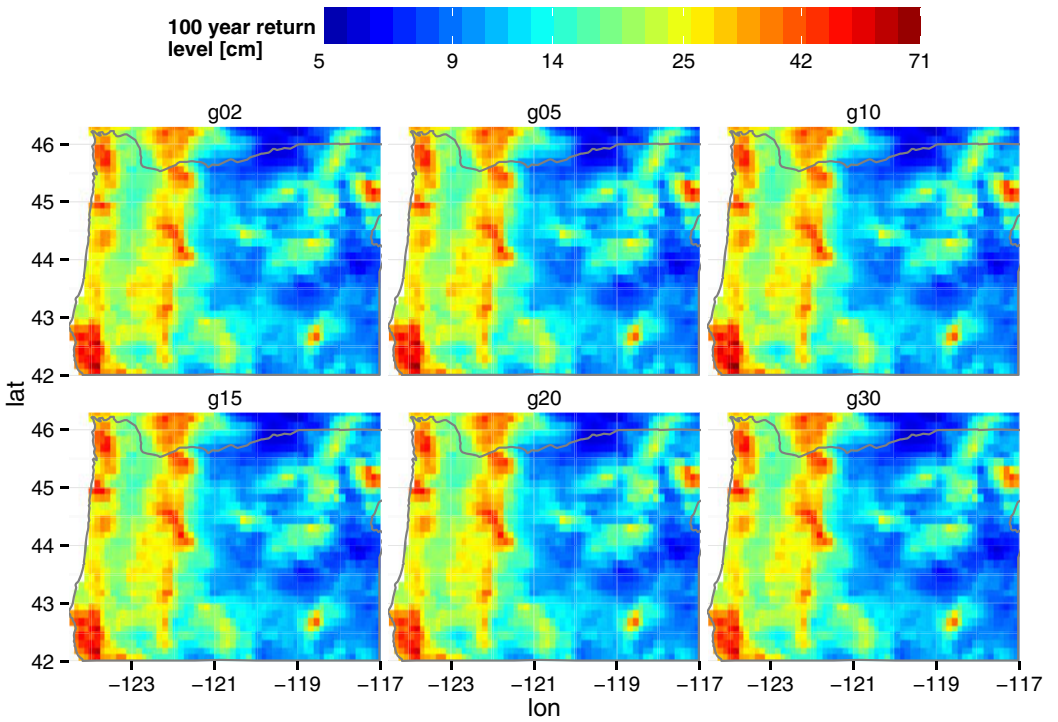

Figure 5. Median return levels using a group sizes of 2, 5, 10, 15, 20, and 30. Note the logarithmic color scale.

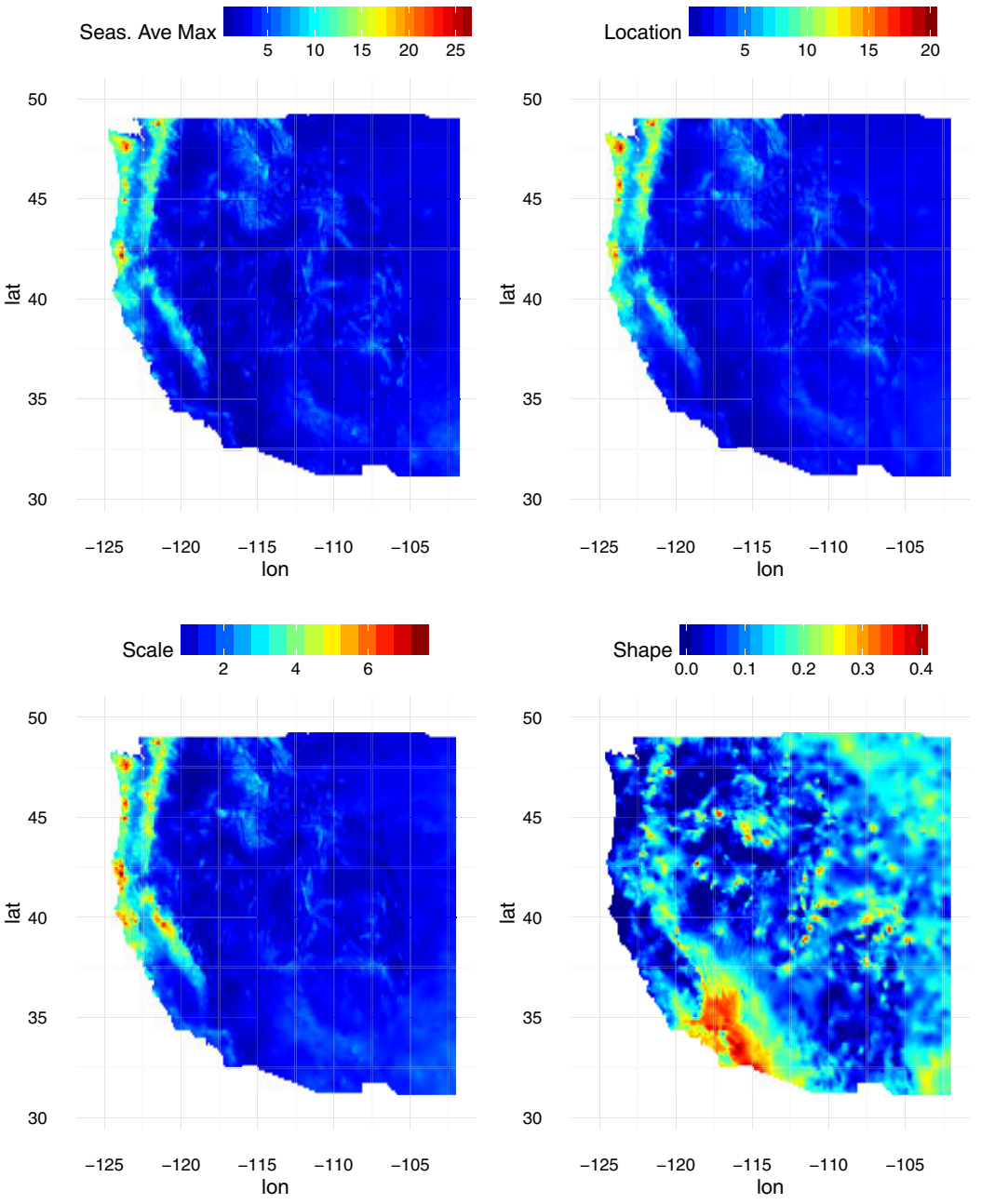

Figure 6. Median of underlying GEV parameters, location (μ, upper right), scale (σ, lower left), and shape (ξ, lower right). The figure on the top left shows the average fall 3 day maximum precipitation computed from the *Maurer et al.* [2002] data set.

region. In some years the seasonal maxima in that region are identically zero, while other season has large events, causing the heaviest tailed distributions in the western US.

We find reasonable agreement with *Cooley and Sain* [2010] (Figures 1 and 3) in terms of spatial patterns of GEV parameters and return levels, though magnitudes are somewhat different, which is expected from winter versus fall. Similar patterns of high location and scale parameters can be seen in the pacific northwest and high shape values in the southwest can be seen in both models.

5.4. Validation

Cross validation was conducted by dropping 885 stations or approximately 35% of the total stations. Gridded return levels were computed for this subset of data. Figure 7 shows the difference between the

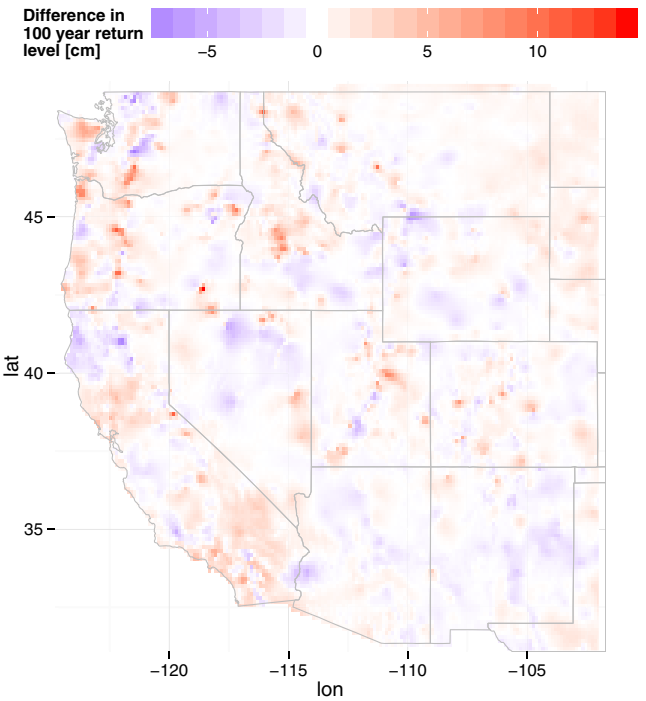

Figure 7. Difference between fiftieth percentile return levels from the full model and the validation model dropping 35% of the data.

median return level for the full data and subset data. The difference map shows some spatial coherence but none that indicates any strong bias in a single region (states for example). The largest differences occur in areas in the northwest where influential stations were dropped randomly. For example in the Pacific Northwest, some of the largest differences are seen where stations with extremes higher than surrounding stations were dropped. These differences do not suggest any systematic bias in the method. The method is expected to perform well even with smaller subsets of data.

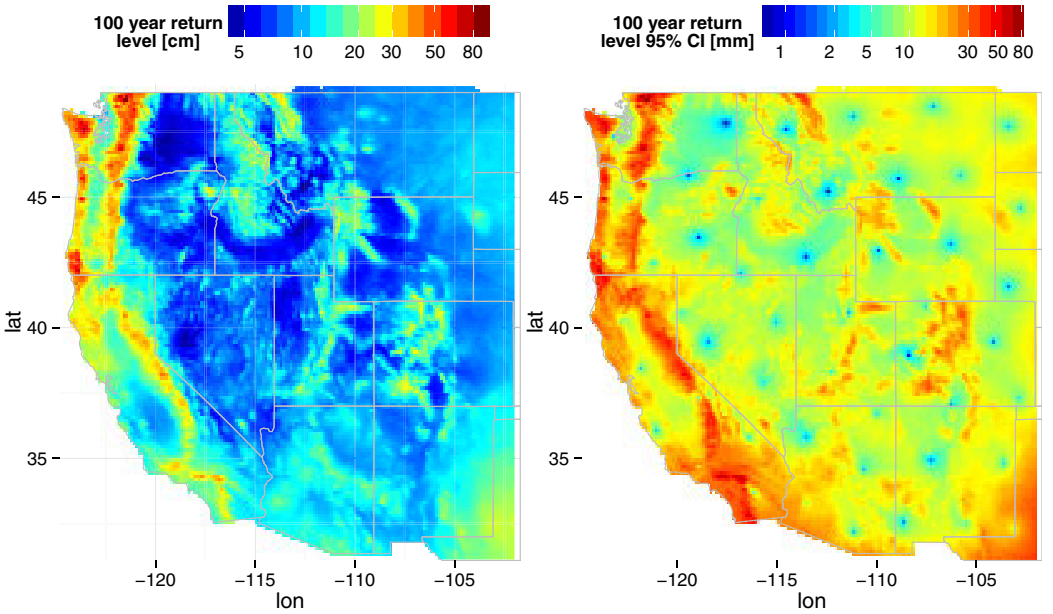

Figure 8. Return levels maps produced using latent Gaussian predictive processes.

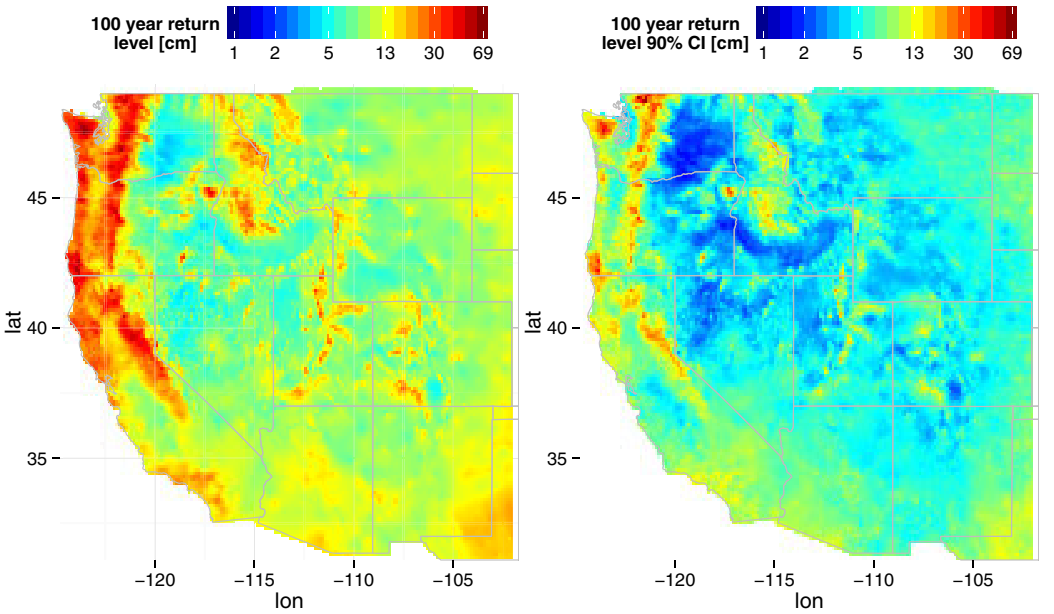

Figure 9. Median 100 year return levels for fall (left) and width of corresponding 95% credible interval (right). Note the logarithmic color scale.

5.5. A Case for Composite Likelihood

To highlight the usefulness of the composite likelihood approach for this application, we present results using a Gaussian predictive process (GPP) model [*Banerjee et al.*, 2008] for the latent GEV parameter processes (Figure 8). A Gaussian predictive process model approximates the likelihood at a small set of knots to reduce the dimension of the covariance matrix and the computational burden of inverting it. The GP values at knot locations are treated as model parameters, and sampled in the MCMC procedure. During the likelihood computation, the GP is interpolated via kriging to each station location in order to obtain the value of the GP at each station. We originally set out using latent GPPs for this application for each GEV parameter but switched to a composite likelihood approach when we realized the uncertainty was unacceptably large away from knot locations. While this section is not intended to be an exhaustive comparison of these two methods, it is intended as a cautionary note.

The median return levels with the GPP approach (Figure 8, left) were nearly identical to those from the composite likelihood method (Figure 9, left) but large differences are apparent when looking at the credible intervals of the return levels. Clear artifacts are present in Figure 8 (right) where the locations of knots are apparent. Uncertainty away from knot locations was typically large, rendering this method much less useful than the composite likelihood approach. The reason for the large uncertainty away from knot locations is due to the conditional simulation procedure to generate GEV parameter fields. The GP is simulated from as if there are no stations in between knots, when in reality there are many stations in between each knot that were unused in the spatial predictions. Thus, we recommend the use of a composite likelihood type approach when low predictive uncertainty is required.

6. Discussion and Conclusions

We describe a general Bayesian hierarchical model for extreme data observed over space and time. The data are assumed to originate from a Gaussian elliptical copula having generalized extreme value (GEV) marginal distributions. Spatial dependence is further captured by Gaussian processes on the three GEV parameters (location, scale, and shape). Using a composite likelihood approach, we are able to incorporate 2595 observation locations with 54 years of data. With spatially varying regression coefficients, the model can be applied to arbitrarily large regions. The model was applied to extreme 3 day precipitation in fall in the western United States, a climatologically and geographically diverse region. The model was fit using a

standard Bayesian methodology, implicitly capturing uncertainty in the parameter estimates and spatial predictions.

This model is useful as a tool for creating precipitation frequency maps similar to the NOAA Atlas 14 [*National Oceanic Atmospheric Administration*, 2004]. Engineering applications for these types of maps include design of water management infrastructure (dams, levees, storm water control structures), water supply management, and flood control. While we focused on 100 year return levels, the model can easily create maps for any return period, without reestimation of parameters. One estimation run may take 2–3 days but this is a one-time up-font cost.

In section 5.5 we briefly examine results for the same region using a Gaussian predictive process (GPP) model for the latent GEV parameters. In this application, GPPs produced unreasonably large posterior credible intervals when moving away from knot locations. In light of this we recommend a composite likelihood approach for regions of equal or larger size than the western US.

A crux of this model is the use of appropriate spatial covariates. Mean seasonal precipitation (MSP) had a correlation of 95% with the MLE estimates of μ and 75% with the MLE estimates of σ. Even with spatially varying regression coefficients, appropriate covariates are key. The covariates here helped in generating realistic spatial variability and helped to reveal a complex spatial pattern for the shape parameter, ξ. The strongest covariate for ξ was elevation. The spatial variability in ξ shows that it is inappropriate to model without spatial variation for anything but the smallest regions.

A number of extensions can be made to this framework. The most obvious extension is to allow temporal variation in the GEV parameters by including temporal covariates. While this extension remains infeasible for the size of the current study region, it may be feasible for smaller regions, say a single state or moderate sized river basin. Additional spatial covariates could be included; for example, seasonal temperature, winds or evapotranspiration. A model such as the one presented here can be used to investigate changes in risk under specific climate regimes (i.e., ENSO); one would simply include the mean seasonal precipitation field from strong El Niño or La Niña years. Because we incorporate a data layer, this model could be used to simulate realistic fields of extremes under specific climate regimes. Finally, we plan to explore the linking of streamflow data into the hierarchy, so that streamflow extremes can be simultaneously estimated.

Acknowledgments

Funding for this research by a Science and Technology grant from Bureau of Reclamation is gratefully acknowledged. Kleiber's portion was supported by NSF DMS-1406536. This work utilized the Janus supercomputer, which is supported by the National Science Foundation (award CNS-0821794) and the University of Colorado Boulder. The Janus supercomputer is a joint effort of the University of Colorado Boulder, the University of Colorado Denver, and the National Center for Atmospheric Research. The authors are thankful for support from the Janus supercomputer staff at the University of Colorado. Pre- and postprocesseing analysis was conducted using the R language [*R Core Team*, 2014]. Data are available at: http://bechtel.colorado.edu/bracken/spatial_extremes/data/.

References

Apputhurai, P., and A. G. Stephenson (2013), Spatiotemporal hierarchical modelling of extreme precipitation in Western Australia using anisotropic Gaussian random fields, *Environ. Ecol. Stat.*, *20*(4), 667–677.

Aryal, S. K., B. C. Bates, E. P. Campbell, Y. Li, M. J. Palmer, and N. R. Viney (2010), Characterizing and modeling temporal and spatial trends in rainfall extremes, *J. Hydrometrol.*, *10*(1), 241–253, doi:10.1175/2008JHM1007.1.

Atyeo, J., and D. Walshaw (2012), A region-based hierarchical model for extreme rainfall over the UK, incorporating spatial dependence and temporal trend, *Environmetrics*, *23*(6), 509–521.

Banerjee, S., A. E. Gelfand, A. O. Finley, and H. Sang (2008), Gaussian predictive process models for large spatial data sets, *J. R. Stat. Soc. B*, *70*(4), 825–848, doi:10.1111/j.1467-9868.2008.00663.x.

Bradley, A. A. (1998), Regional frequency analysis methods for evaluating changes in hydrologic extremes, *Water Resour. Res.*, *34*(4), 741–750.

Caragea, P. C., and R. L. Smith (2006), Approximate likelihoods for spatial processes, in *proceedings of the Joint Statistical Meeting*, pp. 385–390.

Caragea, P. C., and R. L. Smith (2007), Asymptotic properties of computationally efficient alternative estimators for a class of multivariate normal models, *J. Multivar. Anal.*, *98*(7), 1417–1440.

Castruccio, S., R. Huser, and M. Genton (2014), High-order composite likelihood inference for max-stable distributions and processes, *J. Comput. Graph. Stat.*, 1–32, doi:10.1080/10618600.2015.1086656.

Cooley, D., and S. R. Sain (2010), Spatial hierarchical modeling of precipitation extremes from a regional climate model, *J. Agric. Biol. Environ. Stat.*, *15*(3), 381–402.

Cooley, D., P. Naveau, and P. Poncet (2006), Variograms for spatial max-stable random fields, in *Dependence in Probability and Statistics, Lect. Notes Stat.*, vol. 187, pp. 373–390, Springer-Verlag, N. Y.

Cooley, D., D. Nychka, and P. Naveau (2007), Bayesian spatial modeling of extreme precipitation return levels, *J. Am. Stat. Assoc.*, *102*, 824–840.

Davison, A. C., S. A. Padoan, and M. Ribatet (2012), Statistical modeling of spatial extremes, *Stat. Sci.*, *27*(2), 161–186.

Dyrrdal, A. V., A. Lenkoski, T. L. Thorarinsdottir, and F. Stordal (2014), Bayesian hierarchical modeling of extreme hourly precipitation in Norway, *Environmetrics*, *26*(2), 1–18, doi:10.1002/env.2301.

Fuentes, M. (2007), Approximate likelihood for large irregularly spaced spatial data, *J. Am. Stat. Assoc.*, *102*(477), 321–331.

Ghosh, S., and B. K. Mallick (2011), A hierarchical Bayesian spatio-temporal model for extreme precipitation events, *Environmetrics*, *22*(2), 192–204.

Heagerty, P. J., and S. R. Lele (1998), A composite likelihood approach to binary spatial data, *J. Am. Stat. Assoc.*, *93*(443), 1099.

Hoffman, M. D., and A. Gelman (2014), The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo, *J. Mach. Learning Res.*, *15*, 1593–1623.

Hosking, J. R. M., and J. R. Wallis (1993), Some statistics useful in regional frequency analysis, *Water Resour. Res.*, *29*(2), 271–281.
Johnson, M. E., L. M. Moore, and D. Ylvisaker (1990), Minimax and maximin distance designs, *J. Stat. Plann. Inference*, *26*(2), 131–148.
Lindsay, B. G. (1988), Composite likelihood methods, *Contemp. Math.*, *80*, 221–239.
Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen (2002), A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States, *J. Clim.*, *15*(22), 3237–3251.
Najafi, M. R., and H. Moradkhani (2013), Analysis of runoff extremes using spatial hierarchical Bayesian modeling, *Water Resour. Res.*, *49*, 6656–6670, doi:10.1002/wrcr.20381.
Najafi, M. R., and H. Moradkhani (2014), A hierarchical Bayesian approach for the analysis of climate change impact on runoff extremes, *Hydrol. Processes*, *28*(26), 6292–6308.
National Oceanic Atmospheric Administration (2004), Precipitation-frequency atlas of the United States, technical report, NOAA Atlas 14, vol. 1, version 5.0, edited by G. M. Bonnin et al., NOAA, National Weather Service, Silver Spring, Md.
Nychka, D., and N. Saltzman (1998), Design of air-quality monitoring networks, in *Case Studies in Environmental Statistics*, pp. 51–76, Springer, N. Y.
Padoan, S. A., M. Ribatet, and S. A. Sisson (2010), Likelihood-based inference for max-stable processes, *J. Am. Stat. Assoc.*, *105*(489), 263–277, doi:10.1198/jasa.2009.tm08577.
R Core Team (2014), *R: A Language and Environment for Statistical Computing*, R Found. for Stat. Comput., Vienna.
Reich, B. J., and B. Shaby (2012), A hierarchical max-stable spatial model for extreme precipitation, *Ann. Appl. Stat.*, *6*(4), 1430–1451.
Reiss, R.-D., and M. Thomas (2007), *Statistical Analysis of Extreme Values: With Applications to Insurance, Finance, Hydrology and Other Fields*, 3rd ed., Basel Birkhäuser, Berlin.
Renard, B. (2011), A Bayesian hierarchical approach to regional frequency analysis, *Water Resour. Res.*, *47*, W11513, doi:10.1029/2010WR010089.
Renard, B., and M. Lang (2007), Use of a Gaussian copula for multivariate extreme value analysis: Some case studies in hydrology, *Adv. Water Resour.*, *30*(4), 897–912.
Ribatet, M., D. Cooley, and A. C. Davison (2012), Bayesian inference from composite likelihoods, with an application to spatial extremes, *Stat. Sinica*, *22*, 813–845.
Rue, H., S. Martino, and N. Chopin (2009), Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations, *J. R. Stat. Soc. Ser. B*, *71*(2), 319–392.
Sang, H. (2015), Composite likelihood for extreme values, in *Extreme Value Modeling and Risk Analysis Methods and Applications*, Chapman Hall/CRC, Boca Raton.
Sang, H., and A. E. Gelfand (2009), Hierarchical modeling for extreme values observed over space and time, *Environ. Ecol. Stat.*, *16*(3), 407–426.
Sang, H., and A. E. Gelfand (2010), Continuous spatial process models for spatial extreme values, *J. Agric. Biol. Environ. Stat.*, *15*(1), 49–65.
Schlather, M. (2002), Models for stationary max-stable random fields, *Extremes*, *5*(1), 33–44.
Shang, H., J. Yan, and X. Zhang (2011), El Niño–Southern Oscillation influence on winter maximum daily precipitation in California in a spatial model, *Water Resour. Res.*, *47*(11), 1–9.
Stan Development Team (2015a), *Stan: A C++ Library for Probability and Sampling, Version 2.7.0*, Stan Development Team.
Stan Development Team (2015b), *RStan: The R Interface to Stan, Version 2.7.0*.
Stein, M. L., Z. Chi, and L. J. Welty (2004), Approximating likelihoods for large spatial data sets, *J. R. Stat. Soc. Ser. B*, *66*(2), 275–296.
Varin, C., N. Reid, and D. Firth (2011), An overview of composite likelihood methods, *Stat. Sinica*, *21*, 5–42.
Vecchia, V. (1988), Estimation and model identification of continuous spatial processes, *J. R. Stat. Soc. B*, *50*(2), 297–312.
Wang, Z., J. Yan, and X. Zhang (2014), Incorporating spatial dependence in regional frequency analysis, *Water Resour. Res.*, *50*, 9570–9585, doi:10.1002/2013WR014849.
Xia, Y., et al. (2012a), Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow, *J. Geophys. Res.*, *117*, D03110, doi:10.1029/2011JD016051.
Xia, Y., et al. (2012b), Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products, *J. Geophys. Res.*, *117*, D03109, doi:10.1029/2011JD016048.
Yan, H., and H. Moradkhani (2015), A regional Bayesian hierarchical model for flood frequency analysis, *Stochastic Environ. Res. Risk Assess.*, *29*(3), 1019–1036.