NIH Public Access
Author Manuscript
J R Stat Soc Ser C Appl Stat. Author manuscript; available in PMC 2013 July 14.

Published in final edited form as:
J R Stat Soc Ser C Appl Stat. 2012 August ; 61(4): 535–553. doi:10.1111/j.1467-9876.2011.01025.x.

Spatiotemporal quantile regression for detecting distributional changes in environmental processes

Brian J Reich
Department of Statistics, North Carolina State University

Abstract

Climate change may lead to changes in several aspects of the distribution of climate variables, including changes in the mean, increased variability, and severity of extreme events. In this paper, we propose using spatiotemporal quantile regression as a flexible and interpretable method for simultaneously detecting changes in several features of the distribution of climate variables. The spatiotemporal quantile regression model assumes that each quantile level changes linearly in time, permitting straight-forward inference on the time trend for each quantile level. Unlike classical quantile regression which uses model-free methods to analyze a single quantile or several quantiles separately, we take a model-based approach which jointly models all quantiles, and thus the entire response distribution. In the spatiotemporal quantile regression model, each spatial location has its own quantile function that evolves over time, and the quantile functions are smoothed spatially using Gaussian process priors. We propose a basis expansion for the quantile function that permits a closed-form for the likelihood, and allows for residual correlation modeling via a Gaussian spatial copula. We illustrate the methods using temperature data for the southeast US from the years 1931–2009. For these data, borrowing information across space identifies more significant time trends than classical non-spatial quantile regression. We find a decreasing time trend for much of the spatial domain for monthly mean and maximum temperatures. For the lower quantiles of monthly minimum temperature, we find a decrease in Georgia and Florida, and an increase in Virginia and the Carolinas.

Keywords

Bayesian hierarchical model; climate change; non-Gaussian data; US temperature data; warming hole

1 Introduction

A vast literature on climate change research has emerged in recent years. One aspect of climate change research that has received considerable attention is global warming, which refers to increasing trends in the mean temperature. However, climate change may lead to changes in several aspects of the distribution of climate variables, including changes in the mean (e.g., Soloman et al, 2007), increased variability (e.g., Katz and Brown, 1992; Chandler, 2005), and severity of extreme events (e.g, Easterling et al., 2000). Simultaneous analysis of many aspects of climate change is a challenging statistical problem. In this paper, we propose using spatiotemporal quantile regression as a flexible and interpretable method for detecting changes in the distribution of climate variables.

We illustrate this method by analyzing monthly temperature data for the southeast US for years 1931–2009. Despite a global increase, this region has experienced little to no increase in mean summer temperature (Folland et al., 2001; Robinson, Reudy, and Hansen, 2002; Pan et al., 2004; Kunkel et al., 2006). Pan et al. (2004) describe this as a “warming hole”,

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

and suggest this may be due to a feedback process in the hydrological cycle. Here we extend the analysis of US temperatures beyond the mean to other features of the distribution. Each month of each year we observe the monthly minimum, maximum, and mean temperatures for 191 locations covering Florida, Georgia, North Carolina, South Carolina, and Virginia. For example, Figure 1 plots the minimum temperature for July by year for a representative location. This plot suggests that the distribution of temperature is gradually changing with time at this location. A simple linear regression with covariate year gives an estimated increase of 0.04 degrees (F) per year (the p-value for the test that the slope is zero is 0.053) in the mean. Quantile regression provides a more comprehensive analysis. In a linear quantile regression, the $100\tau^{th}$ percentile in year t is modeled as $q_0(\tau) + q_1(\tau)t$. By studying the slopes $q_1(\tau)$ for different quantile levels τ, we analyze changes in different aspects of the temperature distribution. For example, Figure 1b shows that there is a statistically significant increase in distribution's lower tail (10th percentile), but not the center of the distribution (50th percentile, i.e., median) or the magnitude of extremely high temperatures (90th percentile). Combining these results suggests there is less variability in the current climate. Of course these results are specific to this spatial location, and the distribution of temperature may be evolving differently in other locations. The objective of our analysis of these data is to identify distributional changes in temperature over time and compare these changes across different spatial locations.

We develop a spatiotemporal model that allows the entire distribution of the response to vary over space and time. Several nonparametric Bayesian methods have been developed to model a spatially-varying distribution function (Gelfand et al., 2005; Griffin and Steel, 2006; Reich and Fuentes, 2007; Dunson and Park, 2008). These methods treat the conditional distribution of the response at each spatial location as an unknown quantity to be estimated from the data. Although it may be possible to extend these nonparametric methods to allow the distribution to vary with both space and time, we elect to follow Reich, Fuentes, and Dunson (2011; henceforth RFD) and model the distribution via its quantile function.

Rather than restricting the analysis to changes in the mean response, quantile regression permits simultaneous analysis of several features of the response distribution, including the median, variance, and severity of extreme events. Our spatiotemporal quantile regression model assumes that each quantile level changes linearly in time, permitting straight-forward inference on the time trend for each quantile level. Classical quantile regression (Yu, Lu, and Stander, 2003; Koenker, 2005) uses model-free methods to analyze a single quantile or several quantiles separately. This approach of analyzing single quantile levels has recently been extended to handle spatial data (Lee and Neocleous, 2010; Lum, 2010; Oh, Lee, and Nychka, 2011; Sobotka and Kneib, 2011). In contrast, we take a model-based approach which jointly models all quantiles, and thus the entire response distribution. This has been shown to provide more accurate coefficient estimation and higher power for detecting important predictors than analyzing individual quantiles separately (Bondell et al., 2011; Reich et al., 2011; Tokdar and Kadane, 2011). In the spatial quantile regression model, each spatial location has its own quantile function that evolves over time, and the quantile functions are smoothed spatially using Gaussian process priors.

The proposed spatial quantile regression model differs from RFD in several important ways. Although the RFD model for the quantile function has a relatively simple form, the resulting probability density function (i.e., the derivative of the inverse quantile function) does not have a closed form. As a result, the likelihood does not have closed form, which leads to computational difficulty. RFD resort to a two-stage model rather than a fully-Bayesian analysis for large datasets. In constrast, the model for the quantile function in this paper permits a closed-form expression for the response distribution and leads to conjugate full conditional distributions for the parameters in the quantile function. This allows the fully-

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Bayesian model to be fit to large spatiotemporal datasets. Also, although RFD allow the quantile process to vary over space and time, residual spatiotemporal correlation given the quantile process is ignored. As a result, the model performs well for prediction, but the posterior credible sets of model parameters have undercoverage in the presence of strong spatiotemporal association. This is problematic because the primary objective in this paper is to test for trends the quantile process, and the data exhibit strong spatiotemporal association. Therefore, we propose a spatiotemporal copula model to account for residual spatiotemporal association. We show that the quantile regression model is amenable to the copula approach, since it is the quantile function that is required for inverse transform sampling. The proposed copula model gives conjugate full conditional distributions, which leads to straight-forward MCMC coding and tuning.

The remainder of the paper is organized as follows. Section 2 details the spatiotemporal quantile regression model. Section 3 describes the computational algorithm. The method is applied to southeast US temperature data for years 1931–2009 in Section 4. Section 5 concludes.

2 Spatiotemporal quantile regression

Let $y_t(\mathbf{s})$, the response at spatial location $\mathbf{s} \in \mathbb{R}^2$ and time $t \in [0, 1]$, be $y_t(\mathbf{s}) = \mu_t(\mathbf{s}) + \varepsilon_t(\mathbf{s})$, where $\mu_t(\mathbf{s})$ is the true value at time t and location $\mathbf{s}$, and $\varepsilon_t(\mathbf{s}) \overset{iid}{\sim} \mathrm{N}(0, \sigma^2)$ is small-scale error. For example, the temperature data analyzed in Section 4 are rounded to the nearest degree, and so the small-scale errors may include these rounding errors. As such, we treat these measurement errors as independent and attribute all spatial and temporal dependence to μ. Classical quantile regression with independent observations focuses on estimating the quantile function of the response, $y_t(\mathbf{s})$. However, in the presence of strong spatial or temporal correlation, it is possible to bypass the error process and directly model the quantile function of the true underlying spatiotemporal process of interest, $\mu_t(\mathbf{s})$, as is the norm for geostatistical analysis of Gaussian data (Cressie, 1993; Banerjee et al., 2004; Gelfand et al., 2010). Therefore, we develop a model for the quantile function of $\mu_t(\mathbf{s})$. However, including small-scale error is not required, and Appendix A.1 outlines an approach for modeling the quantiles of $y_t(\mathbf{s})$ directly.

The quantile function $q(\tau|\mathbf{s}, t)$ is the function that satisfies $\mathrm{P}[\mu_t(\mathbf{s}) < q(\tau|\mathbf{s}, t)] = \tau \in [0, 1]$. To test for changes in the quantile function over time, each quantile is a linear function of time, $q(\tau|\mathbf{s}, t) = q_0(\tau|\mathbf{s}) + tq_1(\tau|\mathbf{s})$. We note that a bounded time variable t is required. If the time variable is unbounded, then the only valid quantile model is the location-shift model with $q_1(\tau|\mathbf{s})$ constant for all τ, as discussed further below. Both the intercept function, q_0, and the trend function, q_1, are taken to be linear combinations of L basis functions,

$$q_k(\tau|\mathbf{s}) = \beta_k(\mathbf{s}) + \sum_{l=1}^{L} B_l(\tau)\theta_{kl}(\mathbf{s}), \quad (1)$$

where $\beta_k(\mathbf{s})$ is the center (the median in the specification in Section 2.1) of the quantile function at location $\mathbf{s}$, $B_l(\tau)$ are fixed basis functions, and $\theta_{kl}(\mathbf{s})$ are unknown coefficients that determine the shape of the quantile functions. Both $\beta_k(\mathbf{s})$ and $\theta_{kl}(\mathbf{s})$ are modeled as spatial processes, so the center and shape of the quantile functions vary spatially.

Before describing the proposed quantile regression model, we consider the important special case with $L = 1$ and $B_1(\tau) = \Phi^{-1}(\tau)$, where $\Phi^{-1}(\tau)$ is the standard normal quantile function. In this case

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

$$q(\tau|\mathbf{s},t)=[\beta_0(\mathbf{s})+\beta_1(\mathbf{s})t]+[\theta_{01}(\mathbf{s})+\theta_{11}(\mathbf{s})t]\Phi^{-1}(\tau), \quad (2)$$

and similar to Chandler (2005), the responses are Gaussian with spatially-varying mean $\beta_0(\mathbf{s}) + \beta_1(\mathbf{s})t$ and standard deviation $\theta_{01}(\mathbf{s}) + \theta_{11}(\mathbf{s})t$. In Section 2.1, we generalize the model for the quantile process by allowing $L > 1$, while maintaining this heteroskedastic Gaussian model as an important special case. This generalization allows us to characterize the entire quantile function with a flexible model, rather than modeling only the mean and standard deviation as functions of the covariates (here time) as in classical regression. Another consequence of this model for the quantile process is that it can accommodate non-Gaussianity, such as asymmetry and skewness, in $\mu_t(\mathbf{s})$ and thus $y_t(\mathbf{s})$.

2.1 Semiparametric model for the quantile process

The intuition for this model is that rather than specifying the mean and standard deviation as linear functions of time as in (2), we specify the quantile function for quantile levels $0 = \kappa_1 < \ldots < \kappa_{L+1} = 1$ as linear functions of time, i.e., $q(\kappa_l|\mathbf{s}, t) = q_0(\kappa_l|\mathbf{s}) + tq_1(\kappa_l|\mathbf{s})$, as in the classical analysis in Section 1. By allowing the coefficients q_1 to vary by κ, we characterize the evolution of different quantile levels, and thus different aspects of the response density. In contrast with the classical analysis in Section 1, these quantile levels are analyzed simultaneously with a hierarchical Bayesian model. This requires a model not only for the quantiles at the κ_j, but also a model for the quantile function for all $\tau \in [0, 1]$. To specify a complete density with the given quantile levels in a way that generalizes (2), we make use of piece-wise Gaussian basis functions to interpolate the quantile function at the κ_l to other quantile levels. By doing so, we obtain a semiparametric prior for the quantile function over all quantile levels. Other basis functions are possible, for example piecewise linear or spline basis functions. However, these piece-wise Gaussian basis functions have several attractive features including a natural centering on the heteroskedastic Gaussian model, a closed form for the density function, and a wide span of possible quantile functions.

If $L = 1$, then we define $B_1(\tau) = \Phi^{-1}(\tau)$ as in the heteroskedastic Gaussian model (2). To define the basis functions for $L > 1$, let $0 = \kappa_1 < \ldots < \kappa_{L+1} = 1$ be a grid of equally-spaced knots covering [0, 1]. Then for l with $\kappa_l < 0.5$,

$$B_l(\tau)=\begin{cases} \Phi^{-1}(\kappa_l)-\Phi^{-1}(\kappa_{l+1}) & \tau<\kappa_l \\ \Phi^{-1}(\tau)-\Phi^{-1}(\kappa_{l+1}) & \kappa_l \le \tau<\kappa_{l+1} \\ 0 & \kappa_{l+1} \le \tau, \end{cases} \quad (3)$$

and for l such that $\kappa_l \geq 0.5$,

$$B_l(\tau)=\begin{cases} 0 & \tau<\kappa_l \\ \Phi^{-1}(\tau)-\Phi^{-1}(\kappa_l) & \kappa_l \le \tau<\kappa_{l+1} \\ \Phi^{-1}(\kappa_{l+1})-\Phi^{-1}(\kappa_l) & \kappa_{l+1} \le \tau. \end{cases} \quad (4)$$

Figure 2a plots $B_l(\tau)$ for $L = 6$. For these basis functions, $q(\tau|\mathbf{s}, t)$ is continuous in τ for any $\theta_{kl}(\mathbf{s})$.

If $\theta_{01}(\mathbf{s}) = \ldots = \theta_{0L}(\mathbf{s})$ and $\theta_{11}(\mathbf{s}) = \ldots = \theta_{1L}(\mathbf{s})$, then $q(\tau|\mathbf{s}, t) = \beta_0(\mathbf{s}) + t\beta_1(\mathbf{s}) + [\theta_{01}(\mathbf{s}) + t\theta_{11}(\mathbf{s})]\Phi^{-1}(\tau)$, and $\mu_t(\mathbf{s})$ is Gaussian with mean $\beta_0(\mathbf{s}) + t\beta_1(\mathbf{s})$ and standard deviation $\theta_{01}(\mathbf{s}) + t\theta_{11}(\mathbf{s})$. This permits prior smoothing towards the heteroskedastic Gaussian spatial model. If $\theta_{11}(\mathbf{s}) \neq \ldots \neq \theta_{1L}(\mathbf{s})$, then the quantile function is no longer Gaussian. However, in this case if L is even then $B_l(0.5) = 0$ for all l, and thus the median of $\mu_t(\mathbf{s})$ is $q(0.5|\mathbf{s}, t) = \beta_0(\mathbf{s}) + t\beta_1(\mathbf{s})$, identifying $\beta_0(\mathbf{s})$ and $\beta_1(\mathbf{s})$ as the intercept and slope, respectively, of the median.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

To define a valid density function, the quantile function must be increasing in $\boldsymbol{\tau}$. Since $d\Phi^{-1}(\boldsymbol{\tau})/d\boldsymbol{\tau} > 0$ and the derivative of the quantile function is

$$\frac{dq(\tau|\mathbf{s},t)}{d\tau}=[\theta_{0j}(\mathbf{s})+t\theta_{1j}(\mathbf{s})]\frac{d\Phi^{-1}(\tau)}{d\tau}, \quad (5)$$

where $\boldsymbol{\kappa}_j \leq \boldsymbol{\tau} < \boldsymbol{\kappa}_{j+1}$, $q(\boldsymbol{\tau}|\mathbf{s}, t)$ is increasing in $\boldsymbol{\tau}$ if and only if $\theta_{0l}(\mathbf{s}) + t\theta_{1l}(\mathbf{s}) > 0$ for all l and t. This facilitates specifying a prior that yields an increasing quantile function, as described in detail below. We note again that if t was unbounded then $\theta_{1l}(\mathbf{s})$ would have to be zero to satisfy this constraint for all t, but with t bounded to [0,1], non-zero $\theta_{1l}(\mathbf{s})$ are possible.

This model spans a wide class of quantile functions. For example, if we restrict attention to the quantile function at quantile break points $\boldsymbol{\kappa}_2,\ldots,\boldsymbol{\kappa}_L$ and time points $t = 0, 1$, then Appendix A.2 shows that this model spans the entire set of valid quantile functions defined on this support. Therefore, for large L it should be possible to approximate a wide class of quantile functions.

To graphically illustrate the flexibility of the model, Figure 2b plots the quantile function $q(\tau)=\sum_{l=1}^{L} B_l(\tau)b_l$ for three choices of coefficients $\mathbf{b} = (b_1, \ldots b_L)$. The first quantile function (solid line) with all coefficients equal to 2 is the familiar Gaussian quantile function with mean zero and standard deviation 2. The second quantile function (thin dashed) is skewed with the final three coefficients larger than the first three coefficients. The third quantile function (thick dashed) is heavy-tailed with large first and last coefficients.

These basis functions also permit a closed-form expression for the density of $\mu_t(\mathbf{s})$. The quantile function (1) can be written

$$q(\tau|\mathbf{s},t)=\sum_{l=1}^{L}[a_l(\mathbf{s},t)+b_l(\mathbf{s},t)\Phi^{-1}(\tau)]I[\kappa_l \leq \tau<\kappa_{l+1}], \quad (6)$$

where $b_l(\mathbf{s}, t) = \theta_{0l}(\mathbf{s}) + t\theta_{1l}(\mathbf{s})$ and $a_l(\mathbf{s}, t) = q(\boldsymbol{\kappa}_{l+1}|\mathbf{s}, t) - b_l(\mathbf{s}, t)\Phi^{-1}(\boldsymbol{\kappa}_{l+1})$ if $\boldsymbol{\kappa}_l < 0.5$ and $a_l(\mathbf{s}, t) = q(\boldsymbol{\kappa}_l|\mathbf{s}, t) - b_l(\mathbf{s}, t)\Phi^{-1}(\boldsymbol{\kappa}_l)$ if $\boldsymbol{\kappa}_l \geq 0.5$. Recalling that the quantile function of a normal random variable with mean a and standard deviation b is $a + b\Phi^{-1}(\boldsymbol{\tau})$, the density becomes

$$p[\mu_t(\mathbf{s})]=\sum_{l=1}^{L}I[q(\kappa_l) \leq \mu_t(\mathbf{s})<q(\kappa_{l+1})]\mathrm{N}\left[\mu_t(\mathbf{s})|a_l(\mathbf{s},t),b_l(\mathbf{s},t)^2\right], \quad (7)$$

where $\mathrm{N}(\cdot|a, b^2)$ denotes the density of a normal with mean a and standard deviation b. We refer to this as a multiply-split normal density. Here again the need for the constraint that $b_l(\mathbf{s}, t) = \theta_{0l}(\mathbf{s}) + t\theta_{1l}(\mathbf{s}) > 0$ for all l and t is apparent. Figure 2c plots the density function for several examples. Unlike the quantile function, the density function may have discontinuities at the breakpoints $q(\boldsymbol{\kappa}_l|\mathbf{s}, t)$. Discontinuity in the density is a common artifact of Bayesian density estimation, for example, Dirichlet process (Ferguson, 1973, 1974), Polya tree (Lavine, 1992, 1994), and quantile pyramid (Hjort and Walker, 2009) priors. Although discontinuity in the density may not be ideal, our primary objective is to study the quantile function, which is continuous in all cases.

The quantile function is allowed to vary spatially by allowing both the intercepts β_k and basis function coefficients θ_{kl} to be spatial processes. The β_k are independent Gaussian processes with mean $\bar{\beta}_k$ and exponential covariance $\mathrm{Cov}[\beta_k(\mathbf{s}),\beta_k(\mathbf{s}')]=\tau_k^2\exp(-\|\mathbf{s}-\mathbf{s}'\|/\rho_k)$. The θ_{kl} are also modeled as spatial processes, however, these coefficients must satisfy the constraint that $\theta_{0l}(\mathbf{s}) + t\theta_{1l}(\mathbf{s}) > 0$ for all l and t. Since it is assumed that $t \in [0, 1]$, this

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

constraint is satisfied for all t if $\theta_{0l}(\mathbf{s}) > 0$ and $\theta_{0l}(\mathbf{s}) + \theta_{1l}(\mathbf{s}) > 0$. To satisfy the constraints, an unconstrained Gaussian processes θ^*_{kl} is introduced, and

$$\theta_{kl}(\mathbf{s}) = \begin{cases} \theta^*_{kl}(\mathbf{s}) & \theta^*_{0l}(\mathbf{s}) > \varepsilon \text{and} \theta^*_{0l}(\mathbf{s}) + \theta^*_{1l}(\mathbf{s}) > \varepsilon \\ \varepsilon & \text{otherwise} \end{cases} \quad (8)$$

for small constant $\varepsilon > 0$. In Section 4, $\varepsilon = 0.01$. The latent processes θ^*_{kl} are independent Gaussian processes with mean $\bar{\theta}_{kl}$ and exponential covariance $\mathrm{Cov}[\theta^*_{kl}(\mathbf{s}), \theta^*_{kl}(\mathbf{s}')] = \delta_k^2 \exp(-\|\mathbf{s} - \mathbf{s}'\|/\phi_k)$. The variances δ_k^2 control the amount of spatial variability in the quantile process. If $\delta_k^2 = 0$ then the shape of the quantile function is constant across space and determined by the spatial means $\bar{\theta}_{kl}$. The spatial means have hyperpriors $\bar{\theta}_{kl} \sim \mathrm{N}(\tilde{\theta}_k, \omega_k^2)$. Setting both the hypervariance $\omega_k^2 = 0$ and the spatial variances $\delta_k^2 = 0$ gives $\theta_{kl}(\mathbf{s}) = \tilde{\theta}_k$. As shown in Figure 2, constant coefficients gives the Gaussian model with mean $\beta_0(\mathbf{s}') + t\beta_1(\mathbf{s}')$ and standard deviation $\tilde{\theta}_0 + t\tilde{\theta}_1$. Therefore, the spatial quantile regression model is centered on the usual Gaussian spatial model with spatially-varying time trend.

2.2 Residual correlation model

Allowing the quantile function, and thus the density, to vary spatially accounts for some of the spatiotemporal association in the data. However, there may be additional residual correlation present after accounting for spatial variation in the density function. To account for this source of correlation while preserving the marginal quantile function $q(\boldsymbol{\tau}|\mathbf{s}, t)$, we implement a spatial copula (Nelson, 1999), which is often used to account for correlation in multivariate data with fixed marginal distributions. Specifying the marginal distribution via its quantile function is conducive to the copula approach, and the piece-wise Gaussian basis function in (3) and (4) leads to convenient computation. Let $\upsilon_t(\mathbf{s})$ be a latent Gaussian process, modeled using the dynamic spatial model

$$\upsilon_1(\mathbf{s}) = w_1(\mathbf{s}) \;\; \upsilon_t(\mathbf{s}) = \psi_\upsilon \upsilon_{t-1} + \sqrt{1 - \psi_\upsilon^2} w_t(\mathbf{s}), \text{for} t > 1, \quad (9)$$

where $|\psi_\upsilon| < 1$ and w_t are independent spatial processes with mean zero and covariance $\mathrm{Cov}[w_t(\mathbf{s}), w_t(\mathbf{s}')] = \exp(-\|\mathbf{s} - \mathbf{s}'\|/\psi_w)$. Then $\upsilon t(\mathbf{s})$ has the standard normal marginal distribution for all $\mathbf{s}$ and t, and $u_t(\mathbf{s}) = \Phi[\upsilon_t(\mathbf{s})] \sim \mathrm{Unif}(0,1)$. Using the probability integral transformation, $q(u_t(\mathbf{s})|\mathbf{s}, t)$ has quantile function $q(\boldsymbol{\tau}|\mathbf{s}, t)$, as desired. Therefore, we fit the model $\mu_t(\mathbf{s}) = q(u_t(\mathbf{s})|\mathbf{s}, t)$.

Substituting into (6), $\mu_t(\mathbf{s}) = \sum_{l=1}^{L} [a_l(\mathbf{s}, t) + b_l(\mathbf{s}, t)\upsilon_t(\mathbf{s})] I[\upsilon_t(\mathbf{s}) \in D_l]$, and thus

$$y_t(\mathbf{s}) \sim \mathrm{N}\left(\sum_{l=1}^{L} [a_l(\mathbf{s}, t) + b_l(\mathbf{s}, t)\upsilon_t(\mathbf{s})] I[\upsilon_t(\mathbf{s}) \in D_l], \sigma^2 \right), \quad (10)$$

where $D_l = [\kappa_{l-1}, \kappa_l)$. Therefore, as discussed further in Section 3, the full conditional distribution of $\upsilon_t(\mathbf{s})$ is a mixture of truncated normals, which permits Gibbs updates for these parameters.

The marginal density of $\mu_t(\mathbf{s})$ in (7) is discontinuous. However, $\mu_t(\mathbf{s})$ constructed using copula methods may be a smoothly-varying spatial process. Smoothness of spatial processes is often quantified using mean square continuity (Banerjee and Gelfand, 2003; Banerjee et al, 2003). The process $\mu_t(\mathbf{s})$ is mean square continuous at $\mathbf{s}$ if $\lim_{h \to 0} \mathrm{E}[\mu_t(\mathbf{s}) - \mu_t(\mathbf{s} + h\mathbf{u})]^2 = 0$, where $h \in \mathbb{R}$ and $\mathbf{u} \in \mathbb{R}^2$ is a unit vector. The latent process $\upsilon_t(\mathbf{s})$ relates to $\mu_t(\mathbf{s})$ via the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

continuous piecewise linear function $\mu_t(\mathbf{s})=\sum_{l=1}^{L}[a_l(\mathbf{s},t)+b_l(\mathbf{s},t)\upsilon_t(\mathbf{s})]I[\upsilon_t(\mathbf{s})\in D_l]$, which is Lipschitz continuous. Therefore, if the latent process υ_t is mean square continuous, then $\mu_t(\mathbf{s})$ is also mean square continuous.

3 Computational details

We perform MCMC sampling using R (http://www.r-project.org/), although it may be possible to use WinBUGS (http://www.mrc-bsu.cam.ac.uk/bugs/). Gibbs sampling is used for all parameters except for the spatiotemporal range parameters ρ, ϕ, ψ_υ, and ψ_w; these parameters are updated using Metropolis sampling with Gaussian candidate distribution tuned to give roughly 40% acceptance rate. The full conditionals for the latent spatiotemporal effects, $\upsilon_t(\mathbf{s})$, and parameters that control the shape of the quantile process, $\theta^*_{kl}(\mathbf{s})$, are given below. The remaining parameters have standard full conditionals from normal/normal or normal/gamma models. For the analysis of the climate data in Section 4, we generate 25,000 samples and discard the first 10,000 samples as burn-in. Convergence is monitored using trace plots and autocorrelation plots for several representative parameters. R code is available upon request.

The full conditional of the quantile curve coefficient $\theta^*_{kl}(\mathbf{s})$ is a mixture of two truncated normals. Denote $\theta^*_{kl}(\mathbf{s})$'s prior given all other values of $\theta^*_{kl}(\mathbf{s})$ as $\mathrm{N}(M_1, V_1)$. The mean and variance are functions of $\mathbf{P}$, the inverse covariance matrix of $\theta^*_{kl}=(\theta^*_{kl}(\mathbf{s}_1),\ldots,\theta^*_{kl}(\mathbf{s}_n))'$. V_1^{-1} is the diagonal element of $\mathbf{P}$ corresponding to location $\mathbf{s}$, and $M_1=V_1\mathbf{C}(-\mathbf{s})'\theta^*_{kl}(-\mathbf{s})$, where $\mathbf{C}$ is the column of $\mathbf{P}$ corresponding to $\mathbf{s}$, and $\mathbf{C}(-\mathbf{s})$ and $\Theta^*(-\mathbf{s})$ are the vectors $\mathbf{C}$ and Θ^* after removing the element corresponding to $\mathbf{s}$. The posterior is also a function of $x_t = B_l(u_t(\mathbf{s}))[I(k=0)+tI(k=1)]$, the residual not accounting for term l, $r_t(\mathbf{s}) = y_t(\mathbf{s})-\beta_0(\mathbf{s})-t\beta_1(\mathbf{s})-\Sigma_{j\neq l}B_j[u_t(\mathbf{s})][\Theta_{0j}(\mathbf{s})-t\Theta_{1j}(\mathbf{s})]-B_l(u_t(\mathbf{s}))(1+t)\varepsilon$, and residuals accounting for all terms except the term corresponding to Θ_{kl}, $R_t(\mathbf{s})=y_t(\mathbf{s})-\mu_t(\mathbf{s})+x_t\theta^*_{kl}(\mathbf{s})$. Twice the negative log of the full conditional for $\theta^*_{kl}(\mathbf{s})$ is the sum of a constant that does not depend on $\theta^*_{kl}(\mathbf{s})$ and

$$\begin{cases} \sum_t r_t(\mathbf{s})^2/\sigma^2+(\theta^*_{kl}(\mathbf{s})-M_1)^2/V_1 & \theta^*_{kl}(\mathbf{s})\le c \\ \sum_t[R_t(\mathbf{s})-x_t\theta^*_{kl}(\mathbf{s})]^2/\sigma^2+(\theta^*_{kl}(\mathbf{s})-M_1)^2/V_1 & \theta^*_{kl}(\mathbf{s})>c \end{cases} \quad (11)$$

where $c=\max\{\varepsilon,\varepsilon-\Theta_{1l}(\mathbf{s})\}$ if $k=0$ and $c=\varepsilon-\Theta_{0l}$ if $k=1$. Sampling proceeds by first determining if $\theta^*_{kl}(\mathbf{s})\le c$ or $\theta^*_{kl}(\mathbf{s})>c$. If $\theta^*_{kl}(\mathbf{s})\le c$, then $\theta^*_{kl}(\mathbf{s})\sim N_{[-\infty,c]}(M_1,V_1)$; otherwise $\theta^*_{kl}(\mathbf{s})\sim N_{[-\infty,c]}(M_2,V_2)$, where $V_2=1/(\Sigma_t x_t^2/\sigma^2+1/V_1)$, $M_2=V_2[\Sigma_t x_t R_t(\mathbf{s})/\sigma^2+M_1/V_1]$, and $N_A(a,b^2)$ is the truncated normal density with location a, scale b, and domain A. From (11), the probability that $\theta^*_{kl}(\mathbf{s})\le c$ is

$$\frac{\Phi\left(\frac{c-M_1}{\sqrt{V_1}}\right)}{\Phi\left(\frac{c-M_1}{\sqrt{V_1}}\right)+\sqrt{\frac{V_2}{V_1}}\left[1-\Phi\left(\frac{c-M_1}{\sqrt{V_1}}\right)\right]\exp[\sum_t r_t(\mathbf{s})^2/\sigma^2+M_1^2/V_1-M_2^2/V_2-\sum_t R_t(\mathbf{s})^2/\sigma^2]}. \quad (12)$$

Sampling from $\upsilon_t(\mathbf{s})$ is similar. Denote $\upsilon_t(\mathbf{s})$'s prior conditioned all other values of the latent process as $\upsilon_t(\mathbf{s})|\upsilon_{t'}(\mathbf{s}')\forall(t',\mathbf{s}')\neq(t,\mathbf{s})\sim\mathrm{N}(\bar{\upsilon},\sigma_\upsilon^2)$, $V_l^{-1}=1/[b_l(\mathbf{s},t)^2/\sigma^2+1/\sigma_\upsilon^2]$, and $M_l=V_l[b_l(\mathbf{s},t)(y_t(\mathbf{s})-a_l(\mathbf{s},t))/\sigma^2+\bar{\upsilon}\sigma_\upsilon^2]$. Let $g=j$ if $\upsilon_t(\mathbf{s})\in\mathscr{D}_j=[\Phi^{-1}(\kappa_j),\Phi^{-1}(\kappa_{j+1})]$. We sample $\upsilon_t(\mathbf{s})$ by first drawing g, and then $\upsilon_t(\mathbf{s})|g$. The full conditional probability of $g=j$ is proportional to

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

$$V_l^{1/2}\exp\left[\frac{M_l^2}{2V_l}-\frac{(y_t(\mathbf{s})-a_l(\mathbf{s},t))^2}{2\sigma^2}\right]\left[\Phi\left(\frac{\Phi^{-1}(\kappa_{j+1})-M_l}{\sqrt{V_l}}\right)-\Phi\left(\frac{\Phi^{-1}(\kappa_j)-M_l}{\sqrt{V_l}}\right)\right]. \quad (13)$$

Given g, $\upsilon_l(\mathbf{s}) \sim \mathrm{N}_{\mathscr{D}_g}(M_g, V_g)$.

4 Analysis of southeast US temperature data

4.1 Description of the data and exploratory analysis

To illustrate the proposed method, we analyze monthly climate data from the Florida, Georgia, North Carolina, South Carolina, and Virginia for the years 1931–2009. The data are downloaded from the National Climate Data Center at http://www7.ncdc.noaa.gov/CDO/cdo. We analyze separately three monthly responses: monthly mean temperature (F) and monthly extreme maximum and minimum temperature (F). We restrict analysis to observations from June, July, and August, and to locations missing less than 80% of the observations, leaving 188 sites for mean temperature, 191 sites for minimum temperature, and 190 sites for maximum temperature. The locations are plotted in Figure 3. To study more recent trends, we also consider an analysis of data from 1980–2009. Applying the same missing data threshold leaves 343, 359, and 358 sites for mean, minimum, and maximum temperature, respectively.

We begin presenting the analysis using standard frequentist quantile regression, implemented by the `rq` function in R, separately by spatial location and quantile level. Figure 3 plots the estimated linear time trend by quantile level for data in July. The estimates show considerable variation across both space and quantile level. For example, in the western Carolinas there is an increasing trend in minimum temperature for $\tau = 0.1$ (Figure 3c) but not $\tau = 0.9$ (Figure 3d). Similarly in northern Georgia there is a decreasing trend in maximum temperature for $\tau = 0.1$ (Figure 3e) but not $\tau = 0.9$ (Figure 3f). Also, several locations show a statistically significant time trend, but significance often differs between nearby sites. Temperature is a fairly smooth spatial process, so a spatial model that pools information across sites is desirable.

4.2 Priors and model-fitting

Section 2 outlines a general model for spatial quantile regression. This model is slightly modified for these temperature data. Rather than assuming a single quantile function for each month, we allow the quantile function to vary by month. Define $y_{tm}(\mathbf{s})$ as the observation in year t and month $m \in \{1, 2, 3\}$ for June, July, and August. Then $y_{tm}(\mathbf{s}) \sim \mathrm{N}(\mu_{tm}(\mathbf{s}), \sigma^2)$, where $\mu_{tm}(\mathbf{s})$'s quantile function is $q_{0m}(\tau|\mathbf{s}) + tq_{1m}(\tau|\mathbf{s})$ and $q_{km}(\tau|\mathbf{s})=\beta_{km}(\mathbf{s})+\sum_{l=1}^{L} B_l(\tau)\theta_{klm}(\mathbf{s})$. Since the quantile functions are likely to be similar, but not identical across month, we pool information across month by replacing the constant means $\mathrm{E}[\beta_{km}(\mathbf{s})] = \bar{\beta}_k$ and $\mathrm{E}[\theta_{klm}(\mathbf{s})] = \bar{\theta}_{kl}$ with spatial processes $\mathrm{E}[\beta_{km}(\mathbf{s})] = \bar{\beta}_k(\mathbf{s})$ and $\mathrm{E}[\theta_{klm}(\mathbf{s})] = \bar{\theta}_{kl}(\mathbf{s})$. The remaining parameters in the coefficient processes are the same, i.e., the variance and correlation parameters are shared across month. The average median processes are modeled as Gaussian processes with mean $\mathrm{E}[\bar{\beta}_k(\mathbf{s})] = \mathbf{X}(\mathbf{s})^T\gamma_k$ and covariance $\mathrm{Cov}[\bar{\beta}_k(\mathbf{s}),\bar{\beta}_k(\mathbf{s}')]=\tau\frac{2}{\beta}\exp(-\|\mathbf{s}-\mathbf{s}'\|/\rho_{\bar{\beta}})$, where $\mathbf{X}(\mathbf{s})$ are spatial covariates for the intercept, scaled latitude and longitude, and log elevation; the $\bar{\theta}_{kl}(\mathbf{s})$ are modeled similarly. Finally, since the analysis only includes three of twelve months, we assume the residual process follows the dynamic model in (9) within each year, and is independent across years.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

We use uninformative priors for most of the hyperparameters. The regression coefficients in the mean of $\bar{\beta}_k(\mathbf{s})$ and $\bar{\theta}_{kl}(\mathbf{s})$ have $N(0,10^2)$ priors, and all variances have InvGamma(0.1, 0.1) priors. The autocorrelation has flat prior $\psi_\upsilon \sim U(-1,1)$. The spatial range is notoriously difficult to estimate, and we therefore choose a mildly informative prior. To apply a stationary spatial model we first project the spatial locations to a two-dimensional surface using the Mercator projection, and then scale them to the unit square $[0, 1]^2$. Each of the log spatial ranges have priors $N(-2,1)$, which implies a (0.1, 1.5) prior 80% interval for the effective spatial range ρ log(1/0.05), i.e., the distance corresponding to correlation 0.05. This prior distributes mass for the effective range over short (10% of the spatial domain) and long ranges (150% of the spatial domain).

We fit the heteroskedastic Gaussian model with $L = 1$, as well as non-Gaussian models with $L = 4$ and $L = 10$ basis functions. These models are compared using several model comparison criteria, described below. To compare models, we randomly (across space and time) withhold 10% of the observations as a test set. For the test set, we compare mean squared error, $MSE=\frac{1}{N}\Sigma[y_{tm}(\mathbf{s})-\hat{y}_{tm}(\mathbf{s})]^2$, and mean absolute deviation $MAD=\frac{1}{N}\Sigma|y_{tm}(\mathbf{s})-\hat{y}_{tm}(\mathbf{s})|$, where the sum is over the N withheld observations and $\hat{y}_{tm}(\mathbf{s})$ is the posterior predictive mean. We also report the coverage probability of the predictive 90% intervals.

In addition to test-set criteria, the models are compared using several model-fit criteria for the training set. The deviance information criteria (DIC; Spiegelhalter et al., 2002) is

$$DIC=\bar{D}+p_D$$

where $\bar{D}$ is the posterior mean of the deviance, $pD=\bar{D}-\hat{D}$ is the effective number of parameters, and $\hat{D}$ is the deviance evaluated at the the posterior mean of the parameters in the likelihood, (10). The model's fit is measured by $\bar{D}$, while the model's complexity is captured by p_D. Gelfand and Ghosh (1998) propose the expected posterior deviation (EPD) criterion. Let $\tilde{y}_i^{(l)}$ be a replication of the ith observation drawn from (10) at iteration $l = 1, \ldots, L$ in the MCMC algorithm. EPD is defined as

$$EPD=G+\frac{a}{a+1}P,$$

where $G=\frac{1}{nL}\sum_{i=1}^{n}\sum_{l=1}^{L}(\tilde{y}_i^{(l)}-y_i)^2$ measures goodness of fit, $P=\frac{1}{nL}\sum_{i=1}^{n}\sum_{l=1}^{L}(\tilde{y}_i^{(l)}-\hat{y}_i)^2$ measures predictive variance, and $\hat{y}_i=\sum_{l=1}^{L}\tilde{y}_i^{(l)}/L$. We take $a = \infty$ to give equal weight to G and P. Gneiting et al. (2007) consider the continuous ranked probability score (CRPS). $CRPS$ is defined as

$$CRPS=E|y_i-\tilde{y}_i|+\frac{1}{2}E|\tilde{y}_i-\acute{y}_i|,$$

where $\tilde{y}_i$ and $\acute{y}_i$ are independent replicates from the model. CRPS can be approximated using MCMC output as $\frac{1}{nL}\sum_{i=1}^{n}\sum_{l=1}^{L}|y_i-\tilde{y}_i^{(l)}|+\frac{1}{2}|\tilde{y}_i^{(l)}-\acute{y}_i^{(l)}|$. Models with smaller DIC, EPD, and $CRPS$ are preferred.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

The results are in Table 1. For the 1931–2009 data, all five model comparison criteria are minimized using $L = 4$ basis functions for mean temperature. There is discrepancy between the criteria for minimum and maximum temperature. The cross-validation criteria favor $L = 1$ and $L = 10$ basis functions, respectively, for minimum and maximum temperature, while the remaining criteria favor $L = 10$ and $L = 1$ basis functions, respectively. The *DIC* criteria accounts for the entire fit of the model to the data, while the cross-validation criteria focus primarily on estimating the center of the distribution. Since the objective is to identify changes in the entire distribution over time, we present the results for number of basis functions minimized by *DIC*. For the 1980–2009 data, *DIC* favors $L = 4$ basis functions for all three responses.

4.3 Results

Figures 4–6 plot the posterior mean of the time trend $q_{1m}(\mathbf{s}|\boldsymbol{\tau})$ by quantile level $\boldsymbol{\tau}$ and month for the 1931–2009 data. Since the time variable is scaled to the unit interval, these estimates represent the change in temperature (F) quantiles from 1931 to 2009. For all three variables the quantile estimates have the same general pattern as in the classical quantile regression approach in Figure 3. For example, there is a negative trend for the lower quantile of the mean temperature distribution in most of the spatial region, but an increasing trend in southern Florida. However, using the Bayesian spatial model, the estimated time trends are much smoother across space and there are more sites identified as having a statistically significant time trend since information is pooled across spatial locations and months.

For states other than Florida, there is a slightly decreasing trend in mean temperature, especially for lower quantile levels. For example, the site in central North Carolina (longitude −79.3, latitude 35.8) in Figure 4a shows a statistically significant (represented by solid symbols) decreasing trend for lower quantile levels, and no significant trends for quantiles above the median.

The time trends for minimum temperature in Figure 5 display the most variability across space and quantile level. For $\boldsymbol{\tau} = 0.1$, there is an increasing trend in the northern part of the domain, and a decreasing trend in the south. This pattern varies significantly by quantile level. For example, the site in central Florida (longitude 81.5, latitude 28.1) in Figure 5a has a significant decreasing trend for lower quantiles, and a significantly increasing trend for upper quantiles for June and August. In contrast, the site in northern Georgia (longitude −83.8, latitude 34.8) in Figure 5a has a significant increasing trend for lower quantiles, and smaller effects for upper quantiles.

Unlike minimum temperature, the trends for maximum temperature in Figure 6 are relatively stable across space and quantile level. For most of the spatial domain there is a decreasing trend for all quantile levels. The decreasing trend is the most pronounced in northern Georgia and central North and South Carolina.

Figure 7 plots the posterior mean of the time trend for the 1980–2009 data for July. Again the time variable is scaled to be between zero and one, so that these estimates represent the change in temperature quantiles from 1980 to 2009. The slopes are similar to the full data analysis. Mean and maximum temperature show a decreasing trend for all states other than Florida. Estimates for minimum temperature display the most variability across quantile level and mean. There is an increasing trend in North Carolina and Virginia, especially for $\boldsymbol{\tau} = 0.1$. The trend in northern Georgia is positive for $\boldsymbol{\tau} = 0.1$ and negative for $\boldsymbol{\tau} = 0.9$.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

5 Discussion

This paper presents a new method for spatiotemporal quantile regression. The quantile function is allowed to vary with space and time to detect local climate changes. The form of the quantile function permits convenient Bayesian computation. For the southeast US data, there is a decreasing trend in mean and maximum temperature for most of the spatial domain and most quantile levels. The exception is Florida, which shows an increasing trend for mean temperature and no trend for maximum temperature. Minimum temperature displays more spatial variability and non-linear effects. For example, in Northern Georgia and the western Carolinas, the lower quantiles and median of minimum temperature show an increasing trend, but the trend in the upper quantiles are less pronounced. Also, in Florida the extremely low temperatures are becoming more extreme, while the median and upper quantiles show no change over time.

Several interesting modeling extensions are possible. For example, the quantile function could be a non-linear function of time, perhaps a polynomial function of time. This gives a more flexible model, but would require more complex constraints to ensure an increasing quantile function for each spatiotemporal location. Also, a multivariate extension which simultaneously analyzes several climate variables would provide a more powerful method for detecting climate change. This should be possible by using multivariate spatial priors for the parameters in the quantile functions and latent residual process.

References

Banerjee S, Gelfand AE. On smoothness properties of spatial processes. Journal of Multivariate Analysis. 2003; 84:85–100.

Banerjee S, Gelfand AE, Sirmans CF. Directional rates of change under spatial process models. Journal of the American Statistical Association. 2003; 98:946–954.

Banerjee, S.; Carlin, BP.; Gelfand, AE. Hierarchical modeling and analysis for spatial data. Chapman & Hall/CRC; 2004.

Bondell HD, Reich BJ, Wang H. Non-crossing quantile regression curve estimation. Biometrika. 2011; 97:825–838. [PubMed: 22822254]

Chandler RE. On the use of generalized linear model for interpreting climate variability. Environmetrics. 2005; 16:699–715.

Cressie, N. Statistics for spatial data. New York: Wiley-Interscience; 1993.

Dunson DB, Park JH. Kernel stick-breaking processes. Biometrika. 2008; 95:307–323. [PubMed: 18800173]

Easterling DR, Meehl GA, Parmesan C, Changnon SA, Karl TR, Mearns LO. Climate Extremes: Observations, Modeling, and Impacts. Science. 2000; 289:2068–2074. [PubMed: 11000103]

Folland, CK., et al. Observed climate variability and change. In: Houghton, JT., et al., editors. Climate Change 2001: The Scientific Basis. Cambridge, UK: Cambridge University Press; 2001. p. 99-181.

Ferguson TS. A Bayesian analysis of some nonparametric problems. The Annals of Statistics. 1973; 1:209–230.

Ferguson TS. Prior distribution on spaces of probability measures. The Annals of Statistics. 1974; 2:615–629.

Gelfand AE, Ghosh SK. Model choice: A minimum posterior predictive loss approach. Biometrika. 1998; 77:1–11.

Gelfand AE, Kottas A, MacEachern SN. Bayesian Nonparametric Spatial Modeling with Dirichlet Process Mixing. Journal of the American Statistical Association. 2005; 100:1021–1035.

Gelfand, AE.; Diggle, P.; Guttorp, P.; Fuentes, M. Handbook of spatial statistics. Chapman & Hall/CRC; 2010.

Gneiting T, Balabdaoui F, Raftery AE. Probabilistic forecasts, calibration and sharpness. Journal of the Royal Statistical Society: Series B. 2007; 69:243–268.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Griffin JE, Steel MFJ. Order-based dependent Dirichlet processes. Journal of the American Statistical Association. 2006; 101:179–194.

Hjort NL, Walker SG. Quantile Pyramids for Bayesian Nonparametrics. The Annals of Statistics. 2009; 37:105–131.

Katz RW, Brown BG. Extreme events in a changing climate: Variability is more important than averages. Climate Change. 1992; 31:289–302.

Koenker, R. Quantile Regression. Cambridge, U.K.: Cambridge University Press; 2005.

Kunkel KE, Liang XZ, Zhu J, Lin Y. Can CGCMs simulate the twentieth-century ”warming hole” in the central United States? Journal of Climate. 2006; 19:4137–4153.

Lavine M. Some aspects of Polya tree distributions for statistical modelling. The Annals of Statistics. 1992; 20:1222–1235.

Lavine M. More aspects of Polya tree distributions for statistical modelling. The Annals of Statistics. 1994; 22:1161–1176.

Lee D, Neocleous. Bayesian quantile regression for count data with application to environmental epidemiology. Journal of the Royal Statistical Society: Series C. 2010; 59:905–920.

24. Lum, K. Ph.D. thesis. Department of Statistical Science, Duke University; 2010. Bayesian spatial quantile regression.

Nelsen, R. An introduction to copulas. New York: Springer-Verlag; 1999.

Oh H-S, Lee TCM, Nychka DW. Fast nonparametric quantile regression with arbitrary smoothing methods. Journal of Computational and Graphical Statistics. 2011; 20:510–526.

Pan Z, Arritt RW, Takle ES, Gutowski WJ, Anderson CJ, Segal M. Altered hydrologic feedback in a warming climate introduces a “warming hole”. Geophysical Research Letters. 2004; 31:2–5.

Robinson WA, Reudy R, Hansen JE. General circulation model simulations of recent cooling in the east-central United States. Journal of Geophysical Research. 2002; 107:47–48.

Reich BJ, Fuentes M. A multivariate semiparametric Bayesian spatial modeling framework for hurricane surface wind fields. Annals of Applied Statistics. 2007; 1:249–264.

Reich BJ, Fuentes M, Dunson DB. Bayesian spatial quantile regression. Journal of the American Statistical Association. 2011; 106:6–20. [PubMed: 23459794]

Sobotka F, Kneib T. Geoadditive expectile regression. Computational Statistics and Data Analysis. 2011 In press.

Solomon, S.; Qin, D.; Manning, M.; Alley, RB.; Berntsen, T.; Bindoff, NL.; Chen, Z.; Chidthaisong, Z.; Gregory, JM.; Hegerl, GC.; Heimann, M.; Hewitson, B.; Hoskins, BJ.; Joos, F.; Jouzel, J.; Kattsov, V.; Lohmann, U.; Matsuno, T.; Molina, M.; Nicholls, N.; Overpeck, J.; Raga, G.; Ramaswamy, V.; Ren, J.; Rusticucci, M.; Somerville, R.; Stocker, TF.; Whetton, P.; Wood, RA.; Wratt, D. Technical Summary. In: Solomon, S.; Qin, D.; Manning, M.; Chen, Z.; Marquis, M.; Averyt, KB.; Tignor, M.; Miller, HL., editors. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press; 2007.

Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit (with discussion and rejoinder). Journal of the Royal Statistical Society: Series B. 2002; 64:583–639.

Tokdar ST, Kadane JB. Simultaneous Linear Quantile Regression: A Semiparametric Bayesian Approach. 2011 Submitted.

Yu K, Lu Z, Stander J. Quantile regression: Applications and current research areas. The Statistician. 2003; 52:331–350.

Appendix A.1 - Likelihood for the model without small-scale errors

Modeling the quantile process of the responses directly by setting $\varepsilon_t(\mathbf{s}) = 0$ requires a different computational algorithm. In this case, we derive an expression for the joint density of all observations over space and time as a function of β and θ, and then update β and θ

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

using Metropolis sampling. First define the cumulative distribution function corresponding to (7) as

$$F[y_t(\mathbf{s})]=\kappa_l+\Phi\left[\frac{y_t(\mathbf{s})-a_l(\mathbf{s},t)}{b_l(\mathbf{s},t)}\right]-\Phi\left[\frac{q(\kappa_l|\mathbf{s},t)-a_l(\mathbf{s},t)}{b_l(\mathbf{s},t)}\right]$$

where $\kappa_l < y_t(\mathbf{s}) < \kappa_{l+1}$ Then $\upsilon_t(\mathbf{s}) = \Phi^{-1}\{F[y_t(\mathbf{s})]\}$, and the joint density for all observations (over space and time), $\mathbf{y} = \{y_t(\mathbf{s}_i)|t, i\}$, is

$$p(\mathbf{y}|\beta,\theta)=p(\mathbf{v})\prod_{i,t}\frac{f[y_t(\mathbf{s}_i)|\beta,\theta]}{\phi[\upsilon_t(\mathbf{s}_i)]}$$

where $p(\mathbf{v})$ is the multivariate normal density corresponding to the spatial-temporal model for $\mathbf{v} = \{\upsilon_t(\mathbf{s}_i)|t, i\}$, f is given by (7), and ϕ is the standard normal density function. Computing $p(\mathbf{v})$ may be slow for large spatiotemporal datasets, but using a separable space-time correlation structure eases this computation.

Appendix A.2 - Span of the quantile function

Consider the quantile function at the breakpoints $\kappa_2, \ldots, \kappa_L$ and times $t = 0, 1$. Let $d_{kj} = q(\tau_j|t = k)$ and assume L is even. We show that any valid d_{kj} (i.e., with $d_{kj} > d_{kj-1}$ for all k and j) can be obtained within the constraints of the proposed model. Since β_0 and $\beta_0 + \beta_1$ define the median at $t = 0$ and $t = 1$, respectively, we may assume without loss of generality that $\beta_k = q(0.5|t = k) = 0$ for $k = 0, 1$, and thus $d_{kj} = 0$ if $\kappa_j = 0.5$ and $d_{kj} > 0$ if $\kappa_j > 0.5$ and vice versa. Also, since the model is symmetric in τ, we deal only with d_{kj} above the median, and results follow by symmetry for quantiles below the median.

For l so that $\kappa_l > 0.5$, the basis function (3) at the breakpoints $\kappa_2, \ldots, \kappa_L$ reduces to $B_l(\kappa_j) = c_l I(j > l)$, where $c_l = \Phi^{-1}(\kappa_{l+1}) - \Phi^{-1}(\kappa_l) > 0$. Therefore, the quantile functions at $t = 0$ and $t = 1$, omitting notational dependence on $(\mathbf{s})$, become

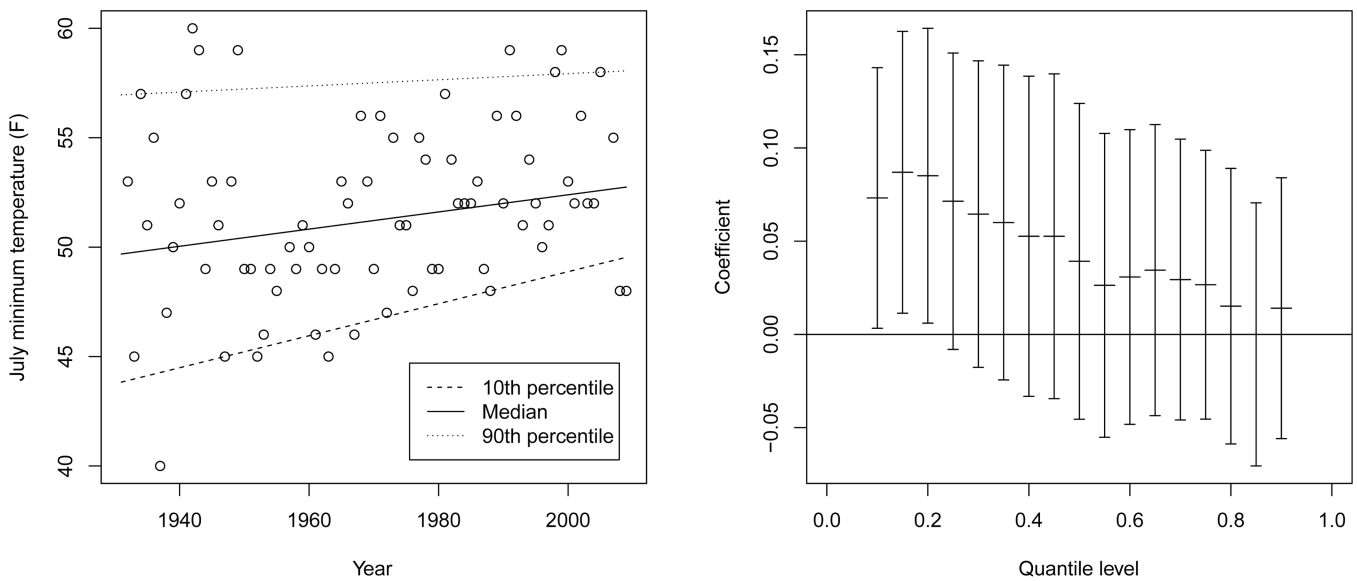
$$q(\tau_j|t=0)=\sum_{l=p-1}^{j-1}c_l\theta_{0l}\text{and}q(\tau_j|t=1)=\sum_{l=p-1}^{j-1}c_l(\theta_{0l}+\theta_{1l}),$$

where $p = \min\{j|\kappa_j > 0.5\}$. Then for $j = p$ we have $d_{0p} = c_{p-1}\theta_{0p-1}$ and $d_{1p} = c_{p-1}(\theta_{0p-1} + \theta_{1p-1})$. Thus $\theta_{0p-1} = d_{0p}/c_{p-1} > 0$ and $\theta_{1p-1} = (d_{1p} - d_{0p})/c_{p-1}$, which satisfies $\theta_{1p-1} + \theta_{0p-1} = d_{1p}/c_{p-1} > 0$. Next, for $j > p$, assume that θ_{kq} for $q < j - 1$ are fixed to match d_{kq} for $q < j$. Then $d_{0j} = d_{0j-1} + c_{0j-1}\theta_{0j-1}$ and $d_{1j} = d_{1j-1} + c_{j-1}(\theta_{0j-1} + \theta_{1j-1})$. Thus $\theta_{0j-1} = [d_{0j} - d_{0j-1}]/c_{j-1} > 0$ and $\theta_{1j-1} = ([d_{1j} - d_{1j-1}] - [d_{0j} - d_{0j-1}])/c_{j-1}$, which satisfies $\theta_{1j-1} + \theta_{0j-1} = [d_{1j} - d_{1j-1}]/c_{j-1} > 0$. Therefore, by recursion, θ_{jk} that satisfy model constraints are defined that match the quantiles at the breakpoints for $t = 0$ and $t = 1$.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

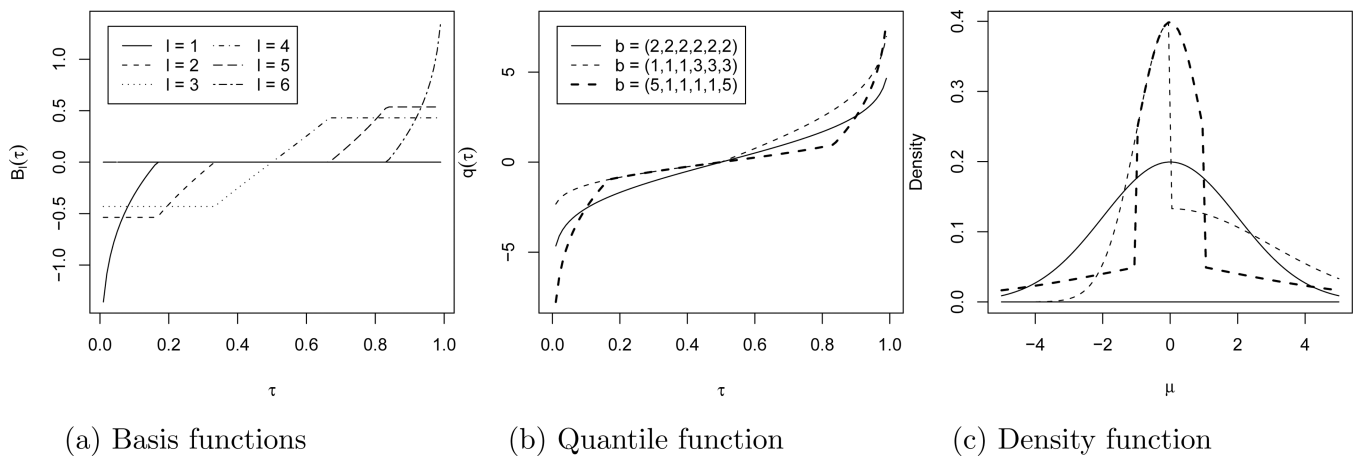
NIH-PA Author Manuscript



(a) Data and fitted quantile functions (b) Time trend estimate by quantile level

Figure 1.

July minimum temperature data by year for a site in northern Georgia with long/lat (−83.71, 34.85). Panel (a) shows the classical linear quantiles estimates $\hat{q}_0(\tau)+\hat{q}_1(\tau)t$, fit separately by quantile level and ignoring spatiotemporal dependence. Panel (b) plots the estimate of the time trend $\hat{q}_1(\tau)$ (±2*se) for various quantile levels.



(a) Basis functions (b) Quantile function (c) Density function

Figure 2.

Basis functions $B_l(\tau)$ for $L = 6$ (left), and the quantile function $q(\tau)=\sum_{l=1}^{L} B_l(\tau)b_l$ (middle) and corresponding density functions (right) for several coefficients $\mathbf{b} = (b_1, \ldots, b_6)$.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

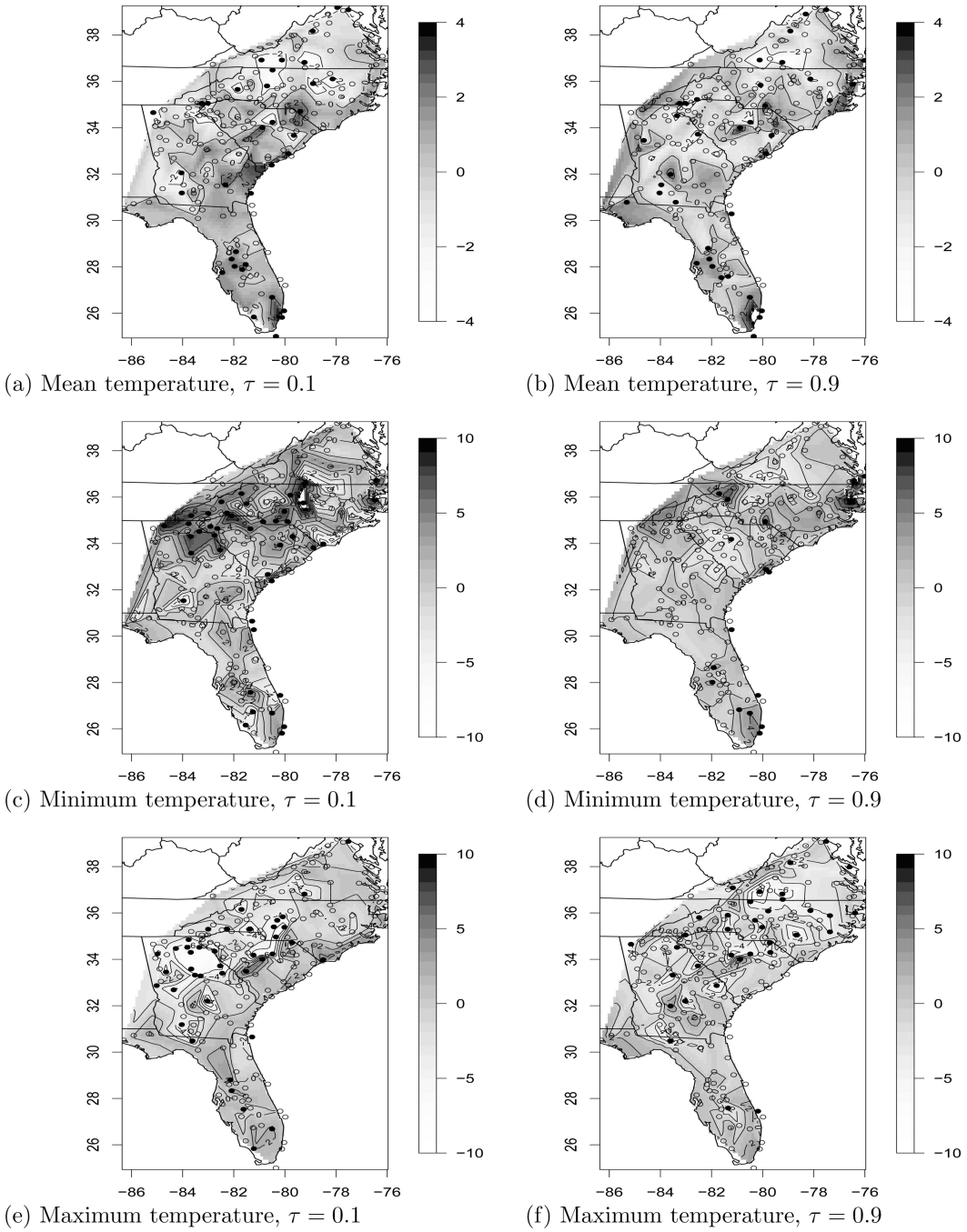


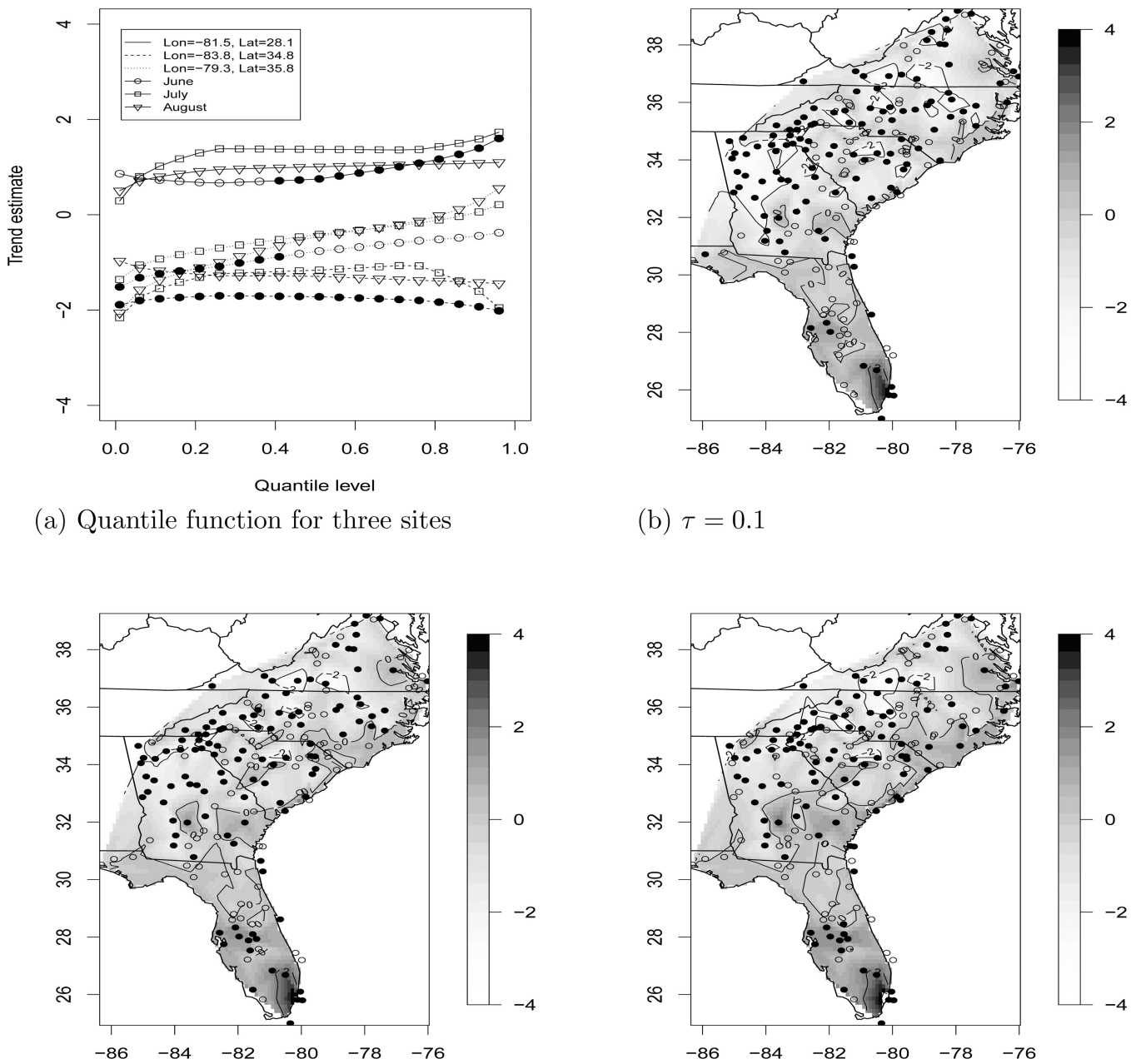
Figure 3.

Classical quantile regression estimates for the time trend by location and quantile level. The coefficients represent the change in temperature (F) quantiles from 1931 to 2009. Circles represent data locations, closed circles indicate a p-value less than 0.05 for the time trend, and the shading represents the estimated time trend coefficient.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript



(a) Quantile function for three sites

(b) $\tau = 0.1$

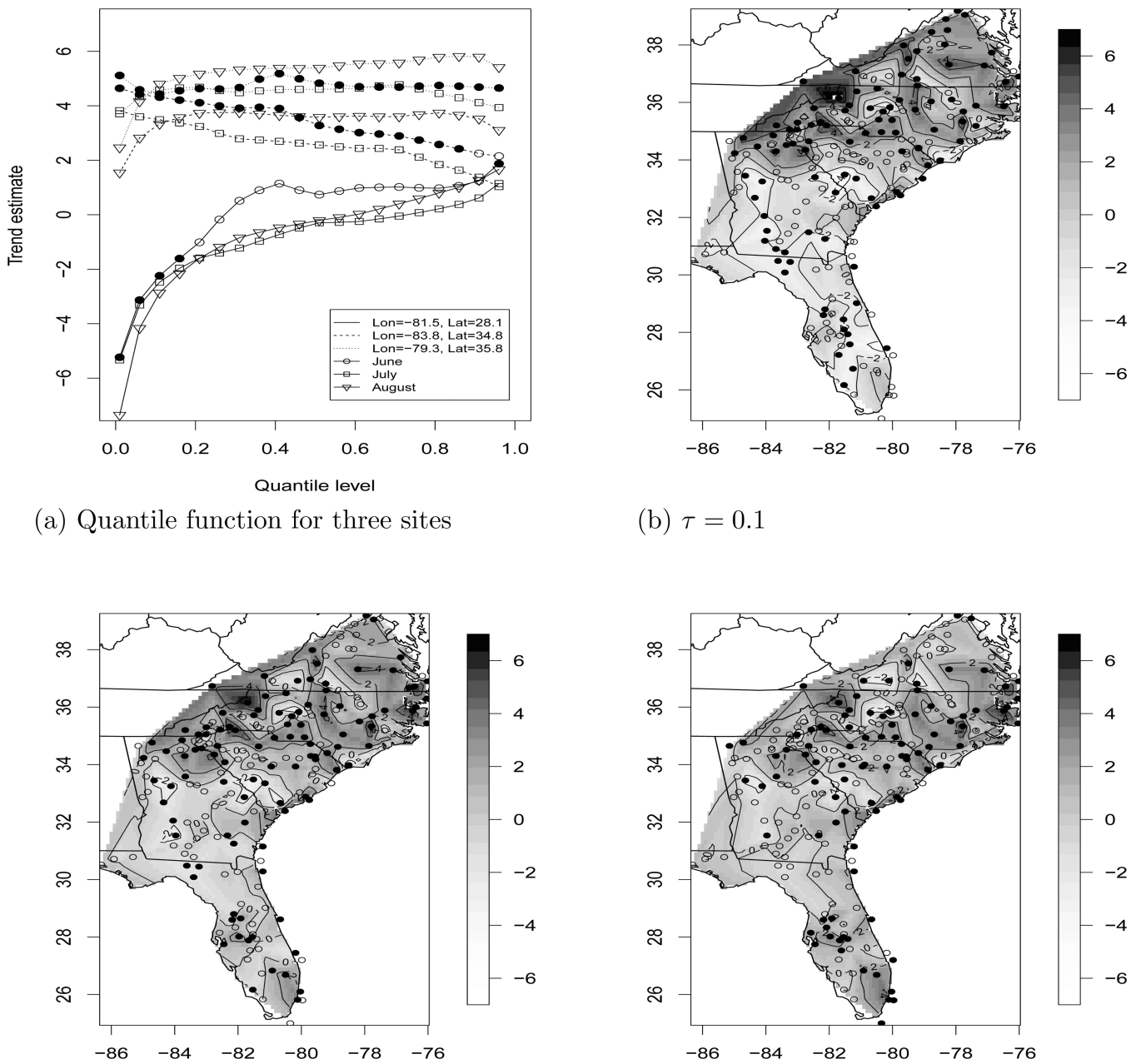
Figure 4.

Posterior mean trend coefficients $q_{1m}(\mathbf{s}|\boldsymbol{\tau})$ by quantile level $\boldsymbol{\tau}$ and month for 1931–2009 mean temperature (F). The coefficients represent the change in temperature (F) quantiles from 1931 to 2009. Panels (b)–(d) map the trend for July. Shaded symbols indicate the corresponding 95% interval excludes zero.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript



(a) Quantile function for three sites

(b) $\tau = 0.1$

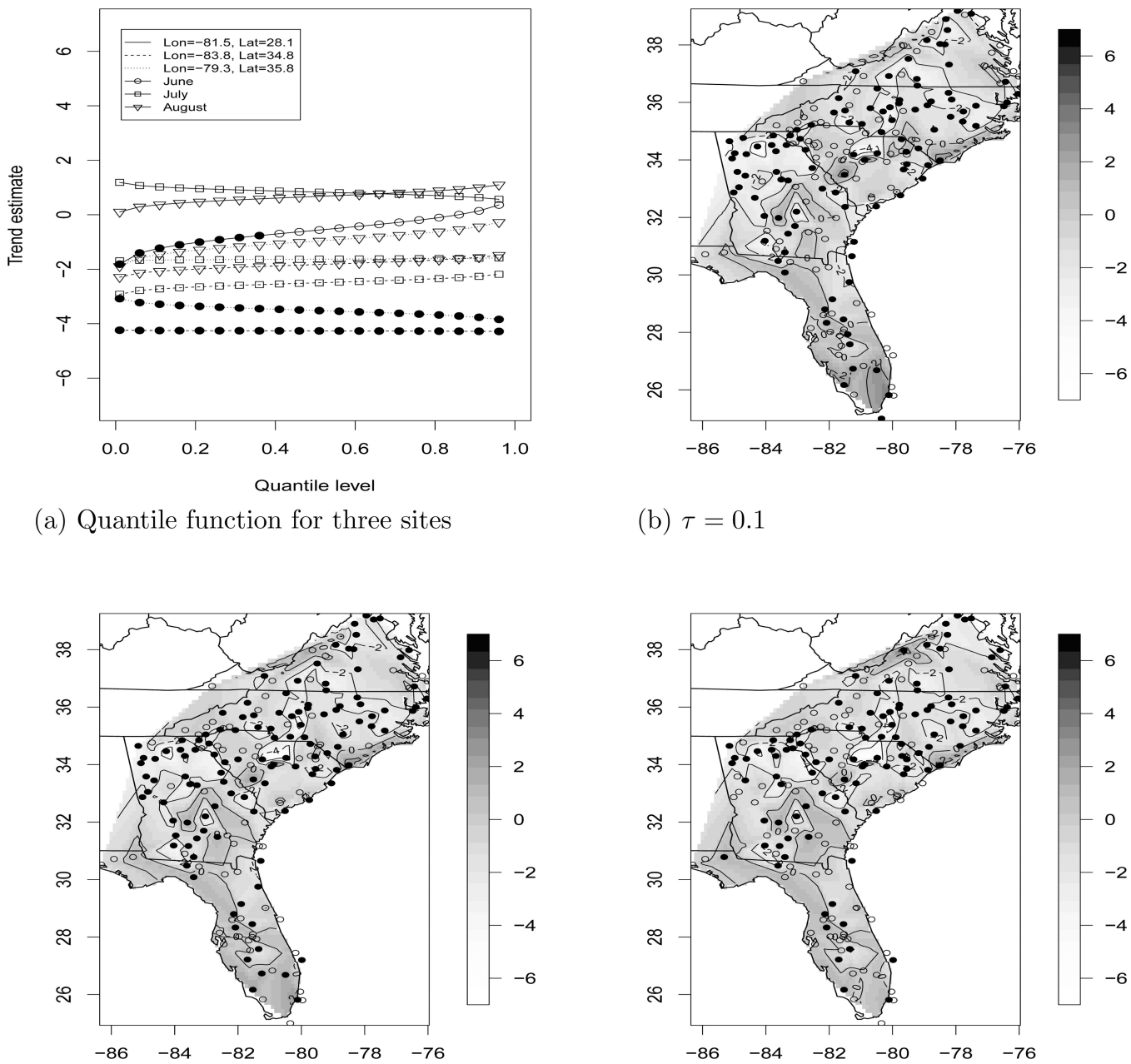
Figure 5.

Posterior mean trend coefficients $q_{1m}(\mathbf{s}|\boldsymbol{\tau})$ by quantile $\boldsymbol{\tau}$ level and month for 1931–2009 minimum temperature (F). The coefficients represent the change in temperature (F) quantiles from 1931 to 2009. Panels (b)–(d) map the trend for July. Shaded symbols indicate the corresponding 95% interval excludes zero.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript



(a) Quantile function for three sites

(b) $\tau = 0.1$

Figure 6.

Posterior mean trend coefficients $q_{1m}(\mathbf{s}|\boldsymbol{\tau})$ by quantile level $\boldsymbol{\tau}$ and month for 1931–2009 maximum temperature (F). The coefficients represent the change in temperature (F) quantiles from 1931 to 2009. Panels (b)–(d) map the trend for July. Shaded symbols indicate the corresponding 95% interval excludes zero.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

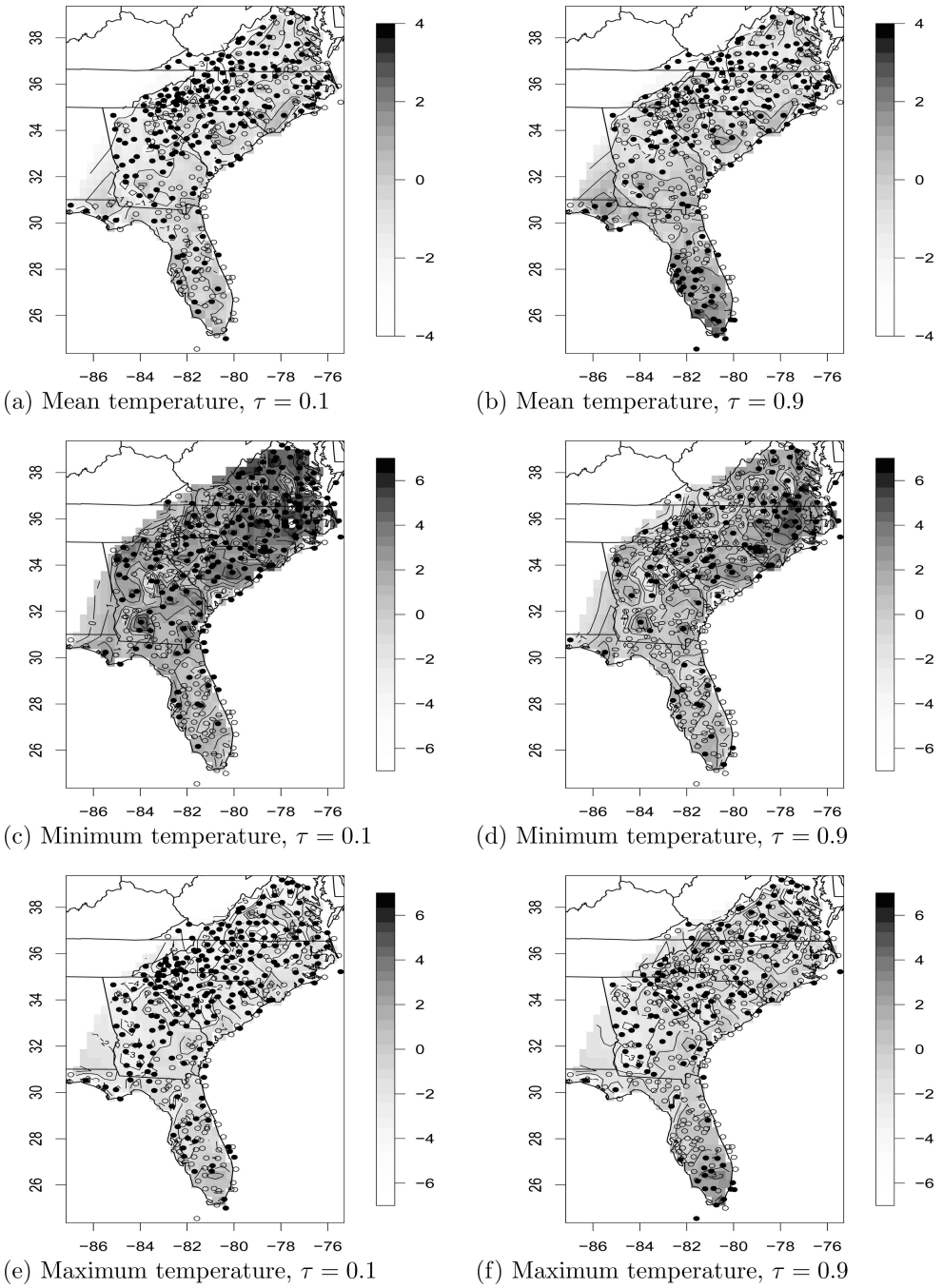


Figure 7.

Posterior mean trend coefficients $q_{1m}(\mathbf{s}|\boldsymbol{\tau})$ by quantile level $\boldsymbol{\tau}$ of July for 1980–2009 temperature (F). The coefficients represent the change in temperature (F) quantiles from 1980 to 2009. Shaded symbols indicate the corresponding 95% interval excludes zero.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

Model fit criteria for the number of basis functions, L.

Data	L	DIC	pD	$\bar{D}$	EPD	CRPS	MSE	MAD	COV90
(a) 1931–2009 data									
Mean temp	1	52,134	21,153	30,982	0.270	0.620	0.347	0.432	0.912
	4	**51,727**	22,187	29,540	**0.259**	**0.608**	**0.343**	**0.431**	0.916
	10	51,956	29,719	29,719	0.260	0.609	0.347	0.433	0.910
Min temp	1	155,779	12,905	142,873	5.162	2.706	**3.883**	**1.463**	0.912
	4	153,692	14,021	169,371	4.742	2.595	3.986	1.474	0.916
	10	**153,522**	14,862	138,660	**4.633**	**2.566**	4.038	1.502	0.907
Max temp	1	**132,352**	14,434	117,918	**2.687**	**1.955**	2.341	1.154	0.907
	4	132,717	14,502	118,215	2.704	1.960	2.479	1.177	0.909
	10	132,903	14,313	118,590	2.725	1.967	**2.193**	**1.132**	0.914
(b) 1980–2009 data									
Data	**L**	**DIC**	**pD**	**$\bar{D}$**	**EPD**	**CRPS**	**MSE**	**MAD**	**COV90**
Mean temp	1	40,080	13,847	26,232	0.314	0.668	**0.318**	**0.420**	0.923
	4	**38,245**	14,423	23,822	**0.285**	**0.637**	0.411	0.446	0.909
	10	38,304	14,442	23,862	**0.285**	**0.637**	0.372	0.434	0.900
Min temp	1	116,135	7,973	108,161	5.554	2.796	4.043	1.461	0.908
	4	**114,634**	8,860	105,774	**5.097**	**2.682**	4.098	1.447	0.912
	10	115,174	8,560	106,614	5.244	2.720	**3.936**	**1.435**	0.916
Max temp	1	97,047	9,844	87,203	2.659	1.944	**2.008**	**1.083**	0.914
	4	**96,388**	10,243	86145	2.549	1.904	2.128	1.090	0.903
	10	96,391	10,213	86,177	**2.535**	**1.899**	2.199	1.107	0.909