Supplementary materials for this article are available at https://doi.org/10.1007/s13253-020-00391-6.

Projecting Flood-Inducing Precipitation with a Bayesian Analogue Model

Gregory P. Bopp, Benjamin A. Shaby, Chris E. Forest, and Alfonso Mejía

The hazard of pluvial flooding is largely influenced by the spatial and temporal dependence characteristics of precipitation. When extreme precipitation possesses strong spatial dependence, the risk of flooding is amplified due to catchment factors such as topography that cause runoff accumulation. Temporal dependence can also increase flood risk as storm water drainage systems operating at capacity can be overwhelmed by heavy precipitation occurring over multiple days. While transformed Gaussian processes are common choices for modeling precipitation, their weak tail dependence may lead to underestimation of flood risk. Extreme value models such as the generalized Pareto processes for threshold exceedances and max-stable models are attractive alternatives, but are difficult to fit when the number of observation sites is large, and are of little use for modeling the bulk of the distribution, which may also be of interest to water management planners. While the atmospheric dynamics governing precipitation are complex and difficult to fully incorporate into a parsimonious statistical model, non-mechanistic analogue methods that approximate those dynamics have proven to be promising approaches to capturing the temporal dependence of precipitation. In this paper, we present a Bayesian analogue method that leverages large, synoptic-scale atmospheric patterns to make precipitation forecasts. Changing spatial dependence across varying intensities is modeled as a mixture of spatial Student-t processes that can accommodate both strong and weak tail dependence. The proposed model demonstrates improved performance at capturing the distribution of extreme precipitation over Community Atmosphere Model (CAM) 5.2 forecasts.
Supplementary materials accompanying this paper appear online.

Key Words: Dynamical system; Extreme value analysis; Stochastic weather generator; Student-t mixture.

G. P. Bopp (✉), Department of Statistics, Pennsylvania State University, 326 Thomas Building, University Park, PA 16802, USA (E-mail: *gxb951@psu.edu*). B. A. Shaby, Department of Statistics, Colorado State University, 211 Statistics Building, Fort Collins, CO 80523, USA (E-mail: *bshaby@colostate.edu*). Chris E. Forest, Department of Meteorology and Atmospheric Science; Department of Geosciences; Earth and Environmental Systems Institute, Pennsylvania State University, University Park, PA 16802, USA (E-mail: *ceforest@psu.edu*). A. Mejía, Department of Civil and Environmental Engineering, Pennsylvania State University, University Park, PA 16802, USA (E-mail: *aim127@psu.edu*).

Journal of Agricultural, Biological, and Environmental Statistics, Volume 25, Number 2, Pages 229–249
https://doi.org/10.1007/s13253-020-00391-6

1. INTRODUCTION

In this paper, we develop a mixture model for daily precipitation allowing for varying spatial dependence for different storm magnitudes. Instead of constructing a parametric formulation of the atmospheric dynamics regulating precipitation, we develop a Bayesian analogue method that can accommodate complex temporal dependence patterns, wherein precipitation analogues are established by identifying similar synoptic-scale atmospheric patterns from a historical library of observed climate states that are associated with precipitation outcomes.

Our goal is to make pluvial flood risk assessments under future climate projections. As such, we focus our modeling efforts on two key, intricately interacting, features. The first is the right tail of the precipitation distribution, since intense rains, all other things being equal, present the greatest flood risk. This does not mean we can ignore the bulk of the distribution, as moderately intense rainfall can also cause severe flooding, if the storm is widespread or persistent enough. This dovetails with our second key consideration, which is spatial dependence, because spatial characteristics are a key determinant of flood risk in a drainage basin.

Global climate models (GCMs) are the main tool for forecasting future climate conditions, providing a means to assess potential changes in the frequency and magnitudes of events such as heat waves, droughts and floods, all of which can have profound impacts on human health. Unfortunately, the coarse spatial resolution of GCMs cannot resolve fine scale hydrometeorological processes associated with precipitation extremes (Boé et al. 2006; Maraun et al. 2010). However, GCMs also produce smoothly varying atmospheric variables such as atmospheric pressure and temperature whose spatiotemporal patterns are closely linked with precipitation outcomes (Xoplaki et al. 2004; Raziei et al. 2012).

Analogue methods try to address the problem of making forecasts in the presence of complex temporal dependence without a model parameterization of any geophysical dynamics. In their simplest form, analogue methods match the current climate state to observed climate states from a library of historical observations in order to forecast some future quantity (e.g., precipitation tomorrow) with the observed successor of the historical match (e.g., precipitation on the day following the historical match). While the atmosphere is known to be a chaotic dynamical system that is unstable under slight perturbations of initial conditions (Lorenz 1969), analogue approaches are justified by the tendency of that system to regularly revisit subsets of the phase space over time. Analogue methods were originally developed as empirical tools for short-term weather forecasting (Krick 1942) and climate modeling (Barnett and Preisendorfer 1978), but researchers have begun to recognize their utility in a variety of contexts, including machine learning (Zhao and Giannakis 2016; Lguensat et al. 2017), wind-speed modeling (Nagarajan et al. 2015), and air quality monitoring (Delle Monache et al. 2014). Historically, analogue methods have been empirical, somewhat ad hoc tools, but recently there have been attempts at putting these into a probabilistic framework (McDermott and Wikle 2016; McDermott et al. 2018).

Precipitation has been the subject of extensive study as it plays a central role in agriculture, flood risk, and hydrology. Distributional forecasts of precipitation are critical for water management, infrastructure design, and developing disaster preparedness strategies. While

modern numerical forecasting weather models have achieved considerable success at making short-term forecasts by approximating solutions to the complex atmospheric processes governing the generation of precipitation, they typically do not account for uncertainty in their inputs. The function of stochastic weather generators, in contrast, is not to make short-term forecasts, but to accurately reproduce the distributional characteristics of precipitation on a fine spatial grid, while quantifying the uncertainty associated with those estimates (see Ailliot et al. 2015 for a review).

As precipitation is an inherently spatial phenomenon, several approaches have been proposed to model the spatial dependence features of both occurrences (the presence/absence of rain) and intensities (positive rain accumulations). A popular stochastic weather generator was proposed in a seminal paper by Wilks (1998) that models precipitation across multiple sites in two parts: (1) a two-state Markov process controlling the occurrence of precipitation at a given site and (2) a precipitation intensity model that accounts for spatial dependence between sites but ignores temporal dependence. Berrocal et al. (2008) and Kleiber et al. (2012) develop similar two-stage spatiotemporal models for observations from rain-gauge stations, wherein a latent Gaussian process is thresholded to model precipitation occurrence, while another marginally transformed Gaussian process controls precipitation intensity. While most models assume the same spatial dependence structure across precipitation intensities, Bárdossy and Pegram (2009) raise the issue of varying spatial dependence types across different intensities and aim to address it with an empirical copula approach. In this paper, we address the issue of varying spatial dependence types by modeling precipitation as a mixture of Student-t processes with different spatial correlation structures. Our approach to this problem is similar to that of Gelfand et al. (2005), who treat precipitation intensities as a Dirichlet process (DP) mixture of Gaussian processes. In the context of extreme value modeling, Fuentes et al. (2013) have also explored DP mixtures of Gaussian processes to flexibly model spatial dependence, but in the context of modeling maxima, which support marginal transformations to generalized extreme value (GEV) margins. Gaussian processes and their mixtures, however, are characterized by weak tail dependence, and in the presence of strong spatial dependence among extremes, their application can lead to underestimation of flood risk. To allow for stronger tail dependence, similar approaches have been taken by Morris et al. (2017) and Hazra et al. (2018) who use skew-t processes and their mixtures, which possess strong tail dependence, to model ozone and fire threat extremes, respectively.

Since extreme precipitation stands to do the most damage, accurately capturing the spatial dependence for high intensities necessitates special consideration. The last decade has produced many new methods for modeling spatial extremes. Two common approaches to modeling spatial extremes are based on liming results: max-stable processes (see Davison et al. 2012 for a review), which are appropriate for component-wise maxima over large temporal blocks, and generalized Pareto processes for high threshold exceedances (Ferreira and De Haan 2014). These models possess restrictive spatial dependence due to their max-stability and threshold stability properties. Moreover, while they are theoretically justified models for modeling asymptotic tail distributions, they are less useful for modeling the bulk of the distribution.

Along with limiting models, several Bayesian hierarchical models that borrow ideas from classical geostatistics have been developed (Cooley et al. 2007; Sang and Gelfand 2009;

Sang 2010; Reich and Shaby 2012; Zhang et al. 2019). Unlike in the classical setup to modeling extremes, which requires a somewhat arbitrary qualification of an extreme event (e.g., the threshold in a peaks-over-threshold model or block size in a model for block maxima), our proposed mixture model incorporates the entire distribution of data, while accommodating different dependence types for different storm intensities.

In addition to accurately representing the spatial dependence, capturing temporal dependence has also been central to the field of precipitation modeling. Several hidden Markov models for unobserved weather states have been proposed that aim to capture temporal dependence at various scales, including those attributable to large, synoptic-scale weather patterns (Bellone et al. 2000; Ailliot et al. 2009; Flecher et al. 2010). Latent multivariate autoregressive models are also common approaches to modeling temporal dependence for both occurrence and intensity processes (Bardossy and Plate 1992; Makhnin and McAllister 2009; Rasmussen 2013). Methods based in physics have also appeared in the statistical literature; recently, Liu et al. (2018) have proposed a Lagrangian advection reference frames approach to modeling storm dynamics that couples radar reflectivity and wind field data to make short-term precipitation forecasts.

In the remainder of the paper, we take up a similar aim to that of Gao et al. (2014) and Gao and Schlosser (2019) who use analogue methods to couple GCM forecasts of predictive atmospheric variables with the historical precipitation so as to understand the changes in the distributions of extreme precipitation under different climate forcing scenarios. Our method is different from these earlier efforts in that it is a hierarchical Bayesian model-based approach that makes use of the full data likelihood and can easily account for multiple sources of uncertainty. The following sections are outlined as follows: in Sect. 2, we develop a Bayesian models for precipitation occurrences and intensities that capture temporal dependence with a probabilistic analogue formulation. In Sect. 3, we apply the model to precipitation data over the northeastern USA and compare it with climate model and reanalysis distributional forecasts of extreme precipitation. Finally, in Sect. 4 we provide some concluding remarks and summary of the proposed method.

2. MODEL DEFINITIONS

One of the main challenges of modeling precipitation fields is that their distribution consists of a mixture of a preponderance of zeros and positive precipitation amounts. To account for this, the proposed spatiotemporal model for precipitation is made up of two parts: (1) an occurrence model for the presence versus absence of precipitation and (2) an intensity model for positive precipitation amounts (Wilks 1990; Berrocal et al. 2008). We begin by describing the occurrence model for precipitation.

2.1. OCCURRENCE MODEL

A common approach to modeling spatially varying, binary data is via data augmentation, wherein a continuous latent process is thresholded into two categories (Albert and Chib 1993; Heagerty and Lele 1998; Collett 2002). To model the point referenced, binary occurrence

of precipitation, we use a Gaussian process for the unobserved, latent component, such that it is positive at a location $\boldsymbol{s}$ whenever there is precipitation at $\boldsymbol{s}$ and negative otherwise. This commonly used model is referred to as a clipped Gaussian process or spatial probit model (De Oliveira 2000, 2020).

The spatiotemporal occurrence process $\{O_t(\boldsymbol{s}),\ \boldsymbol{s} \in \mathcal{S}\},\ t = 1, \ldots, T, \mathcal{S} \subset \mathbb{R}^2$, consists of spatial random fields that encode the presence versus absence of precipitation (1: presence, 0: absence) at a location $\boldsymbol{s}$ and time t. It is defined in terms of a zero-thresholded, latent spatiotemporal Gaussian process $\{Z_t(\boldsymbol{s}),\ \boldsymbol{s} \in \mathcal{S}\},\ t = 1, \ldots, T$:

$$O_t(\boldsymbol{s}) = \begin{cases} 1, & \text{if } Z_t(\boldsymbol{s}) > 0 \\ 0, & \text{if } Z_t(\boldsymbol{s}) \leq 0. \end{cases}$$

The latent processes $Z_t(\boldsymbol{s})$ (by abuse of notation) are parameterized by a mean function $\mu_t(\boldsymbol{s})$ and covariance function $C(\boldsymbol{s}, \boldsymbol{s}')$ that induce spatial dependence in the occurrence process to reflect the fact that nearby locations are more likely to have common presence or absence of rain than distant ones. Conditional on the mean and covariance functions, the spatially dependent processes are assumed to be independent in time, each distributed as:

$$Z_t(\boldsymbol{s}) \overset{\text{indep.}}{\sim} \text{GP}(\mu_t(\boldsymbol{s}), C(\boldsymbol{s}, \boldsymbol{s}')), \quad \text{for } t = 1, \ldots, T.$$

Temporal dependence is contained in the mean function. The covariance function can be expressed as a product of a variance parameter σ_Z^2 and correlation function $c(\boldsymbol{s}, \boldsymbol{s}')$ as $C(\boldsymbol{s}, \boldsymbol{s}') = \sigma_Z^2 c(\boldsymbol{s}, \boldsymbol{s}')$. However, since the variance term of this model is unidentifiable (De Oliveira 2020), it is fixed at $\sigma_Z^2 = 1$, making it sufficient to focus on the spatial correlation function. Due to the flexibility provided by a parameter governing the smoothness of the process, we assume a correlation function from the general Matérn class (Stein 1999), which is both stationary and isotropic, making it expressible as a function of distance between locations $h = ||\boldsymbol{s} - \boldsymbol{s}'||$: $c(\boldsymbol{s}, \boldsymbol{s})' = c_\nu(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu}\frac{h}{\rho}\right)^\nu K_\nu \left(\sqrt{2\nu}\frac{h}{\rho}\right),\ h \geq 0$ for range $\rho > 0$ and smoothness $\nu > 0$ parameters, where K_ν is a modified Bessel function of the second kind.

Using this construction, the marginal probability of occurrence at time t and site $\boldsymbol{s}$ can be expressed in terms of the mean function of the latent $Z_t(\boldsymbol{s})$ process as $\Pr(O_t(\boldsymbol{s}) = 1) = \Pr(Z_t(\boldsymbol{s}) > 0) = \Phi(\mu_t(\boldsymbol{s}))$, where Φ is a standard normal distribution function. To allow for spatially varying marginal occurrence probability, the mean function is further modeled as a linear combination of L spatial basis functions $\{\psi_l(\boldsymbol{s}) : \mathbb{R}^2 \to \mathbb{R}; \boldsymbol{s} \in \mathcal{S},\ l = 1, \ldots, L\}$. For the spatial basis functions, we use Gaussian kernels, although other choices are viable. For generic knot locations $\mathbf{v}_l \in \mathcal{S}, l = 1, \ldots, L$, the basis functions are defined as $\psi_l(\boldsymbol{s}) = \phi\left(\frac{||\boldsymbol{s} - \mathbf{v}_l||}{\Delta}\right)$, where ϕ denotes a standard Gaussian density function and $\Delta > 0$ is a bandwidth parameter. Denoting the vector of basis functions at location $\boldsymbol{s}$ as $\boldsymbol{\psi}(\boldsymbol{s}) = (\psi_1(\boldsymbol{s}), \ldots, \psi_L(\boldsymbol{s}))'$, we model the mean function at time t as a sum of an offset term $\gamma_t^{(O)}$ and a spatially varying term $\boldsymbol{\psi}(\boldsymbol{s})'\boldsymbol{\beta}_t^{(O)}$ as $\mu_t(\boldsymbol{s}) = \gamma_t^{(O)} + \boldsymbol{\psi}(\boldsymbol{s})'\boldsymbol{\beta}_t^{(O)}$. For the vector of spatial basis coefficients $\boldsymbol{\beta}_t^{(O)}$, we assume an independent normal prior $\boldsymbol{\beta}_t^{(O)} \overset{\text{iid}}{\sim} \text{N}_L(\mathbf{0}, \sigma^2_{\beta^{(O)}}\text{I}),\ t = 1, \ldots, T$. Note that in the presence of seasonality, a more complex prior that allows for

differing variances for different seasons may be necessary. However, in the sequel, we consider data from a single season in multiple years.

Since the presence or absence of precipitation is determined by whether $Z_t(\boldsymbol{s})$ is positive or negative, the $\gamma_t^{(O)}$ offset term can be thought of as governing the overall (non-spatially varying) probability of precipitation on day t. We will use this term to capture the temporal dependence in occurrence of precipitation by leveraging synoptic-scale climate forcings (e.g., geopotential height and temperature) that are concomitant with precipitation by construction of a prior on $\gamma_t^{(O)}$. We defer discussion of the analogue prior on $\gamma_t^{(O)}$, $t = 1, \ldots, T$ until Sect. 2.3 as it is also used in the model for precipitation intensities. The priors for the Matérn dependence parameters are taken to be $\rho \sim \text{Uniform}(\rho_l, \rho_u)$ and $\nu \sim \text{Uniform}(0, 2)$, where in subsequent sections ρ_l and ρ_u are taken to be the minimum and maximum distance between observation locations.

2.2. INTENSITY MODEL

In this section, we develop a Bayesian model for the positive spatial precipitation intensity process. The intensity process is treated as a continuous process defined on the entire spatial domain, $\mathcal{S}$, that is masked by the occurrence process. In other words, the intensity process is only observed at locations where the occurrence process is positive. We make use of the Student-t process because of its flexibility in capturing heavy tailed behavior, which is commonly observed in precipitation data (Vrac and Naveau 2007; Naveau et al. 2016).

Denote the precipitation intensity process by $\{Y_t^*(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}$ for times $t = 1, \ldots, T$. While precipitation intensities are strictly positive, it is much more convenient to work with a spatial process defined on the whole real line. The softplus function $f(x) = \log(\exp(x) - 1)$, maps $f : (0, \infty) \to \mathbb{R}$ and is strictly increasing, preserving the natural ordering of the data. Both the softplus function and its inverse leave moderate to large values effectively unchanged. In the remainder of this section, we define a model for the transformed precipitation intensities, $Y_t(\boldsymbol{s}) \equiv f(Y_t^*(\boldsymbol{s}))$.

A location-zero Student-t process can be expressed as a Gaussian process scale mixture (see e.g., Shah et al. 2014):

$$
\begin{aligned}
U(\boldsymbol{s}) &= \sigma \epsilon(\boldsymbol{s}) \\
\epsilon(\boldsymbol{s}) &\sim \text{GP}(\mathbf{0}, C(\boldsymbol{s}, \boldsymbol{s}')) \\
\sigma^2 &\sim \text{IG}\left(\frac{a}{2}, \frac{ab}{2}\right)
\end{aligned}
$$

where $\text{IG}(\frac{a}{2}, \frac{ab}{2})$ denotes a inverse-gamma distribution with shape $a/2$ and scale $ab/2$. After marginalizing over σ, $U(\boldsymbol{s})$ is a Student-t process with degrees of freedom a and scale b. A Gaussian process is a limiting case of a Student-t process as $a \to \infty$.

To allow flexible spatial dependence types across different precipitation intensity levels, the transformed precipitation amounts $Y_t(\boldsymbol{s})$ are modeled as a finite mixture of Student-t processes. Finite mixtures can easily be accommodated via data augmentation. Let $\xi_t \in \{1, \ldots, K\}$ denote the latent mixture label for time t, where K is the total number of mixture components. A multinomial logistic model, also referred to as the Luce model in

the probabilistic-choice econometrics literature when modeling latent classes (Luce 1959; McFadden 1973), is used for the mixture class membership. Denote covariates (e.g., containing synoptic-scale atmospheric information) by $\mathbf{u}_t \in \mathbb{R}^p$ for times $t = 1, \ldots, T$, vectors of coefficients for each class $k = 1, \ldots, K$ by $\boldsymbol{\alpha}_k \in \mathbb{R}^p$. Then, for linear predictor $\eta_{tk} = \mathbf{u}_t' \boldsymbol{\alpha}_k$, the probability $\pi_{t,k}$ that the process at time t belongs to mixture component class k is modeled as:

$$\pi_{tk} \equiv \Pr(\xi_t = k) = \frac{\exp(\eta_{tk})}{\sum_{j=1}^{K} \exp(\eta_{tj})}, \quad k = 1, \ldots, K;\ t = 1, \ldots, T. \tag{1}$$

For identifiability, the loadings for the Kth class are fixed $\boldsymbol{\alpha}_K = \mathbf{0}$, while the remaining loadings have independent normal priors:

$$\boldsymbol{\alpha}_k \overset{\text{iid}}{\sim} \mathrm{N}_p(\mathbf{0}, \sigma_\alpha^2 \mathrm{I}_p), \quad k = 1, \ldots, K-1.$$

Instead of using a common spatial dependence structure, each mixture class is free to have different spatial covariance, degrees of freedom a_k, and scale b_k parameters. Any spatial variation in the location surface at time t is captured by an offset term and a linear expansion of spatial basis functions $\mu_t(\boldsymbol{s}) = \gamma_t^{(I)} + \boldsymbol{\psi}(\boldsymbol{s})' \boldsymbol{\beta}_t^{(I)}$, just as was done in Sect. 2.1. Conditional on the mixture label at time t, $\xi_t = k$, the intensity process is modeled as follows:

$$\begin{aligned} Y_t(\boldsymbol{s}) &= \gamma_t^{(I)} + \boldsymbol{\psi}(\boldsymbol{s})' \boldsymbol{\beta}_t^{(I)} + \sigma_t \epsilon_t(\boldsymbol{s}) \\ \epsilon_t(\boldsymbol{s}) &\sim \mathrm{GP}(\mathbf{0}, c_k(\cdot, \cdot)) \\ \sigma_t^2 &\overset{indep}{\sim} \mathrm{IG}\left(\frac{a_k}{2}, \frac{a_k b_k}{2}\right). \end{aligned}$$

For each mixture class, we assume an isotropic Matérn correlation function with potentially different smoothness ν_k and range ρ_k parameters:

$$c_k(h) = \frac{2^{1-\nu_k}}{\Gamma(\nu_k)} \left(\sqrt{2\nu_k}\frac{h}{\rho_k}\right)^{\nu_k} K_{\nu_k}\left(\sqrt{2\nu_k}\frac{h}{\rho_k}\right), \quad k = 1, \ldots, K.$$

Just as in Sect. 2.1, we assume independent normal priors for the spatial basis coefficients $\boldsymbol{\beta}_t^{(I)} \overset{\text{iid}}{\sim} N(\mathbf{0}, \sigma_\beta^2 \mathrm{I})$, $t = 1, \ldots, T$, and we use an analogue prior for $\gamma_t^{(I)}$, $t = 1, \ldots, T$, which we discuss in Sect. 2.3. The following priors are used for the degrees of freedom and scale parameters: $a_k \overset{\text{iid}}{\sim} \mathrm{Uniform}(0, 30)$, $k = 1, \ldots, K$, and $b_k \overset{\text{iid}}{\sim} \mathrm{Gamma}(0.1, 10)$, where the Gamma distribution is parameterized with shape and scale, respectively. Priors for the Matérn covariance parameters ρ_k and ν_k are assumed to be independent across k and the same as those in the occurrence model.

Because we use a mixture of Student-t processes, different mixture components are capable of capturing different spatial dependence types, allowing for varying dependence strengths across different intensities. The upper tail dependence of a generic stationary and isotropic spatial process $\{W(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}$ can be characterized by the tail dependence function at level $u \in (0, 1)$, defined as:

$$\chi_u(h) = \Pr\left\{ W(\boldsymbol{s}+h) > F^{-1}(u) | W(\boldsymbol{s}) > F^{-1}(u) \right\}, \tag{2}$$

where F denotes the marginal distribution function of $W(\boldsymbol{s})$. The limit $\chi(h) = \lim_{u \to 1} \chi_u(h)$ determines the tail dependence class. The process is said to be asymptotically independent at distance h if $\chi(h) = 0$ and asymptotically dependent at distance h otherwise. Gaussian processes are asymptotically independent processes with $\chi(h) = 0$ for all $h > 0$, making them suitable for modeling physical phenomena exhibiting weak tail dependence (Sibuya 1960). Student-t processes and their finite mixtures possess asymptotic dependence making them suitable in the case of strong tail dependence, but also exhibit weakening spatial dependence at extreme but finite levels, making them flexible tools in practice (Nikoloulopoulos et al. 2009). Because a Gaussian process is a limiting case of the Student-t process, a Student-t process with large degrees of freedom is capable of capturing relatively weak dependence at sub-asymptotic levels.

2.3. ANALOGUE PRIOR

Both models described in Sects. 2.1 and 2.2 possess location offset terms $\gamma_t^{(O)}$ and $\gamma_t^{(I)}$, which here we will denote generically as γ_t, $t = 1, \ldots, T$. The priors for the offset parameters have thus far been left unspecified. In this section, we describe how the similarities between climate forcings at different times can be used to model the temporal dependence in the offset term. Following McDermott and Wikle (2016) and McDermott et al. (2018), who take a similar approach to forecasting soil moisture and waterfowl settling behavior, we will refer to this as an analogue prior on γ_t, $t = 1, \ldots, T$.

The presence and intensity of precipitation is closely related to other atmospheric conditions such as atmospheric pressure, temperature, and water vapor. Rather than explicitly model the complex precipitation dynamics, we use closely related atmospheric variables to identify historical analogues (times t', $t' \neq t$) of the atmospheric conditions at time t. Because of concomitance of atmospheric conditions with precipitation, the precipitation conditions during identified historical analogue times can then be used to inform the precipitation for the reference time. In a classical analogue model (Barnett and Preisendorfer 1978; Sugihara and May 1990), identified historical analogue precipitation fields and their weighted averages would be used as the forecasts for the reference time. Instead, we make a weaker assumption, and only borrow information about the location offset terms γ_t from analogue precipitation fields.

To formalize this, for each time $t = 1, \ldots, T$, define a vector of weights $\mathbf{w}_t = (w_{t,1}, \ldots, w_{t,t-1}, w_{t,t+1}, \ldots, w_{t,T})$, quantifying the similarity between the atmospheric conditions at time $t' \neq t$ and those at time t. For identifiability, the elements of the weight vectors are restricted to $w_{t,t'} \in [0,1]$, $w_{t,t} = 0$, and $\sum_{t' \neq t} w_{t,t'} = 1$, for all t. Combining the location parameters for other times $t' \neq t$ into a single vector $\boldsymbol{\gamma}_{-t} = (\gamma_1, \ldots \gamma_{t-1}, \gamma_{t+1}, \ldots, \gamma_T)'$, we specify conditional normal priors on the location parameters $\gamma_t | \boldsymbol{\gamma}_{-t} \sim N(\boldsymbol{\gamma}'_{-t}\mathbf{w}_t, \sigma^2_\gamma)$, $t = 1, \ldots, T$, so that the prior mean is a weighted average of the precipitation location parameters from historical conditions with strong similarity to the reference conditions.

The weights are calculated using a kernel function applied to distances between atmospheric conditions at different times, such that more similar historical atmospheric conditions receive larger weights. We use an unnormalized, compact Gaussian kernel function: $g(d;\theta,\tau) = \exp(\frac{-d^2}{2\theta})1\{\mathrm{d} < \tau\}$, $d > 0$, where d is a measure of distance, θ is a kernel bandwidth parameter, and τ is a threshold ensuring that the kernel is compact, which prevents irrelevant non-analogue days from receiving positive weight. In subsequent sections, it is fixed such that on average (across times t) only the top m nearest analogues receive weight. When the kernel bandwidth is large, many historical analogues will contribute a small weight, and when the bandwidth is small, fewer historical analogues will contribute a larger weight. Given distances $d_{t,t'}$ between all times t', $t' \neq t$ and t, unnormalized weights $w^*_{t,t'}$ are calculated as $w^*_{t,t'} = g(d_{t,t'};\theta,\tau)$, and normalized to give weights $w_{t,t'} = w^*_{t,t'} / \sum_{j\neq t} w^*_{t,j}$.

Due to the high-dimension of geopotential height fields and temperature fields, some dimension reduction is needed before calculating distances between atmospheric covariates. To reduce the dimension while preserving the relevant variation in the data, each of the atmospheric variables is projected onto the first several components of the empirical orthogonal functions (EOFs) calculated from the data (Jolliffe 2002; Hannachi et al. 2007; Demsar et al. 2013). Rather than rely on snapshots of the atmospheric conditions on any given day, we employ the time lagging approach of Takens (1981). Combining lagged EOF loadings into trajectories of atmospheric conditions over time gives a fuller reconstruction of the state-space of the dynamical system (Sugihara and May 1990). Let $\mathbf{x}_t$ denote the loadings for the first several components of relevant EOFs on day t, we construct a new embedding matrix as $X_t = [\mathbf{x}_t, \mathbf{x}_{t-1}, \ldots, \mathbf{x}_{t-r}]$, where r is the number of lagged time steps. Doing this for all time points, the quality of $Y_{t'}$ as an analogue for Y_t is determined by the distance (e.g., Euclidean norm) $d_{t,t'} = ||X_t - X_{t'}||$ between $X_{t'}$ and X_t.

To summarize the proposed approach, the precipitation process is broken into two components, occurrence and intensity, which are modeled independently. The spatial dependence in both processes are modeled using (1) a linear combination of basis functions in the mean function which captures smoothly varying spatial features and (2) a mixture of spatial stochastic processes which capture residual spatial dependence. To model the temporal dependence in both models, an analogue prior is used, wherein the random intercept of the process at each time has a conditional mean that is a weighted average of intercepts from times which share similar covariates to the current time.

For both the occurrence and intensity models, posterior samples are made using Markov chain Monte Carlo (MCMC), the details of which can be found in "Appendix A." Gibbs updates are available for some parameters, and the remaining are made with Metropolis–Hastings updates. Samples of the parameters are made for all observation times, and in-sample posterior predictive draws can be made directly, e.g., $Y_t(\boldsymbol{s}') \mid \mathbf{Y}_{1:T}$ for $t \in \{1,\ldots,T\}$ and non-observation location $\boldsymbol{s}' \notin \{\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n\}$. However, to make out-of-sample posterior predictive draws, for example of $Y_{t'}|\mathbf{Y}_{1:T}$ for time $t' > T$, distances between atmospheric conditions on the future day and historical days must first be calculated, e.g., $d_{t',t}$ as well as their corresponding normalized weights $w_{t',t}$ from $w^*_{t',t} = g(d_{t',t};\theta,\tau)$ for $t = 1,\ldots,T$, conditional on posterior samples of θ.

To assess the utility of the model, we perform two simulation studies, the results of which are summarized in the Supplementary Material.

3. PRECIPITATION ANALYSIS

In this section, we apply our model to daily precipitation observed over the Susquehanna river basin in southern New York and Pennsylvania. By coupling daily precipitation accumulations observed at rain-gauge stations with predictive atmospheric conditions, we can apply the proposed analogue model to make forecasts of precipitation over watersheds. The rain-gauge data come from the National Oceanic and Atmospheric Administration (NOAA) (https://www.ncdc.noaa.gov/ghcnd-data-access) and consist of daily precipitation accumulations (in millimeter) observed between 1986 and 2017 for $n = 174$ gauge stations. While precipitation data are often rounded to the nearest millimeter, we do not consider the role of quantization here. This is in part because our focus is on the bulk and right tail of the distribution where the rounding is comparatively negligible for larger precipitation amounts. To capture the temporal dependence, meteorological reanalysis estimates of geopotential height and surface temperature are used to construct the distance matrix for the analogue prior on location terms in the occurrence and intensity models. Reanalysis computer models infill meteorological fields on a spatial grid by assimilating historical, spatially varying atmospheric observations, which are treated as boundary conditions in a consistent model of the climate system. We consider 500 hPa geopotential height and surface temperature from the Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA-2) project (Gelaro et al. 2017) for identifying analogues. The analogue prior distance matrix is constructed using the lagged EOF loading approach described in Sect. 2.3 with $r = 3$ lagged time steps and the first 10 PCs of each of the two MERRA-2 atmospheric variables. An example of the 500 hPA and surface temperature fields for reference and nearest analogue days is shown in Fig. 1. For covariates $\mathbf{u}_t$, we use the EOF loadings of 500 hPa geopotential height and surface temperature at time t. Based on preliminary MCMC runs, we fix the number of mixture components $K = 5$ to allow for additional flexibility in spatial dependence types, but not so large as to generate nearly empty classes. Finally, we assume a $\Delta \sim \text{TN}(0, \sigma_\Delta^2, 0, \infty)$ prior, where σ_Δ is taken to be the 5th percentile of distances between all pairs of coordinates.

We evaluate three models for daily precipitation: occurrence and intensity models with (M1) independence priors on $\gamma_t^{(O)}$ and $\gamma_t^{(I)}$, (M2) analogue priors on $\gamma_t^{(O)}$ and $\gamma_t^{(I)}$ with distance matrix calculated using MERRA-2 atmospheric variables over the continental United States (CONUS), and (M3) analogue priors on $\gamma_t^{(O)}$ and $\gamma^{(I)}$ with distance matrix calculated using MERRA-2 atmospheric variables over the region surrounding Pennsylvania (local PA). The data are split into training (1986–2000) and holdout periods (2001–2017), during each year of which we consider only the months April, May, and June. During the training period, 20 stations are also held out for model evaluation of in-sample prediction (interpolation). Inference on each is performed with a Metropolis within Gibbs Markov chain Monte Carlo (MCMC) algorithm run for 100,000 iterations. The models are evaluated by two measures: by their ability to capture the distribution of precipitation when

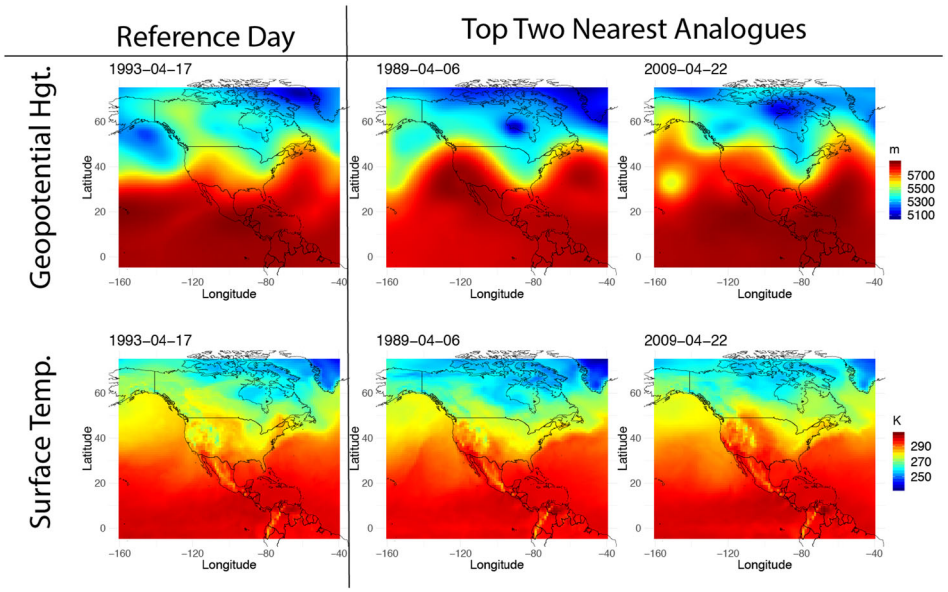

Figure 1. The 500 hPa geopotential heights and surface temperatures for a reference day are shown in the leftmost column. The middle and right columns show the corresponding fields for the top two most similar days based on the Euclidean distance between lagged vectors of the top 10 PC loadings of both variables.

making (1) out-of-sample forecasts and (2) in-sample, kriging predictions. To evaluate (1), out-of-sample posterior predictive draws are made from each model during holdout period at each of the 174 gauge locations, and to evaluate (2), in-sample posterior predictive draws are made at holdout stations for each day of the training period. In addition to these three models, we also consider the distributional forecasts of precipitation from the LENS historical run during the holdout period and MERRA-2 estimates during the training period. LENS fields of 500 hPa geopotential height and surface temperature are projected onto the MERRA-2 EOF basis for comparability of fields when calculating distances. Whereas LENS is used to make comparisons with out-of-sample predictions, MERRA-2 reanalysis estimates are used for in-sample (i.e., kriging) comparisons, because reanalysis models are effectively conditioning on the observed gauge measurements and atmospheric boundary conditions to infill precipitation.

Both MERRA-2 and LENS precipitation forecasts are made on a grid. To resolve the spatial mismatch between grid cells and gauge locations, gauge station forecasts for the LENS and MERRA-2 models are made by assigning each gauge station the precipitation amount from its nearest LENS and MERRA-2 grid cell.

The models are evaluated based on how well they capture the distributional characteristics of extreme precipitation. For each day of the holdout period, the maximum daily precipitation accumulation across all stations is calculated. The distribution of maxima is then compared to the predictions from each model using tail weighted continuously ranked probability scores (TWCRPS) (Gneiting and Katzfuss 2014). For a single sample y, the TWCRPS is calculated as:

Table 1. Tail weighted continuously ranked probability scores (TWCRPS) for the Student-t process mixture, LENS model forecasts of the distribution of daily maximum precipitation among all rain-gauge locations during the holdout period (2001–2017), and MERRA-2 model forecasts of the distribution of daily maximum precipitation among holdout gauge locations during the training period (1986–2000)

Predictions	TW fun.	Indep. prior	CONUS	Local PA	LENS	MERRA-2
In-sample	$w_1(x)$	7.43 (7.35, 7.55)	7.41 (7.32, 7.54)	7.37 (7.28, 7.48)		**7.30** (–,–)
In-sample	$w_2(x)$	6.66 (6.59, 6.77)	6.64 (6.56, 6.77)	6.61 (6.53, 6.70)		**6.54** (–,–)
Out-of-sample	$w_1(x)$	20.9 (13.1, 25.2)	16.3 (15.2, 17.4)	**11.8 (11.2, 12.5)**	13.7 (–,–)	
Out-of-sample	$w_2(x)$	19.6 (12.0, 23.7)	15.0 (14.0, 16.1)	**10.7 (10.2, 11.4)**	12.7 (–,–)	

The better value for each criterion is given in bold. Estimated standard errors are given in parentheses

$$\text{TWCRPS}\{\hat{F}, y\} = \int_{-\infty}^{\infty} (\hat{F}(x) - 1\{y \leq x\})^2 w(x) \mathrm{d}x \,, \tag{3}$$

where $\hat{F}$ refers to the model estimate of the target distribution from which y is drawn, and $w(x)$ is a weight function. In practice, as we do here, the TWCRPS is averaged over a sample $y_1, \ldots, y_{\tilde{T}}$ as $\text{TWCRPS} = \sum_{i=1}^{\tilde{T}} \text{TWCRPS}\{\hat{F}, y_i\}$ (where $\tilde{T} = T$ for in-sample prediction and $\tilde{T}$ is the number of holdout days for out-of-sample prediction). Lower TWCRPS scores correspond to better correspondence between the empirical and model-based distributions. We consider two TWCRPS weight functions: $w_1(x) = 1\{x \geq q_{0.5}\}$ and $w_2(x) = \Phi\{(x - q_{0.5})/s_{emp}\}$ where $q_{0.5}$ and s_{emp} are the sample median and standard deviation of the observed precipitation. The TWCRPS results are summarized in Table 1. Both versions of the analogue prior model outperform the independence prior model in capturing the distribution of precipitation extremes during the holdout period. Interestingly, the model with analogue distance matrix constructed from locally defined EOFs around PA shows greater predictive skill than both the one constructed from CONUS EOFs and the LENS model forecasts. Moreover, the in-sample forecasts of precipitation using the Local PA analogue prior model are also competitive with the MERRA-2 model. GCMs like LENS tend to oversmooth precipitation extremes. The proposed analogue model is able to incorporate information on temporal dependence encoded in the MERRA-2 and LENS models while preserving the heavy tailed nature and spatial features of extreme precipitation. We focus the remainder of our analysis on the analogue prior model using local PA fields.

A map of the pointwise, marginal 99th percentile estimates of the posterior predictive distribution of daily precipitation over the study region is shown in Fig. 2. Estimated precipitation levels tend to be highest over the northeastern part of the state as well as over parts of New Jersey and New York. To assess the correspondence between the empirical and model estimates of the spatial pattern of tail dependence, we also examine the F-madogram for monthly maxima (Cooley et al. 2006). The F-madogram is analogous to the more traditional madogram from classical geostatistics, but is guaranteed to exist even when the first moment of the process under consideration is undefined. For a generic, stationary and isotropic spatial process $\{W(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}\}$, $\mathcal{S} \subset \mathbb{R}^2$, with marginal distribution function F, and spatial lag $h > 0$, the F-madogram, defined as

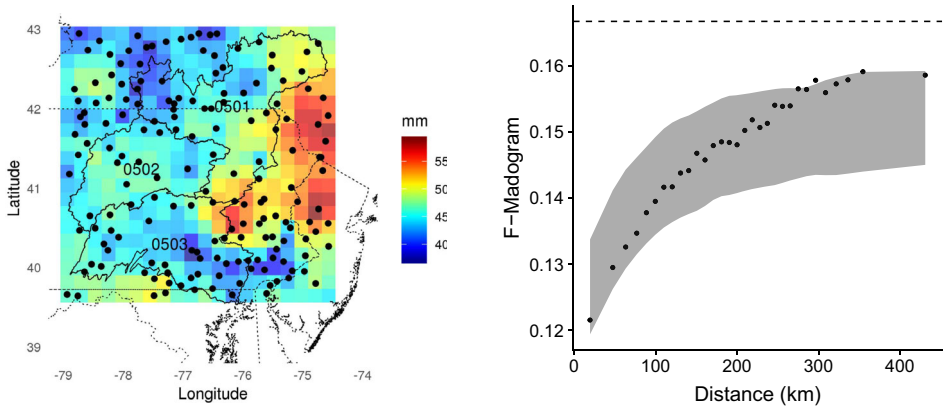

Figure 2. Left: Pointwise estimates of the 99th percentile of daily precipitation over the observation region based on the fitted analogue prior model overlaid with upper (0501), western (0502), and lower (0503) branch Susquehanna drainage basin boundaries. Right: Empirical F-madogram estimates (points) and 95% credible intervals (ribbons) for monthly maxima of daily precipitation accumulations based on the fitted Local PA analogue model. A dashed horizontal line corresponding to independence is plotted for reference.

$$\lambda_F(h) = \frac{1}{2}\mathrm{E}\left| F\{W(\boldsymbol{s}')\} - F\{W(\boldsymbol{s})\} \right|, \quad \text{for } ||\boldsymbol{s} - \boldsymbol{s}'|| = h\,, \tag{4}$$

and describes the dependence in the pairs $(W(\boldsymbol{s}), W(\boldsymbol{s}'))$ that are a distance h apart. When the pair is perfectly dependent, $\lambda_F(h) = 0$, and when the pair is independent, $\lambda_F(h) = 1/6$. Figure 2 shows both good correspondence between the empirical and model estimates of $\lambda_F(h)$ as well as apparent nonzero dependence of monthly maxima even at spatial lags of $h = 400$ km, which suggests that the model is capturing the spatial dependence properties of extreme precipitation well.

To assess the degree to which the Student-t mixture analogue model captures the spatial dependence across precipitation intensities, we compare empirical and model estimates of $\chi_u(h)$ for a fixed lag $h = 200$ km for varying u. The results are shown in Fig. 3. The overlap between model and empirical estimates suggests a good model fit of spatial dependence across intensities.

For illustration of the model-based forecasts, the observed daily precipitation accumulations, a posterior predictive draw, and posterior predictive mean and standard deviation for a single day are shown in Fig. 4. Observed and predicted zero accumulations are shown as gray. The figure shows good correspondence of the general spatial surface and smoothness characteristics in the observed and predicted precipitation amounts. Under the assumption that the library of observed historical climate states is sufficiently rich to match future states, our model can be used to make forecasts of future daily precipitation accumulations under various climate forcing scenarios. To do this, the 500 hPa geopotential height and surface temperature fields of the LENS RCP8.5 run are projected onto the first 10 principal components of their respective MERRA reanalysis fields. The distances between lagged trajectories of LENS EOF loadings and historical MERRA EOF loadings are calculated for all future days (e.g., for some future day t', the distance vector $\mathbf{d}_{t'}$ consists of Euclidean distances between covariates on day t' and all historical days $t = 1, \ldots, T$). LENS RCP8.5 forecasts

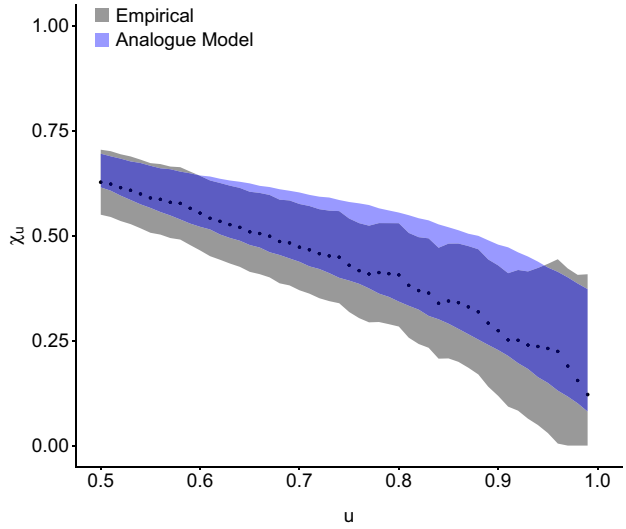

Figure 3. Empirical estimates of $\chi_u(h)$ (dots) at spatial lag $h = 200$ km and pointwise 95% confidence intervals (gray ribbon) are compared with Local PA Student-t mixture analogue model pointwise 95% credible intervals (blue ribbon). The overlap between model and empirical estimates suggests a good model fit of spatial dependence across precipitation intensities (Color figure online).

of geopotential height and surface temperature are coupled with the fitted analogue model to make out-of-sample posterior predictive draws of precipitation on a 10km^2 grid over three Susquehanna drainage basins from 2006–2100. Since the proposed model is defined on a continuous domain, predictions can be made on an arbitrarily fine grid. Summaries of the 3-month (April–May–June) maximum daily total precipitation volumes over basins are shown in Fig. 5. The figure shows high uncertainty in the estimates, with 95% credible bands capturing slight increasing trends over time for all three basins. To quantify the flood risk, the forecasts from this model could be input into a hydrological flow model that accounts for topography and land use among other factors.

4. DISCUSSION

While analogue methods have a long history in meteorology, there have been relatively few attempts at using them in a fully probabilistic framework for precipitation modeling. The analogue prior is a very general approach to modeling temporal dependence that leverages climate model forecasts of atmospheric variables that are concomitant with precipitation. Since this model is developed in a Bayesian framework, it is possible to account for uncertainty in the predictive skill of concomitant atmospheric variables that are used to select analogues. Moreover, the Student-t process mixture is a flexible model that can accommodate a wide variety of spatial dependence types in both the bulk and tail of the distribution. Because we use a fixed data transformation, rather than a classical copula approach, the mixture of t processes must carry the load of representing both the marginal and dependence characteristics of the data. Ordinarily, it is seen as advantageous to separate marginal and

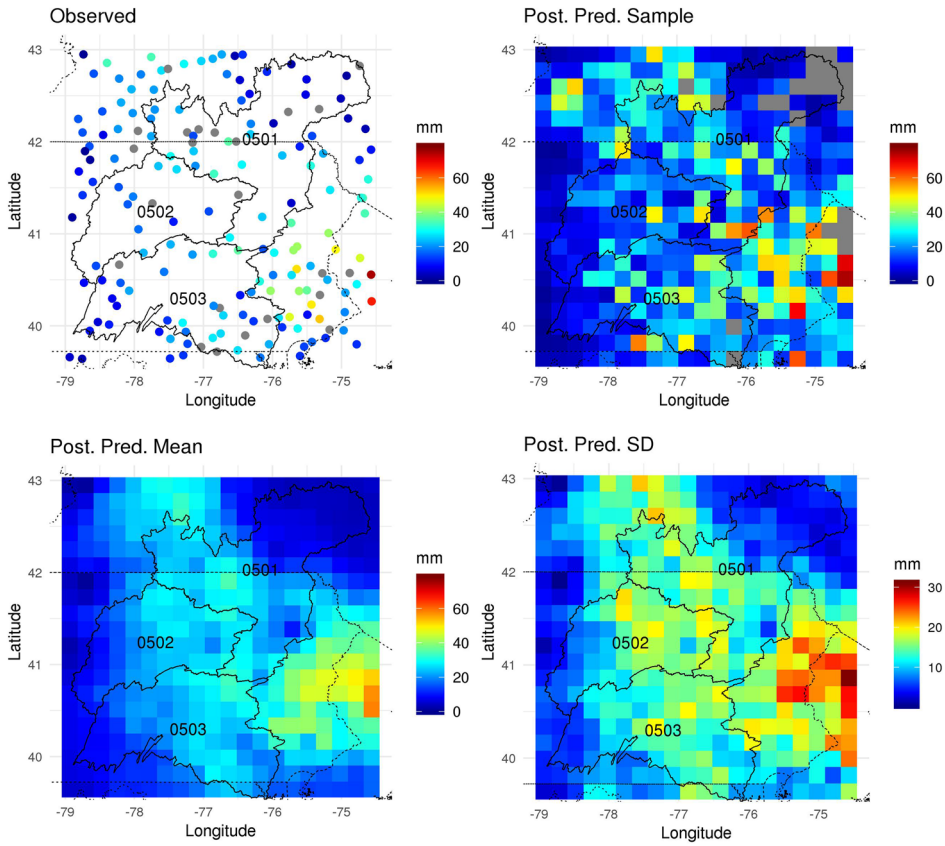

Figure 4. Observed daily precipitation amounts at gauge locations (top left), posterior predictive draw (top right), posterior predictive mean (bottom left), and posterior predictive standard deviation (bottom right) for a single day during the observation period. Gray points and cells correspond to zero precipitation amounts. The general spatial mean surface as well as degree of smoothness in the observed precipitation data is well captured by that of the posterior predictive sample.

dependence sub-models, but the mixture approach that we take here seems to be sufficiently flexible to provide excellent fits to, both marginally and jointly, to data arising from a variety of different mechanisms (Hazra et al. 2018).

This model could be extended further by considering nonlinear dimension reduction techniques for identifying analogues such as Laplacian eigenmaps (Belkin and Niyogi 2002), self-organizing maps (Kohonen 1984), and diffusion maps (Coifman and Lafon 2006). Similar to other analogue methods, the ability of the proposed model to produce accurate forecasts is predicated on the condition that there exist close historical analogues to future conditions (see Van den Dool 1994 for a discussion). As part of our ongoing and future work, we plan to apply our precipitation forecasts to a hydrological water-flow model to more directly interrogate changes in flood risk over the Susquehanna drainage basins.

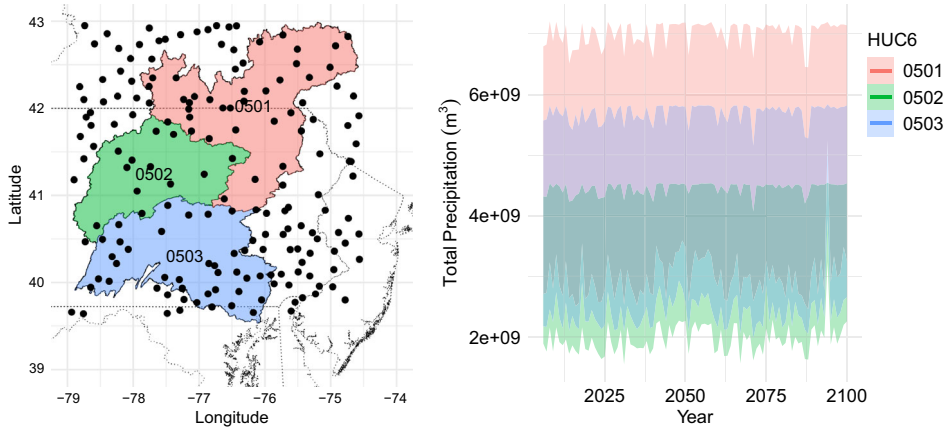

Figure 5. Global Historical Climate Network (GHCN) daily gauge station locations over Pennsylvania and surrounding states are overlaid with upper (0501), western (0502), and lower (0503) branch Susquehanna drainage basin boundaries (left). Forecasted yearly (during April–May–June) maximum total precipitation 95% credible intervals for each basin are plotted on the right.

ACKNOWLEDGEMENTS

This research was supported in part by the National Science Foundation (Grant No. DMS-1752280) and seed grants from the Institute for Computational and Data Sciences and the Institute for Energy and the Environment at the Pennsylvania State University. Computations for this research were performed on the Institute for Computational and Data Sciences Advanced CyberInfrastructure (ICDS-ACI).

[Received June 2019. Accepted March 2020. Published Online March 2020.]

A. MCMC DETAILS

Metropolis–Hastings MCMC algorithms were implemented for making posterior draws of the parameters in both the occurrence and intensity models using the R programming language (http://ww.r-project.org).

A.1. OCCURRENCE MODEL METROPOLIS–HASTINGS ALGORITHM

The parameters ρ, ν, and θ were updated using variable-at-a-time random walks. Truncated normal Gibbs updates are available for the latent Gaussian process at observation locations $\mathbf{Z}_t = (Z_t(\boldsymbol{s}_1), \ldots, Z_t(\boldsymbol{s}_n))$. Denoting occurrence indicators $\mathbf{O}_t = (O_t(\boldsymbol{s}_1), \ldots, O_t(\boldsymbol{s}_n))$ and mean function $\boldsymbol{\mu}_t = (\mu_t(\boldsymbol{s}_1), \ldots, \mu_t(\boldsymbol{s}_n))$, the Gibbs updates for the latent Gaussian process are

$$\mathbf{Z}_t | \mathbf{O}_t, \boldsymbol{\mu}_t, \rho, \nu \sim \mathrm{TN}_n(\boldsymbol{\mu}_t, \Sigma_{\nu,\rho}, \mathbf{l}_t, \mathbf{u}_t)$$

where $\Sigma_{\nu,\rho}$ is the spatial covariance matrix for the n observation locations, lower bounds $\mathbf{l}_t = (l(\boldsymbol{s}_1), \ldots, l(\boldsymbol{s}_n))$ have elements

$$l(\boldsymbol{s}_i) = \begin{cases} 0, & \text{if } O_t(\boldsymbol{s}_i) = 1 \\ -\infty & \text{if } O_t(\boldsymbol{s}_i) = 0 \end{cases}$$

and upper bounds $\mathbf{u}_t = (u(\boldsymbol{s}_1), \ldots, u(\boldsymbol{s}_n))$ have elements

$$u(\boldsymbol{s}_i) = \begin{cases} \infty, & \text{if } O_t(\boldsymbol{s}_i) = 1 \\ 0 & \text{if } O_t(\boldsymbol{s}_i) = 0 \end{cases} .$$

Gibbs updates are also available for both the location parameters $\gamma_t^{(O)}, t = 1, \ldots, T$ and $\boldsymbol{\beta}_t^{(O)}$. For the location parameter, the Gibbs update is

$$\gamma_t^{(O)} | \nu, \rho, \theta, \boldsymbol{\gamma}_{-t}^{(O)}, \sigma_\gamma^2, \boldsymbol{\beta}_t^{(O)}, \mathbf{Z}_t \sim \mathrm{N}\left\{ \left(\mathbf{1}' \Sigma_{\nu,\rho}^{-1} [\mathbf{Z}_t - \boldsymbol{\psi}' \boldsymbol{\beta}_t^{(O)}] + \frac{\mathbf{w}_t' \boldsymbol{\gamma}_{-t}^{(O)}}{\sigma^2_{\gamma^{(O)}}} \right) \times \left(\mathbf{1}' \Sigma_{\nu,\rho}^{-1} \mathbf{1} + \frac{1}{\sigma^2_{\gamma^{(O)}}} \right)^{-1}, \left(\mathbf{1}' \Sigma_{\nu,\rho}^{-1} \mathbf{1} + \frac{1}{\sigma^2_{\gamma^{(O)}}} \right)^{-1} \right\}.$$

For the basis coefficients $\boldsymbol{\beta}_t^{(O)}$, the Gibbs update is

$$\boldsymbol{\beta}_t^{(O)} | \mathbf{Z}_t, \gamma_t^{(O)}, \nu, \rho, \sigma^2_{\beta^{(O)}} \sim N_p \left\{ \left(\boldsymbol{\psi}' \Sigma_{\nu,\rho}^{-1} \boldsymbol{\psi} + \frac{1}{\sigma^2_{\beta^{(O)}}} I_p \right)^{-1} \times \left(\boldsymbol{\psi}' \Sigma_{\nu,\rho}^{-1} [\mathbf{Z}_t - \gamma_t^{(O)} \mathbf{1}] \right), \left(\boldsymbol{\psi}' \Sigma_{\nu,\rho}^{-1} \boldsymbol{\psi} + \frac{1}{\sigma^2_{\beta^{(O)}}} I_p \right)^{-1} \right\}.$$

Inverse-gamma Gibbs updates are used for $\sigma^2_{\gamma^{(O)}}$ and $\sigma^2_{\beta^{(O)}}$. Using inverse gamma parameterized with shape a and scale b, having density $f(x; a, b) = \frac{b^a}{\Gamma(a)} x^{-(a+1)} \exp(-\frac{b}{x}),\ x > 0$. With prior $\sigma^2_{\gamma^{(O)}} \sim \mathrm{IG}(a_\gamma, b_\gamma)$, the Gibbs update for $\sigma^2_{\gamma^{(O)}}$, is

$$\sigma^2_{\gamma^{(O)}} | \boldsymbol{\gamma}_t^{(O)} \sim \mathrm{IG} \left\{ a_\gamma + n/2, b_\gamma + \frac{1}{2} \sum_{t=1}^{T} (\gamma_t - \mathbf{w}_t' \boldsymbol{\gamma}_{-t}{}^{(O)})^2 \right\},$$

and the Gibbs update for $\sigma^2_{\beta^{(O)}}$, assuming prior $\sigma^2_{\beta^{(O)}} \sim \mathrm{IG}(a_\beta, b_\beta)$ is

$$\sigma^2_{\beta^{(O)}} | \boldsymbol{\beta}_{1:T}^{(O)} \sim \mathrm{IG} \left\{ a_\beta + np/2, b_\beta + \frac{1}{2} \sum_{t=1}^{T} \boldsymbol{\beta}_t^{(O)'} \boldsymbol{\beta}_t^{(O)} \right\},$$

A.2. INTENSITY MODEL METROPOLIS–HASTINGS ALGORITHM

The parameters $\rho, \nu, a_k, b_k, \boldsymbol{\alpha}_k$ for $k = 1, \ldots, K$, and θ were updated using variable-at-a-time random walks. The cluster labels ξ_t were also updated variable-at-a-time, but with discrete uniform independence proposals, each on $1{:}K$. The location offset and basis

functions Gibbs updates are similar to those in the occurrence model but are also dependent on the mixture labels.

$$\gamma_t^{(I)} | \nu_k, \rho_k, \theta, \boldsymbol{\gamma}_{-t}^{(I)}, \xi_t = k, \sigma_\gamma^2, \boldsymbol{\beta}_t^{(I)}, \mathbf{Z}_t \sim \mathrm{N}\left\{ \left(\mathbf{1}' \Sigma_k^{-1} [\mathbf{Y}_t - \boldsymbol{\psi}' \boldsymbol{\beta}_t^{(I)}] + \frac{\mathbf{w}_t' \boldsymbol{\gamma}_{-t}^{(I)}}{\sigma^2_{\gamma^{(I)}}} \right) \right.$$
$$\left. \times \left(\mathbf{1}' \Sigma_k^{-1} \mathbf{1} + \frac{1}{\sigma^2_{\gamma^{(I)}}} \right)^{-1}, \left(\mathbf{1}' \Sigma_k^{-1} \mathbf{1} + \frac{1}{\sigma^2_{\gamma^{(I)}}} \right)^{-1} \right\}.$$

where the covariance matrix Σ_k is calculated using dependence parameters ρ_k and ν_k corresponding to mixture class k.

Similarly, the basis coefficients $\boldsymbol{\beta}_t^{(I)}$ have Gibbs update

$$\boldsymbol{\beta}_t^{(I)} | \mathbf{Y}_t, \gamma_t^{(I)}, \nu_k, \rho_k, \xi_t = k, \sigma^2_{\beta^{(I)}} \sim N_p \left\{ \left(\boldsymbol{\psi}' \Sigma_k^{-1} \boldsymbol{\psi} + \frac{1}{\sigma^2_{\beta^{(I)}}} I_p \right)^{-1} \right.$$
$$\left. \times \left(\boldsymbol{\psi}' \Sigma_k^{-1} [\mathbf{Y}_t - \gamma_t^{(I)} \mathbf{1}] \right), \left(\boldsymbol{\psi}' \Sigma_k^{-1} \boldsymbol{\psi} + \frac{1}{\sigma^2_{\beta^{(I)}}} I_p \right)^{-1} \right\}.$$

The Gibbs updates for the prior variances $\sigma^2_{\gamma^{(I)}}$ and $\sigma^2_{\beta^{(I)}}$ are completely analogous to those in the occurrence model.

REFERENCES

Ailliot, P., D. Allard, V. Monbet, and P. Naveau (2015). Stochastic weather generators: an overview of weather type models. *Journal de la Société Française de Statistique 156*(1), 101–113.

Ailliot, P., C. Thompson, and P. Thomson (2009). Space–time modelling of precipitation by using a hidden Markov model and censored Gaussian distributions. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 58*(3), 405–426.

Albert, J. H. and S. Chib (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American statistical Association 88*(422), 669–679.

Bárdossy, A. and G. Pegram (2009). Copula based multisite model for daily precipitation simulation. *Hydrology and Earth System Sciences 13*(12), 2299–2314.

Bardossy, A. and E. J. Plate (1992). Space-time model for daily rainfall using atmospheric circulation patterns. *Water Resources Research 28*(5), 1247–1259.

Barnett, T. and R. Preisendorfer (1978). Multifield analog prediction of short-term climate fluctuations using a climate state vector. *Journal of the Atmospheric Sciences 35*(10), 1771–1787.

Belkin, M. and P. Niyogi (2002). Laplacian eigenmaps and spectral techniques for embedding and clustering. In *Advances in neural information processing systems*, pp. 585–591.

Bellone, E., J. P. Hughes, and P. Guttorp (2000). A hidden Markov model for downscaling synoptic atmospheric patterns to precipitation amounts. *Climate research 15*(1), 1–12.

Berrocal, V. J., A. E. Raftery, and T. Gneiting (2008). Probabilistic quantitative precipitation field forecasting using a two-stage spatial model. *The Annals of Applied Statistics 2*(4), 1170–1193.

Boé, J., L. Terray, F. Habets, and E. Martin (2006). A simple statistical-dynamical downscaling scheme based on weather types and conditional resampling. *Journal of Geophysical Research: Atmospheres 111*(D23106).

Coifman, R. R. and S. Lafon (2006). Diffusion maps. *Applied and computational harmonic analysis 21*(1), 5–30.

Collett, D. (2002). *Modelling binary data*. CRC press, Boca Raton, FL.

Cooley, D., P. Naveau, and P. Poncet (2006). Variograms for spatial max-stable random fields. In *Dependence in probability and statistics*, pp. 373–390. Springer.

Cooley, D., D. Nychka, and P. Naveau (2007). Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association 102*(479), 824–840.

Davison, A. C., S. A. Padoan, and M. Ribatet (2012). Statistical modeling of spatial extremes. *Statistical Science. A Review Journal of the Institute of Mathematical Statistics 27*(2), 161–186.

De Oliveira, V. (2000). Bayesian prediction of clipped Gaussian random fields. *Computational Statistics & Data Analysis 34*(3), 299–314.

De Oliveira, V. (2020). Models for Geostatistical Binary Data: Properties and Connections. *Amer. Statist. 74*(1), 72–79.

Delle Monache, L., I. Djalalova, and J. Wilczak (2014). Analog-based postprocessing methods for air quality forecasting. In *Air Pollution Modeling and its Application XXIII*, pp. 237–239. Springer.

Demsar, U., P. Harris, C. Brunsdon, A. S. Fotheringham, and S. McLoone (2013). Principal component analysis on spatial data: an overview. *Annals of the Association of American Geographers 103*(1), 106–128.

Ferreira, A. and L. De Haan (2014). The generalized Pareto process; with a view towards application and simulation. *Bernoulli 20*(4), 1717–1737.

Flecher, C., P. Naveau, D. Allard, and N. Brisson (2010). A stochastic daily weather generator for skewed data. *Water Resources Research 46*(7).

Fuentes, M., J. Henry, and B. Reich (2013). Nonparametric spatial models for extremes: Application to extreme temperature data. *Extremes 16*(1), 75–101.

Gao, X. and C. A. Schlosser (2019). Mid-western us heavy summer-precipitation in regional and global climate models: the impact on model skill and consensus through an analogue lens. *Climate Dynamics 52*(3), 1569–1582.

Gao, X., C. A. Schlosser, P. Xie, E. Monier, and D. Entekhabi (2014). An analogue approach to identify heavy precipitation events: Evaluation and application to CMIP5 climate models in the United States. *Journal of Climate 27*(15), 5941–5963.

Gelaro, R., W. McCarty, M. J. Suárez, R. Todling, A. Molod, L. Takacs, C. A. Randles, A. Darmenov, M. G. Bosilovich, and R. Reichle (2017). The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate 30*(14), 5419–5454.

Gelfand, A. E., A. Kottas, and S. N. MacEachern (2005). Bayesian nonparametric spatial modeling with Dirichlet process mixing. *Journal of the American Statistical Association 100*(471), 1021–1035.

Gneiting, T. and M. Katzfuss (2014). Probabilistic forecasting. *Annual Review of Statistics and Its Application 1*, 125–151.

Hannachi, A., I. Jolliffe, and D. Stephenson (2007). Empirical orthogonal functions and related techniques in atmospheric science: A review. *International journal of climatology 27*(9), 1119–1152.

Hazra, A., B. J. Reich, B. A. Shaby, and A.-M. Staicu (2018). A semiparametric Bayesian model for spatiotemporal extremes. *arXiv preprint* arXiv:1812.11699.

Heagerty, P. J. and S. R. Lele (1998). A composite likelihood approach to binary spatial data. *Journal of the American Statistical Association 93*(443), 1099–1111.

Jolliffe, I. T. (2002). *Principal component analysis* (Second ed.). Springer Series in Statistics. Springer-Verlag, New York.

Kleiber, W., R. W. Katz, and B. Rajagopalan (2012). Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes. *Water Resources Research 48*(1).

Kohonen, T. (1984). *Self-organization and associative memory*, Volume 8 of *Springer Series in Information Sciences*. Springer-Verlag, Berlin.

Krick, I. P. (1942). *A Dynamical Theory of the Atmospheric Circulation and Its Use in Weather Forecasting: Studies of Persistent Regularities in Weather Phenomena*. California Institute of Technology.

Lguensat, R., P. Tandeo, P. Ailliot, M. Pulido, and R. Fablet (2017). The analog data assimilation. *Monthly Weather Review 145*(10), 4093–4107.

Liu, X., V. Gopal, and J. Kalagnanam (2018). A spatio-temporal modeling framework for weather radar image data in tropical southeast Asia. *The Annals of Applied Statistics 12*(1), 378–407.

Lorenz, E. N. (1969). Atmospheric predictability as revealed by naturally occurring analogues. *Journal of the Atmospheric sciences 26*(4), 636–646.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. Courier Corporation.

Makhnin, O. V. and D. L. McAllister (2009). Stochastic precipitation generation based on a multivariate autoregression model. *Journal of Hydrometeorology 10*(6), 1397–1413.

Maraun, D., F. Wetterhall, A. Ireson, R. Chandler, E. Kendon, M. Widmann, S. Brienen, H. Rust, T. Sauter, and M. Themeßl (2010). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics 48*(3).

McDermott, P. L. and C. K. Wikle (2016). A model-based approach for analog spatio-temporal dynamic forecasting. *Environmetrics 27*(2), 70–82.

McDermott, P. L., C. K. Wikle, and J. Millspaugh (2018). A hierarchical spatiotemporal analog forecasting model for count data. *Ecology and evolution 8*(1), 790–800.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior.

Morris, S. A., B. J. Reich, E. Thibaud, and D. Cooley (2017). A space-time skew-t model for threshold exceedances. *Biometrics 73*(3), 749.

Nagarajan, B., L. Delle Monache, J. P. Hacker, D. L. Rife, K. Searight, J. C. Knievel, and T. N. Nipen (2015). An evaluation of analog-based postprocessing methods across several variables and forecast models. *Weather and Forecasting 30*(6), 1623–1643.

Naveau, P., R. Huser, P. Ribereau, and A. Hannart (2016). Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection. *Water Resources Research 52*(4), 2753–2769.

Nikoloulopoulos, A. K., H. Joe, and H. Li (2009). Extreme value properties of multivariate t copulas. *Extremes 12*(2), 129–148.

Rasmussen, P. (2013). Multisite precipitation generation using a latent autoregressive model. *Water Resources Research 49*(4), 1845–1857.

Raziei, T., A. Mofidi, J. A. Santos, and I. Bordi (2012). Spatial patterns and regimes of daily precipitation in Iran in relation to large-scale atmospheric circulation. *International Journal of Climatology 32*(8), 1226–1237.

Reich, B. J. and B. A. Shaby (2012). A hierarchical max-stable spatial model for extreme precipitation. *The Annals of Applied Statistics 6*(4), 1430–1451.

Sang, H. and A. E. Gelfand (2009). Hierarchical modeling for extreme values observed over space and time. *Environmental and ecological statistics 16*(3), 407–426.

Sang, H. and A. E. Gelfand (2010). Continuous spatial process models for spatial extreme values. *Journal of Agricultural, Biological, and Environmental Statistics 15*(1), 49–65.

Shah, A., A. Wilson, and Z. Ghahramani (2014). Student-t processes as alternatives to Gaussian processes. In *Artificial Intelligence and Statistics*, pp. 877–885.

Sibuya, M. (1960). Bivariate extreme statistics. I. *Ann. Inst. Statist. Math. Tokyo 11*, 195–210.

Stein, M. L. (1999). *Interpolation of spatial data: some theory for kriging*. Springer-Verlag, New York, NY.

Sugihara, G. and R. M. May (1990). Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature 344*(6268), 734.

Takens, F. (1981). Detecting strange attractors in turbulence. In *Dynamical systems and turbulence, Warwick 1980 (Coventry, 1979/1980)*, Volume 898 of *Lecture Notes in Math.*, pp. 366–381. Springer, Berlin-New York.

Van den Dool, H. (1994). Searching for analogues, how long must we wait? *Tellus A 46*(3), 314–324.

Vrac, M. and P. Naveau (2007). Stochastic downscaling of precipitation: From dry events to heavy rainfalls. *Water resources research 43*(7).

Wilks, D. (1998). Multisite generalization of a daily stochastic precipitation generation model. *Journal of Hydrology 210*(1-4), 178–191.

Wilks, D. S. (1990). Maximum likelihood estimation for the gamma distribution using data containing zeros. *Journal of Climate 3*(12), 1495–1501.

Xoplaki, E., J. González-Rouco, J. Luterbacher, and H. Wanner (2004). Wet season mediterranean precipitation variability: influence of large-scale dynamics and trends. *Climate dynamics 23*(1), 63–78.

Zhang, L., B. A. Shaby, and J. L. Wadsworth (2019). Hierarchical transformed scale mixtures for flexible modeling of spatial extremes on datasets with many locations. *arXiv e-prints*, arXiv:1907.09617.

Zhao, Z. and D. Giannakis (2016). Analog forecasting with dynamics-adapted kernels. *Nonlinearity 29*(9), 2888–2939.