Time-Varying Models for Extreme Values

Gabriel Huerta

Department of Mathematics and Statistics, University of New Mexico, Albuquerque, U.S.A.

ghuerta@stat.unm.edu, www.stat.unm.edu/∼ghuerta

Bruno Sansó

Department of Applied Mathematics and Statistics, University of California, Santa Cruz, U.S.A.

bruno@ams.ucsc.edu, www.ams.ucsc.edu/∼bruno

Summary. We propose a new approach for modeling extreme values that are measured in time and space. First we assume that the observations follow a Generalized Extreme Value (GEV) distribution for which the location, scale or shape parameters define the space-time structure. The temporal component is defined through a Dynamic Linear Model (DLM) or state space representation that allows to estimate the trend or seasonality of the data in time. The spatial element is imposed through the evolution matrix of the DLM where we adopt a process convolution form. We show how to produce temporal and spatial estimates of our model via customized Markov Chain Monte Carlo (MCMC) simulation. We illustrate our methodology with extreme values of ozone levels produced daily in the metropolitan area of Mexico City and with rainfall extremes measured at the Caribbean coast of Venezuela.

Some key words: Spatio-temporal process, Extreme values, GEV distribution, Process convolutions, MCMC, ozone levels.

1. Introduction

Consider the time series presented in Figure 1 which corresponds to daily maxima of ozone concentrations (in ppm) over the area of Merced in Mexico City for the period 1990-2002. This plot suggests that the random fluctuations could be superimposed with a decreasing trend. The quantification of such a trend is key for the validation of important environmental policy decisions. The analysis of a time series of maximum values such as the one in Figure 1 would be naturally carried out by considering non-stationary processes with time varying parameters. This is the focus of this paper. Suppose that, in addition to the data at one specific location, we have data available at neighboring locations. The information from these data can be pooled together to obtain a global description of the behavior of maxima over a region. In this paper, we will show how the time-varying models developed for one location can be naturally extended to include a spatial component.

The traditional approach for the analysis of extreme values is based on the use of the Generalized Extreme Value (GEV) distribution. This is given by

$$H(z) = \exp\left\{ -\left[1 + \xi \left(\frac{z-\mu}{\sigma} \right) \right]_+^{-1/\xi} \right\} \tag{1}$$

where μ is a location parameter, σ is a scale parameter and ξ is a shape parameter. The $+$ sign denotes the positive part of the argument. This distribution encompasses the Frèchet ($\xi > 0$), Weibull ($\xi < 0$) and Gumbel ($\xi \to 0$) families. A clear account of the statistical modeling for extreme values using the GEV distribution is presented in Coles (2001). Under the assumption of independence, conditional on μ, σ and ξ, the likelihood of a sample of maxima, say $z_1, \ldots, z_m$ is readily available from (1). Thus, a Bayesian analysis can be carried out by imposing a prior distribution on μ, σ and ξ as in Coles and Tawn (1996).

An alternative approach to modeling extreme data is that of considering the distribution of exceedances over a high threshold. Pickands (1975) obtained the Generalized Pareto Distribution, whilst Pickands (1971) introduced a point process approach. These ideas have been extended and applied in Smith et al. (1997), Smith (1989), Davison and Smith (1990) and more recently Assunção et al. (2004). Gaetan and Grigoletto (2004) present a model with dynamically varying parameters for the analysis of extremes of a univariate

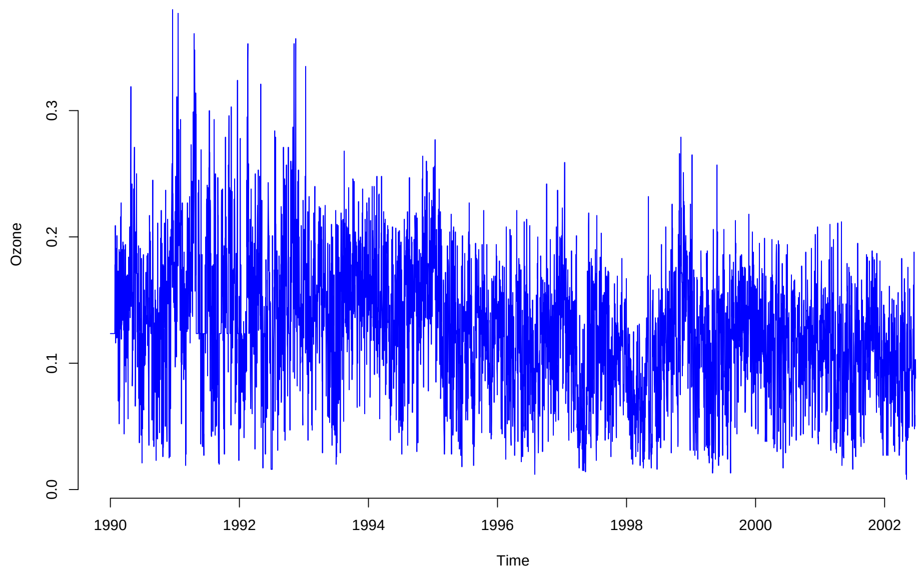

Fig. 1. Daily maxima of ozone concentration at station Merced in Mexico City from 1990 to 2002.

time series. Their model includes dynamic changes for scale/shape parameters and sequential updating with *Particle Filters*. There is no space or space-time structure. Regarding the modeling of data at different locations in space, Casson and Coles (1999) consider the spatial distribution of extremes using a threshold model based on the point process approach. Gilleland et al. (2004) develop a hierarchical model based on the Generalized Pareto Distribution. The present paper shares with the last two papers cited the fact that the modeling is based on hierarchical structures that assume the existence of an underlying spatial model. An alternative would be to consider models based on multivariate distributions for extremes. Families of such distributions have been developed, in Ledford and Tawn (1996, 1997) and Heffernan and Tawn (2004) by focusing on obtaining flexible asymptotic dependence properties. Dupuis (2005) considers an approach based on copulas. So far, models based on multivariate distributions for extremes have been applied only to a few spatial locations.

In this paper we focus on the use of the GEV distribution and allow the parameters to vary in time and space. In Section 2 we propose some approaches for time series of extremes taken at one location in space. We study the time varying behavior of the parameters that govern the distribution of the data using Dynamic Linear Models as presented in West and Harrison (1997). We illustrate the possibilities of capturing time varying trends. The use of a dynamic model allows for the detection of both short and long range changes. We also consider the problem of relating trends to time varying covariates. By doing so we are able to describe how the strength of the linear relationship changes with time. In section 3 we extend the model to allow for spatial variation of the parameters. Our spatio-temporal model for extremes is based on process convolutions using methods presented, for example, in Higdon (2002). This provides a reduction of the parameter space that is effective in handling large number of locations. In the final section we present a discussion of our results.

2. GEV distribution with parameters varying in time

Coles (2001) presents an approach to model changes across time in maxima using the GEV distribution. These changes are defined through deterministic functions and may refer to trend or seasonality in the data.

Consider a sequence of conditionally independent observations $z_1, z_2, \ldots, z_m$ such that each observation follows a GEV distribution with a time varying location parameter, i.e., $z_t \sim GEV(\mu_t, \sigma, \xi)$. To obtain z_i we fix a time length or block of data and calculate the maximum of the observations obtained during the i-th period of time or block. To model changes in time, Coles (2001) mentions several possibilities. For example, $\mu_t = \beta_0 + \beta_1 t$, $\mu_t = \beta_0 + \beta_1 + \beta_2 t + \beta_3 t^2$ or some other higher-order polynomial on t. Also, he suggests the use of covariates with an expression of the form $\mu_t = \beta_0 + \beta_1 X_t$. Non-stationary models can also be formulated for the shape and/or scale parameters. For example, $\sigma_t = \exp(\beta_0 + \beta_1 t)$; $\xi_t = \beta_0 + \beta_1 t$ or $\xi_t = \beta_0 + \beta_1 t + \beta_2 t^2$.

Alternatively, we propose the use of Dynamic Linear Models (DLMs), a very general class of time series models, as in West and Harrison (1997) to model parameter changes across time for location, scale or shape. For a time-varying location model, we assume that the data $z_1, z_2, \ldots, z_m$ defines a conditionally independent sequence and $z_t \sim GEV(\mu_t, \sigma, \xi)$ so the cumulative distribution function of each z_t is

$$H_t(z_t) \quad = \quad \exp\left\{ - \left[1 + \xi \left(\frac{z_t - \mu_t}{\sigma} \right) \right]_+^{-1/\xi} \right\}; \quad t = 1, \ldots, m.$$

This assumption is based on asymptotic results and so it can be considered as an approximation. Instead of a deterministic function on μ_t, we assume that

$$\begin{aligned} \mu_t &= F_t^{'} \theta_t + \epsilon_t; \quad \epsilon_t \sim N(0, V) \\ \theta_t &= G_t \theta_{t-1} + \omega_t; \quad \omega_t \sim N(0, W) \end{aligned}$$

where θ_t is state vector of dimensions $k \times 1$; F_t is a "regressor" of dimensions $k \times 1$; G_t is a $k \times k$ evolution matrix; V is the observational variance and W is a $k \times k$ evolution covariance matrix. This defines a DLM (F_t, V, G_t, W). By using a DLM we obtain more flexibility in that trends are not constrained to having a specific parametric form. We can assess the significance of short term changes together with long term ones. Also, it is possible to quantify how the effects due to covariates change with time.

Posterior inference and structure under our new GEV distribution follows standard Markov chain Monte Carlo (MCMC) methods (see, for example, Gamerman, 1997). We briefly describe the relevant full conditional distributions. Define the vectors $\boldsymbol{Z} = (z_1, z_2, \ldots, z_m)$; $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_m)$ and $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)$. At each step of our MCMC algorithm, we draw,

- Each value of μ_t from $p(\mu_t | z_t, \theta_t, V, W); t = 1, \ldots, m$ via Metropolis-Hastings steps. Given the conditional independence structure of the model, the conditioning on $\boldsymbol{Z}$ reduces to z_t.
- The scale and shape parameters are simulated from $p(\sigma | \boldsymbol{Z}, \boldsymbol{\mu}, \xi)$ and $p(\xi | \boldsymbol{Z}, \boldsymbol{\mu}, \sigma)$ with individual Metropolis-Hastings steps. The scale and shape parameters of the GEV distribution are assumed constant in time.
- The observational variance V can be sampled from an Inverse Gamma distribution.
- The evolution covariance matrix W can be modeled with discount factors or as functions of the observational variance V. In the second case, the evolution can be sampled with an Inverse-Gamma or Inverse-Wishart distribution.
- Conditional on $\boldsymbol{\mu}$, V and W, we sample the state vector $\boldsymbol{\theta}$ using the *Forward Filtering Backward Simulation* (FFBS) algorithm as presented in Carter and Kohn (1994) or Frühwirth-Schnatter (1994). Forward in time, we apply the recursive filtering equations (West and Harrison, 1997, Chapter 4), to obtain $p(\theta_t | D_t, V, W); t = 1, 2, \ldots, m$, the posterior distribution of θ_t given all the information available up to time t (D_t). Backwards in time, we first simulate $p(\theta_m | D_m, V, W)$ and then, recursively we sample the other θ_t values from $p(\theta_t | \theta_{t+1}, D_m, V, W)$ for $t = m - 1, \ldots, 1$.

An implementation of this algorithm, together with the data analyzed in this paper are available from the URL `www.stat.unm.edu/~ghuerta/`.

As an illustration of the proposed model we return to the data in Figure 1. To explore the existence of possible trends we consider a particular case of time-varying GEV. This is given by a first order polynomial

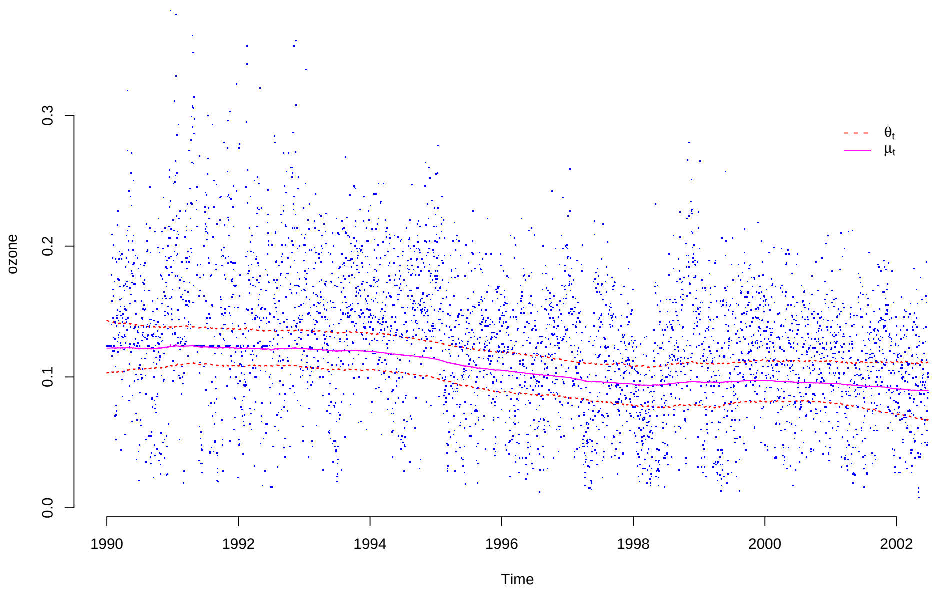

Fig. 2. Daily maxima of ozone concentration at Merced from 1990 to 2002, posterior means for (μ_t, θ_t) and 90% posterior probability interval for θ_t

DLM where $F_t = G_t = 1, \forall t$, V is unknown and $W_t = \tau V, \forall t$, $\tau > 0$, with a one dimensional state-parameter θ_t. Such model is useful to represent changes in level of the location parameter μ_t. We complete the model by considering the priors: $(\theta_0|D_0) \sim N(m_0, C_0), V \sim IG(\alpha_v, \beta_v), \tau \sim IG(\alpha_\tau, \beta_\tau)$. Additionally, we assume that (σ, ξ) are a priori independent where $\log(\sigma)$ and ξ follow a Normal distribution, i.e. $\pi(\sigma) \sim LN(m_\sigma, s_\sigma)$ and $\pi(\xi) \sim N(m_\xi, s_\xi)$.

In Figure 2 we present the results of fitting this GEV- first order polynomial DLM. The figure shows the data (points), posterior means for μ_t and θ_t and the 90% posterior probability interval for θ_t. The decrease in levels across time is quite notorious. These estimates illustrate that our models are very flexible in representing dynamic structure for extreme values.

As we have previously mentioned, the use of the distribution in (1) for extremes is based on asymptotics. That is, z is assumed to be the maximum of n independent random variables and, after normalization, when n tends to infinity, the GEV is obtained. In practice we have neither independence nor a very large number of observations. To check the goodness of fit of the GEV model, the usual approach is to observe the ability of the model to predict the empirical quantiles of the data. Our model produces time-varying quantiles, so, there is only one observation available to obtain the empirical quantile at any given time. To overcome this difficulty we consider two diagnostics. The first one is presented in Figure 3. We observe that most of the data are contained within the (time-varying) intervals whose lower and upper bounds are the 5% and 95% quantiles respectively.

An alternative diagnostic is the one based on the method proposed in Kim et al. (1998). Let $u_t = H_t(z_t), \forall t$ then, according to Rosenblatt (1952), if the model is properly specified then the sequence u_t should be independent and distributed as a uniform in (0,1). Clearly, for our model we have a whole sample of u_t for each t. Kim et al. (1998) propose to do a qq-plot of $\Phi^{-1}(2|u_t - 0.5|)$, where Φ is the cumulative distribution function of a standard normal. The results from the analysis of the u_t are presented in Figure 4. We observe that, according to this diagnostic, the model is providing a fairly good fit of the data. In regards to the tail behavior of Mexico City ozone extremes, a 95% posterior interval for the shape parameter ξ is $(-0.189, -0.146)$, which corresponds to a Weibull family of distributions.

One of the most important questions about ozone daily maxima is to determine if there is a decreasing

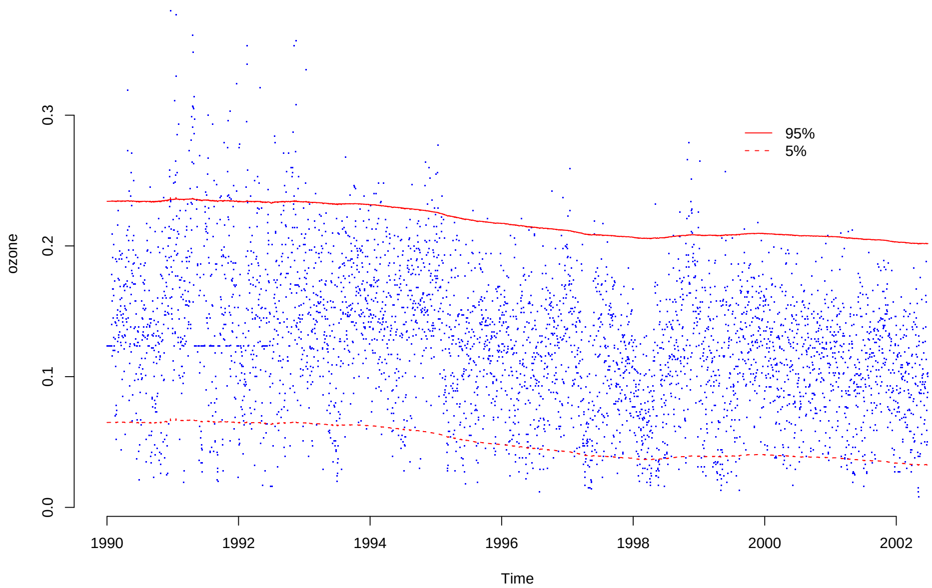

Fig. 3. Daily maxima of ozone concentrations. 95% and 5% quantiles for time-varying model.

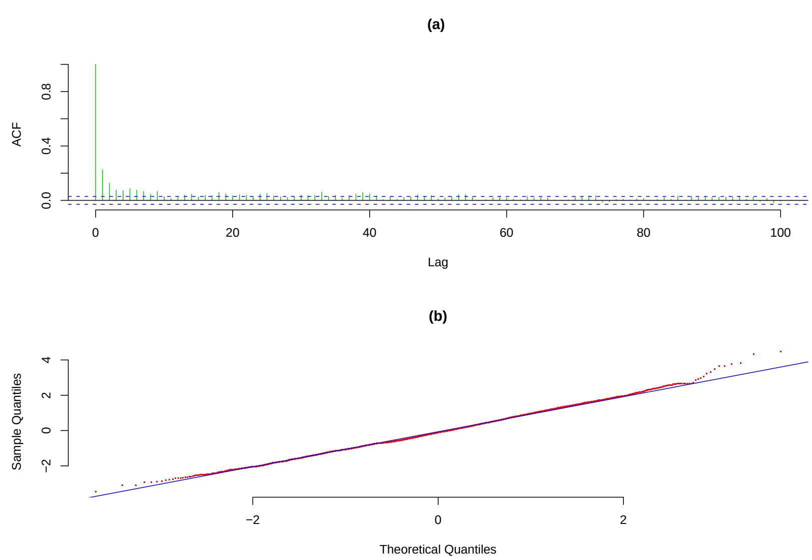

Fig. 4. Diagnostic for the time-varying model fitted to the data observed at Merced between 1990 and 2002. (a) Autocorrelation function of $2|u_t - 0.5|$ (b) qq-plot of $\Phi^{-1}(2|u_t - 0.5|)$.

trend. For this purpose, we considered a *second order polynomial* (West and Harrison, 1997) DLM for μ_t. This second order polynomial is defined by the equations,

$$\begin{aligned}\mu_t &= \delta_t + \epsilon_t \\ \delta_t &= \delta_{t-1} + \beta_{t-1} + \nu_t \\ \beta_t &= \beta_{t-1} + \eta_t\end{aligned}$$

with the typical normal distribution assumptions for the error terms. In DLM notation this model is written $F_t^{'} = (1, 0)$, $\theta_t^{'} = (\delta_t, \beta_t)$,

$$G_t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

$\epsilon_t \sim N(0, V)$ and $\omega_t^{'} = (\nu_t, \eta_t) \sim N(0, W_t)$. δ_t is the level and β_{t-1} represents incremental growth for the series μ_t. We considered a discount factor approach to model the evolution variance W_t. A discount factor measures the amount of information loss, between time $t-1$ and time t, for the state parameters of a DLM (West and Harrison, 1997, Chp. 4). For the daily maxima of ozone and for a discount factor of 0.95, we obtained a posterior estimate for δ_t that resembles the level estimated with the first order polynomial model. However, the estimated incremental growth is practically constant in time with a value close to 0. This implies that there is little evidence of significant short changes in level.

To assess the significance of a long range trend, we considered a regression model for μ_t, with a time varying intercept but with a constant slope

$$\begin{aligned}\mu_t &= \theta_t + \beta(t - \bar{t}) + \epsilon_t \\ \theta_t &= \theta_{t-1} + \omega_t\end{aligned}$$

where $\epsilon_t \sim N(0, V)$, $\omega_t \sim N(0, \tau V), \tau > 0$, ϵ_t and ω_t are assumed independent and $\bar{t} = (1/T)(\sum_{t=1}^{T} t)$ so β represents the change level per unit of time. This new model on μ_t is not a DLM so the FFBS does not apply directly to obtain posterior inferences on β and θ_t. However, notice that conditional on θ_t, $\mu_t - \theta_t$ defines a standard regression model with constant slope and zero intercept. For a non-informative prior, $p(\beta) \propto 1$, the full conditional for β follows a normal distribution. Furthermore, conditional on β, $\mu_t - \beta(t - \bar{t})$ follows a first order polynomial DLM and at this conditional step, FFBS applies to sample θ_t.

In the case of the daily maxima of ozone, Figure 5 shows the posterior distribution of the parameter β. It is clear that there is plenty of posterior mass below zero. In fact, the posterior probability that $\beta < 0$ is 0.7902 and so there is some indication that daily maxima ozone levels have decreased.

As an additional example, we consider maxima monthly rainfall values from January 1961 to November 1999 taken at the Maiquetía station located at the Simón Bolívar Airport near Caracas, Venezuela. Coles et al. (2003) present an analysis of these data, which are very relevant due to the catastrophic events that occurred in the northern coast of Venezuela in December 1999. The rainfall values are shown on Figure 6 (a). It has been suggested that the Monthly North Atlantic Oscillation (NAO) Index can help explain rainfall precipitations across the Atlantic Ocean and could also explain rainfall phenomena nearby the Venezuelan Atlantic coast (personal communication from Lelys Guenni). Based on this we consider the NAO index as a covariate. The NAO index is the difference of normalized sea level air pressures between Ponta Delgada, Azores and Reykjavik, Iceland. More information on the NAO index is available from the URL `tao.atmos.washington.edu/data_sets/nao/`. A time series plot of the NAO index is shown in Figure 6 (b). From a visual comparison of Figure 6 (a)-(b), there is no obvious relationship between the NAO index values and maxima rainfall values at Maiquetía airport.

We considered a model where $Y_t \sim GEV(\mu_t, \sigma, \xi)$, Y_t is the rainfall value at Maiquetía station at time t and μ_t follows a dynamic regression with the monthly NAO index X_t, as the only covariate:

$$\begin{aligned}\mu_t &= \theta_t + \beta_t X_t + \epsilon_t \\ \theta_t &= \theta_{t-1} + \omega_{1t} \\ \beta_t &= \beta_{t-1} + \omega_{2t}.\end{aligned}$$

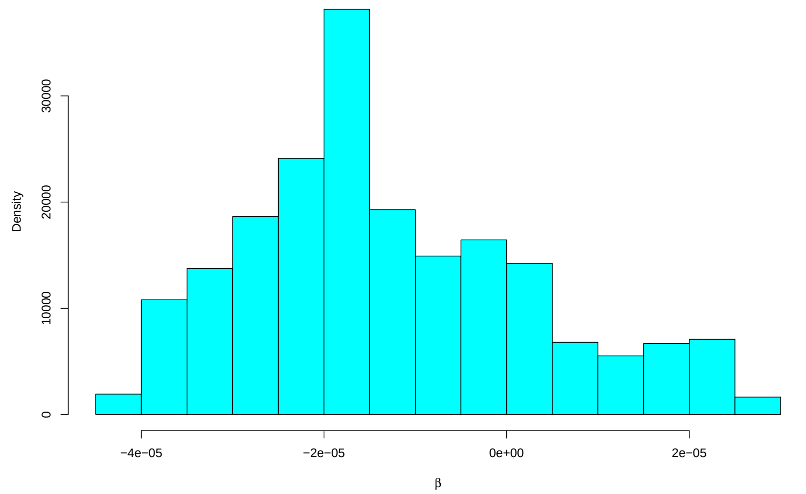

Fig. 5. Posterior distribution of β for ozone data.

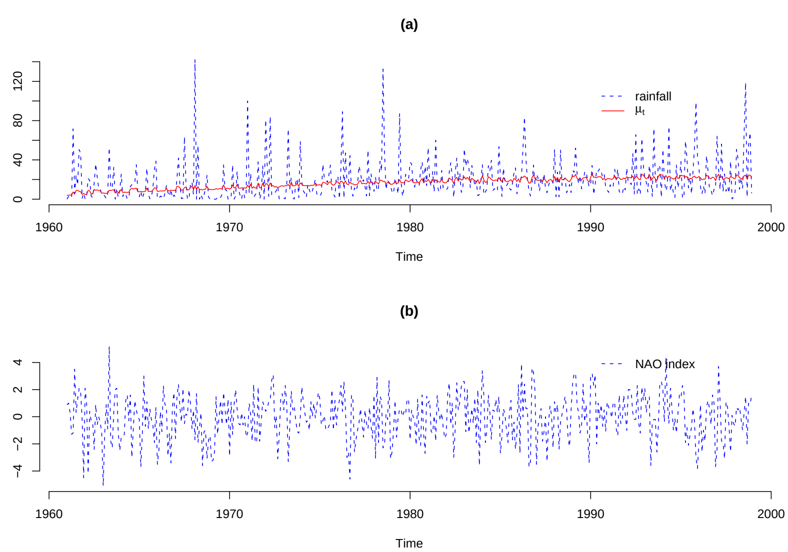

Fig. 6. (a) Maiquetía monthly observations from 1961 and posterior mean of μ_t. (b) Monthly values of NAO index from 1961.

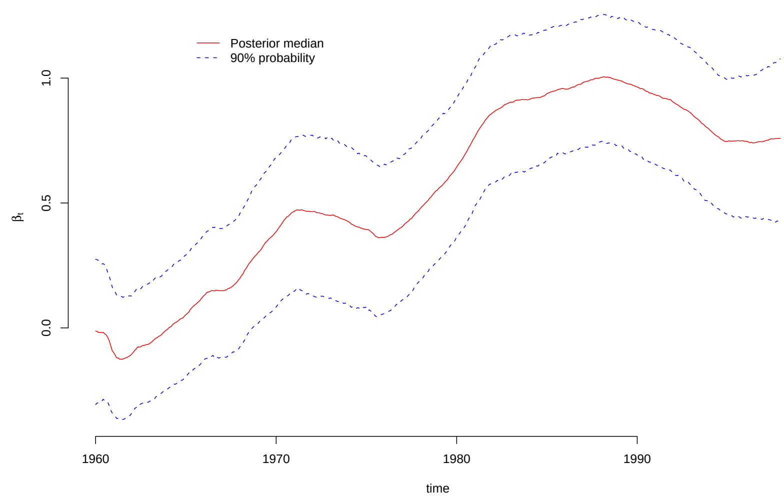

Fig. 7. Posterior median of dynamic regression coefficient β_t and 90% probability bands

The usual assumptions for ϵ_t and $\omega_{it}; i = 1, 2$ are made and the evolution variances are defined via a unique discount factor. For a discount factor of 0.98, Figure 6 (a) shows the posterior mean of μ_t over-imposed to the rainfall data. We observe an increase that is compatible with the fact that a number of large maxima are clustered in the last decade. Rainfall in Northern Venezuela is a very seasonal phenomenon, with a dry and a wet season that are usually very distinguishable. We considered the need of adding a seasonal component to the model but we found no periodical fluctuations in the series of μ_t to support this claim. Also, we calculated the differences between the median of z_t and the corresponding observation and we did not observe any periodicities. Figure 7 shows the posterior median of β_t and 90% probability intervals for each time point. We observe that after 1970 the coefficient is significantly positive. This indicates that there is evidence of a significant linear association between the dynamics of the monthly maxima and the behavior of the NAO index. The strength of such association changes with time. The posterior distribution of the shape parameter ξ (not shown) is mostly concentrated between .3 and .7. Very few samples fell outside the interval (0,1). This is an indication that the distribution of the maxima is Frèchet and that it is unbounded above.

3. GEV distribution with parameters varying in space and time

We can very naturally extend the model presented in the previous section to data in space and time. We base our approach in process convolutions as in Higdon (2002). In fact, process convolutions provide a convenient linear representation of Gaussian processes. Consider that at time t we have a vector $y_t = (y_{1,t}, \ldots, y_{n_t,t})'$ for which observations were recorded at the sites $s_1, \ldots, s_{n_t}$. A space-time formulation for y_t is defined as,

$$\begin{aligned} y_t &= K^t x_t + \epsilon_t \\ x_t &= x_{t-1} + \nu_t \end{aligned}$$

where K^t is a $n_t \times \kappa$ matrix given by $K^t_{ij} = k(s_i - \omega_j),\ t = 1, \ldots, m$, where k is a given kernel. This model is included within the class of DLMs used in previous sections for temporal modeling. We assume that the observation error $\epsilon_t \sim N(0, \sigma^2_\epsilon I_{n_t})$ and the evolution error $\nu_t \sim N(0, \sigma^2_\nu I_\kappa)$, for $t = 1, \ldots, m$. The prior for the state vector x_0 is $x_0 \sim N(0, \sigma^2_x I_\kappa)$.

As mentioned above, the elements $k(\cdot - \omega_j)$ are defined by a smoothing kernel where $\omega_1, \ldots, \omega_\kappa$ are the spatial sites or knots where the kernels are centered. Since the error terms are not spatially correlated, the spatial dependency of this model is completely defined by K. Some recommended kernels are *Gaussian* $k(s) \propto \exp\left\{-||s||^2/2\eta^2\right\}$; *Exponential* $k(s) \propto \exp\left\{-||s||/\eta^2\right\}$;, for some $\eta > 0$ and *Spherical* $k(s) \propto \left(1-||s||^3/r^3\right)^3 I[s \leq r]$, for a given r. Doing a convolution at $\omega_1, \ldots, \omega_\kappa$ provides a finite approximation to the continuous convolution. For the integral convolution it is possible to find an equivalence between the kernels and the covariance function of the process. See Kern (2000) for more details.

In order to merge the spatial DLM with the analysis of the extreme values, assume that $z_{s,t} \sim GEV(\mu_{s,t}, \sigma, \xi)$ for $s = 1, \ldots, S$ and $t = 1, \ldots, m$. The distribution function for $z_{s,t}$ is

$$H_{s,t}(z_{s,t}; \mu_{t,s}, \xi, \sigma) = \exp\left\{ - \left[1 + \xi\left(\frac{z_{s,t} - \mu_{s,t}}{\sigma}\right)\right]_+^{-1/\xi} \right\}.$$

For each t let $\mu_t = (\mu_{1,t}, \mu_{2,t}, \ldots, \mu_{S,t})'$. We now define a DLM on μ_t using the process convolution approach described previously.

$$\begin{aligned} \mu_t &= K'\theta_t + \epsilon_t; \\ \theta_t &= \theta_{t-1} + \nu_t \end{aligned}$$

where the state vector $\theta_t = (\theta_{t,1}, \ldots, \theta_{t,\kappa})'$, $\epsilon_t = (\epsilon_{t,1}, \ldots, \epsilon_{t,\kappa})'$, $\nu_t = (\nu_{t,1}, \ldots, \nu_{t,\kappa})'$. Additionally, $\epsilon_t \sim N(0, \sigma_\epsilon^2 I_S)$ and $\nu_t \sim N(0, \sigma_\nu^2 I_\kappa)$. Using a Gaussian kernel we have that the $S \times \kappa$ matrix K' is given by $K_{ij} = \exp(-d||s_i - \omega_j||^2/2)$ where s_i is the position of station i; ω_j is the center of the jth kernel $j = 1, \ldots, \kappa$ and d is a range parameter.

For the parameters in the first stage of the model we specify the prior distributions $\pi(\sigma) \sim LN(\mu_\sigma, s_\sigma)$ and $\pi(\xi) \sim N(\mu_\xi, s_\xi)$. For the parameters in the second stage, $\theta_0 \sim N(0, \sigma_\theta^2 I_\kappa)$; $1/\sigma_\epsilon^2 \sim Gamma(\alpha_\epsilon, \beta_\epsilon)$; $1/\sigma_\nu^2 \sim Gamma(\alpha_\nu, \beta_\nu)$ and $1/\sigma_\theta^2 \sim Gamma(\alpha_\theta, \beta_\theta)$. The log-likelihood for the first-stage model parameters is

$$\begin{aligned} l(\theta) &= -mS\log\sigma - \sum_{t=1}^{m}\sum_{s=1}^{S}\left[1 + \xi\left(\frac{y_{s,t} - \mu_{s,t}}{\sigma}\right)\right]_+^{-1/\xi} \\ &\quad - \left(1 + \frac{1}{\sigma}\right)\sum_{t=1}^{m}\sum_{s=1}^{S}\log\left[1 + \xi\left(\frac{y_{s,t} - \mu_{s,t}}{\sigma}\right)\right]_+ \end{aligned}$$

Posterior inference and simulation for this space-time model follows a similar structure as for the GEV time-varying model. The full conditional for $\mu_{t,s}$, σ and ξ are sampled each with a Metropolis-Hastings step. We draw the state vector θ_t with Forward Filtering Backward Simulation. The full conditionals of σ_ϵ^2; σ_ν^2; σ_θ^2 are sampled with Inverse Gamma distributions. Traditionally, for process convolutions, kernels are centered around points on a regular grid and the range parameter is fixed as $d = c\phi$, where where ϕ is the distance between kernel and c is a constant, usually between $1/2$ and 2.

We consider the problem of modeling daily maxima of ozone in Mexico City during the year 1999. Daily maxima are obtained from the Red Automática de Monitoreo Ambiental (RAMA) hourly measurements that come from 19 monitoring stations. A spatio-temporal model for the actual hourly ozone observations is developed in Huerta et al. (2004). That paper focuses on interpolation and short term forecasting of the hourly ozone levels and not in extremes. The model is based on the amplitudes of daily fluctuations of ozone. Such amplitudes are assumed to vary in space and time and depend on covariates that also vary in space and time. In Figure 8 we present the RAMA stations along with the kernel positions selected to define the process convolution on the location parameters. The results that we present in this paper are based on a sixteen points grid. We fitted the model using a nine and a four points grid and observed no substantial differences in the diagnostics and quantile estimates. Varying the choice of the parameters that define our priors also produced fairly robust results.

In Figure 9 we show the data for four stations and the estimates of the posterior medians of the time-varying GEV distributions. Since there are no replicate observations for any given day, it is hard to interpret precisely the meaning of the median at time t. Nevertheless, we expect that the curve will split the cloud of points, roughly, in two halves. To asses the validity of the model, we obtained the predictions for each one of the locations after leaving them out, one at a time. For each station we performed a diagnostic based on

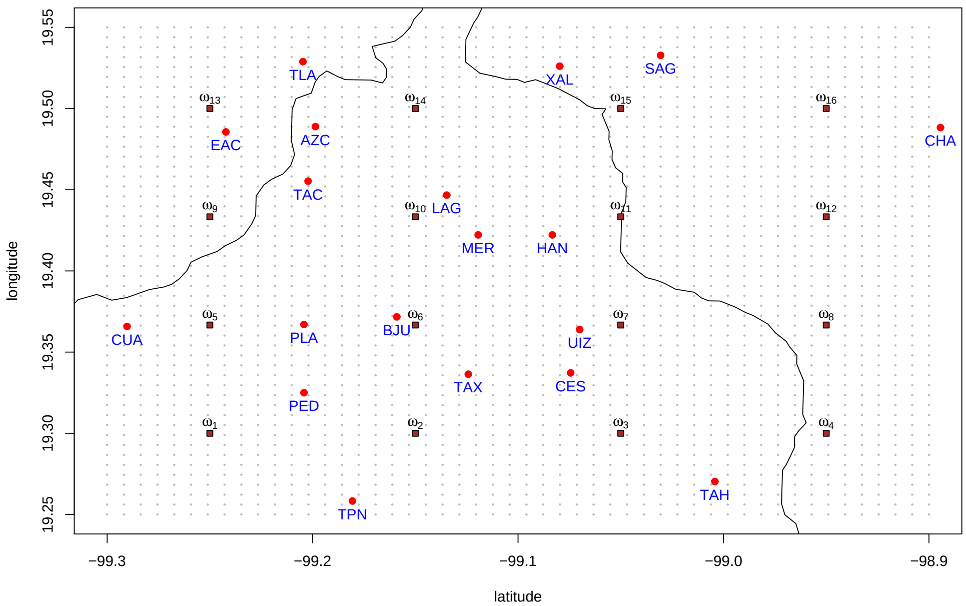

Fig. 8. RAMA stations, kernel and interpolation grid positions. The stations are denoted with a three letters code. The kernel centers are denoted with ω_i. The grid points correspond to the locations used to obtain the values shown in Figure 11. The continuous line corresponds to the boundaries of Mexico City.

the u_t transformation, in analogy to the analysis presented in Section 2. As for the univariate case we did a graphical analysis of the autocorrelations and the qq-plots. The results for the four stations considered in Figure 9 are presented in Figure 10. Overall the diagnostics were fairly acceptable. In addition, we noticed that the fit that used all the data produced results that are very similar to those obtained after deleting any one station.

The spatial and temporal variability of the results is shown in Figure 11. We observe a clear gradient increasing in the NE-SW direction. This is consistent with the results in Huerta et al. (2004). The direction of the gradient remains essentially unchanged in time, but the medians vary substantially from day to day. An exploration of the posterior distribution of ξ shows that the shape is clearly a negative parameter, indicating that the distribution belongs to the Weibull family. A 95% posterior probability interval for ξ is $(-0.219, -0.186)$. This implies that the density fitted to the data has a support that is bounded above. For any station, the bound is time-varying and is given by $\mu_{s,t} - \sigma/\xi$.

Coles et al. (1999) propose a summary of the extremal dependence of two random variables, (X, Y) by transformation to variables, say, (U, V) with uniform marginals. This is given by

$$\chi(u) = 2 - \frac{\log Pr(U < u, V < u)}{\log Pr(U < u)} \quad 0 \leq u \leq 1.$$

This is an approximation to $Pr(V > u|U > u)$. We define a time-varying version of this measure for a pair of locations s and s' using the predictive posterior distributions. Thus

$$\chi_t(u, s, s') = 2 - \frac{\log Pr(U_{t,s} < u, U_{t,s'} < u|D)}{\log Pr(U_{t,s} < u|D)} \quad 0 \leq u \leq 1. \tag{2}$$

where $U_{t,r} = H_{t,r}(z_{t,r}), r = s, s'$ and D denotes all available data. If a sample from the joint posterior distribution of the parameters is available, say, $\mu_{t,s}^{(i)}, \xi^{(i)}, \sigma^{(i)}, i = 1, \ldots, M$, then (2) can be approximated by computing the proportion of pairs $(U_{t,s}^{(i)}, U_{t,s'}^{(i)})$ less than a fixed u and dividing this proportion by the

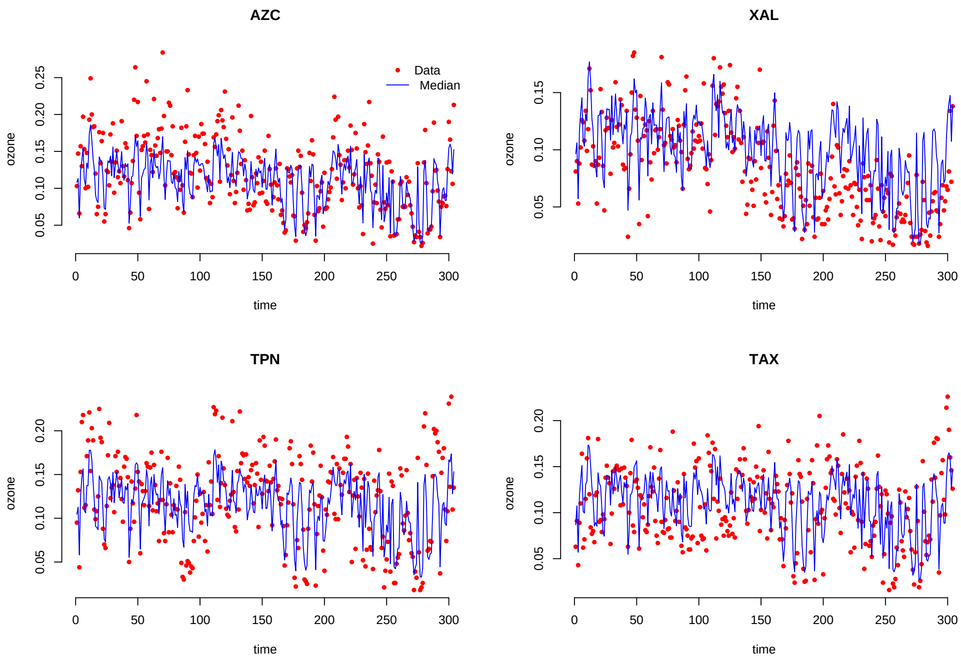

Fig. 9. Daily maxima for 1999 (dots) and posterior estimate of the 0.50 quantile of the GEV distribution at each time point. (continuous line). The quantile is based in a leave-one-out analysis. The chosen stations correspond to locations in the NE, NW, SE and SW respectively, as shown in Figure 8.

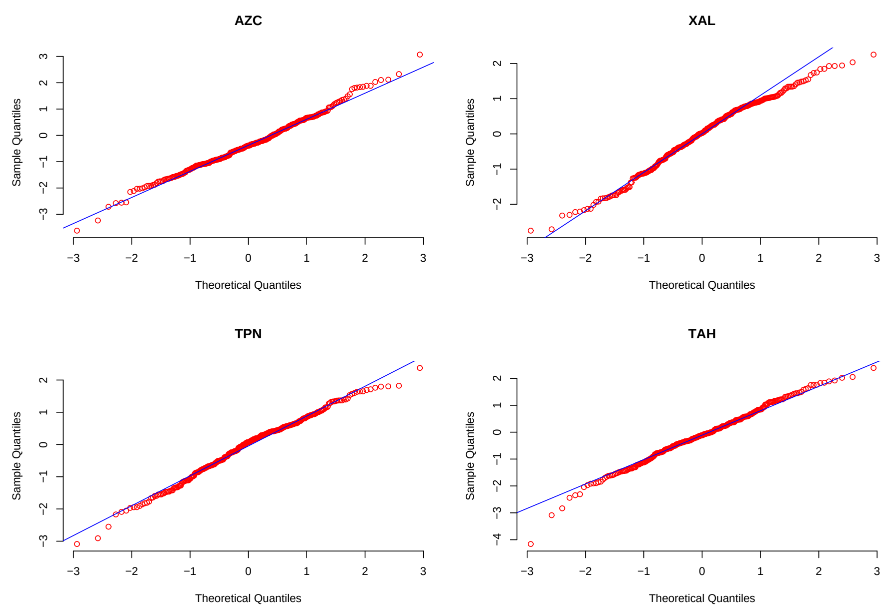

Fig. 10. u_t diagnostics for four of the stations using the posterior predictions based on leaving one station out at a time.

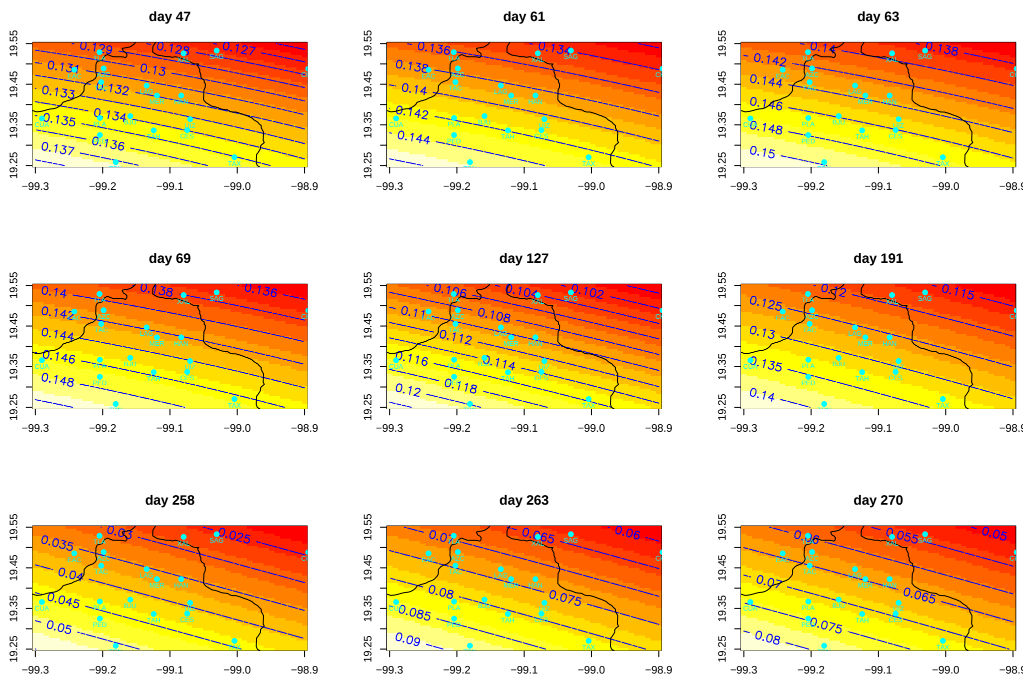

Fig. 11. Posterior estimate of the 0.5 quantile of the space-time GEV distribution for a 50×50 resolution grid. Dots correspond to the locations of the stations. The boundary of Mexico City is shown as a reference. The contour lines correspond to fixed ozone levels in ppm.

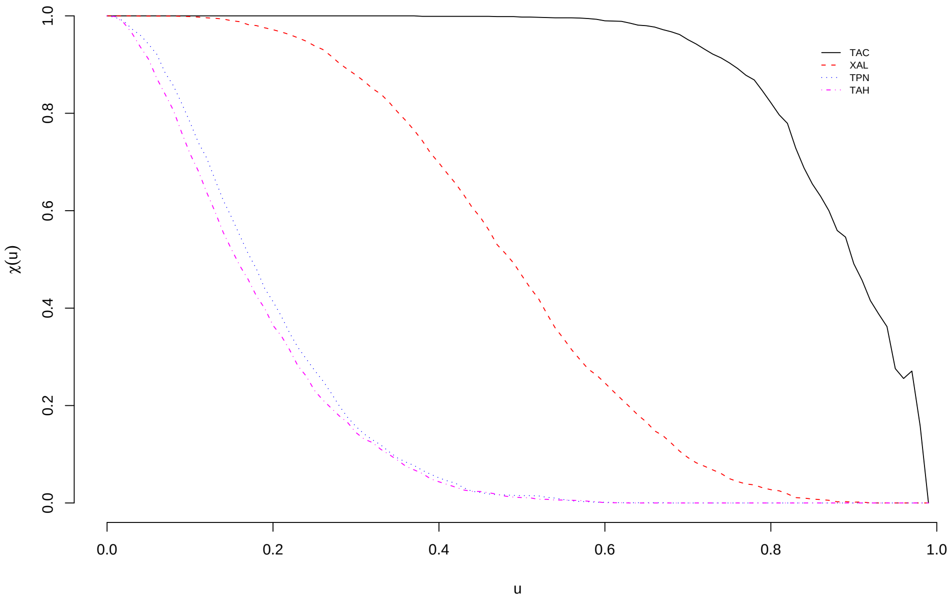

Fig. 12. Dependency between the maxima at AZC and TAL, XAC, TPN and TAH for $t = 200$ and different values of u.

proportion of cases where $U_{t,s}^{(i)} < u$. Figure 12 shows the graph of $\chi_{200}(u, s, s')$ for s corresponding to the location of station AZC and s' to four other different stations. The fact that all curves tend to zero for u close to 1, suggests that the model is producing asymptotic independence. Nevertheless, we observe that the curves have behaviors that depend strongly on the locations of the stations, for a wide range of u values. This indicates that, in practice, the model is able to capture different types of pairwise dependencies.

4. Discussion

We have presented a general framework for modeling extreme values that exhibit a time-varying behavior. The models are based on the use of the GEV distribution and assume that the location parameter varies with time according to a DLM. For the univariate case we presented four different models and two different data sets. Our analysis reveals that the proposed dynamic approach provides a flexible way of capturing short and long range trends as well as time-varying dependencies. The approach can be easily extended to account for spatial variability by using process convolutions. The assumption of a GEV distribution for the data is an approximation, but methods based on quantiles can be used to assess the goodness of fit of the model. The use of the GEV distribution makes inference for quantiles straightforward, since the inverse of the cdf is available in closed form. Within a MCMC approach all that is needed is the evaluation of such inverse for each joint sample of the parameters. So, model checking statistics are easy to compute. The asymptotic dependence induced by the model needs to be theoretically assessed. Empirical measures based on Monte Carlo approximations indicate that the model is fairly flexible.

Many possible extensions are possible to models considered in this paper. As an alternative or in addition to a time-varying location we can consider time-varying shape and scale. This will provide more flexibility for the description of the extremal behaviors. For the scale we can consider an evolution like the one provided by a DLM after transformation to the log scale. In the spatio-temporal case we can consider scale and shape parameters that vary with location. We would like to mention that a station by station analysis of the ozone maxima did not provide grounds for such an extension in the particular case of the data considered in this paper.

Finally it is possible to consider methods based on threshold models. This presents the challenge of dealing with possible time-varying or location dependent thresholds. We would like to mention that for the ozone data we fitted traditional threshold models station by station. We observed, for some stations, strong sensitivity to the choice of the threshold. Also we observed that optimal choices of thresholds suggested values that varied substantially for nearby stations.

Acknowledgments

The authors will like to acknowledge the help of Lelys Guenni on the discussion of the Venezuelan rainfall data. We also thank an anonymous referee whose comments lead to an improved version of the original paper. The second author was partially supported by NSF grant DMS-0504851.

References

Assunção, C. B. N., Gamerman, D. and Lopes, H. F. (2004) Bayesian analysis of extreme events with threshold estimation. *Statistical Modelling*, **4**, 227–244.

Carter, C. and Kohn, R. (1994) On Gibbs sampling for state space models. *Biometrika*, **81**, 541–553.

Casson, E. and Coles, S. (1999) Spatial regression models for extremes. *Extremes*, **1**, 449–468.

Coles, S. (2001) *An Introduction to Statistical Modeling of Extreme Values.* New York, USA: Springer-Verlag.

Coles, S., Heffernan, J. and Tawn, J. (1999) Dependence measures for extreme value analyses. *Extremes*, **2:4**, 339–365.

Coles, S. and Tawn, J. (1996) A Bayesian analysis of extreme rainfall data. *Applied Statistics*, **45**, 463–478.

Coles, S. G., Pericchi, L. R. and Sisson, S. A. (2003) A fully probabilistic approach to extreme value modelling. *Journal of Hydrology*, **273**, 45–50.

Davison, A. C. and Smith, R. L. (1990) Models for exceedances over high thresholds (with comments). *Journal of the Royal Statistical Society, Series B, Methodological*, **52**, 393–442.

Dupuis, D. (2005) Ozone concentrations: a robust analysis of multivariate extremes. *Technometrics.* To appear.

Frühwirth-Schnatter, S. (1994) Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, **15**, 183–202.

Gaetan, C. and Grigoletto, M. (2004) Smoothing sample extremes with dynamic models. *Extremes*, **7**, 221–236.

Gamerman, D. (1997) *Markov Chain Monte Carlo.* London, UK: Chapman and Hall.

Gilleland, E., Nychka, D. and Schneider, U. (2004) Spatial models for the distribution of extremes. *Tech. rep.*, Geophysical Statistics Project, NCAR. Summer Institute at the Center on Global Change, Duke University.

Heffernan, J. and Tawn, J. (2004) A conditional approach for multivariate extreme values (with discussion). *J. R. Statist. Soc. B*, **66**, 497–546.

Higdon, D. (2002) Space and space-time modeling using process convolutions. In *Quantitative Methods for Current Environmental Issues* (eds. C. Anderson, V. Barnett, P. C. Chatwin and A. H. El-Shaarawi), 37–56. London: Springer Verlag.

Huerta, G., Sansó, B. and Stroud, J. R. (2004) A spatio-temporal model for Mexico city ozone levels. *Applied Statistics*, **53**, 231–248.

Kern, J. (2000) *Bayesian Process-Convolution Approaches to Specifying Spatial Dependence Structure.* Ph.D. thesis, Institute of Statistics and Decision Sciences, Duke University, Durham, North Carolina, USA.

Kim, S., Shephard, N. and Chib, S. (1998) Stochastic volatility: likelihood inference and comparison with arch models. *Rev. Fin. Stud.*, **65**, 361–393.

Ledford, A. and Tawn, J. (1996) Statistics for near independence in multivariate extreme values. *Biometrika*, **83**, 169–187.

— (1997) Modeling dependence within joint tail regions. *J. R. Statist. Soc. B*, **59**, 475–499.

Pickands, J. (1971) The two-dimensional Poisson process and extremal processes. *Journal of Applied Probability*, **8**, 745–756.

— (1975) Statistical inference using extreme order statistics. *The Annals of Statistics*, **3**, 119–131.

Rosenblatt, M. (1952) Remarks on a multivariate transformation. *Annals of Mathematical Statistics*, **23**, 470–472.

Smith, R. L. (1989) Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone (C/R: P378-393). *Statistical Science*, **4**, 367–377.

Smith, R. L., Tawn, J. A. and Coles, S. G. (1997) Markov chain models for threshold exceedances. *Biometrika*, **84**, 249–268.

West, M. and Harrison, J. (1997) *Bayesian Forecasting and Dynamic Models.* New York: Springer Verlag, second edn.