Bayesian spatial extreme value analysis to assess the changing risk of concurrent high temperatures across large portions of European cropland

Benjamin A. Shaby¹,* and Brian J. Reich²

¹Department of Statistics, University of California, Berkeley

²Department of Statistics, North Carolina State University

*Correspondence to: Benjamin A. Shaby, Department of Statistics, UC Berkeley, Berkeley, CA 94720, U.S.A. E-mail:bshaby@stat.berkeley.edu

October 1, 2012

Abstract

There is strong evidence that extremely high temperatures are detrimental to the yield and quality of many economically and socially critical crops. Fortunately, the most deleterious conditions for agriculture occur rarely. We wish to assess the risk of the catastrophic scenario in which large areas of croplands experience extreme heat stress during the same growing season. Applying a hierarchical Bayesian spatial extreme value model that allows the distribution of extreme temperatures to change in time both marginally and in spatial coherence, we examine whether the risk of widespread
extremely high temperatures across agricultural land in Europe has increased over the last century.

Keywords: max-stable process, agriculture, climate change

Running head: Bayesian extreme value analysis of temperatures in European crop-land
1 Introduction

As more attention continues to focus on the impact of climate change on weather extremes, the role of extreme value theory grows more prominent. Given the potentially disastrous social, economic, and public health impacts of weather-related events such as heat waves, droughts, and floods, and given the relative rarity of such events, applying the insights of extreme value theory to environmental risk analysis is both a necessity and a natural evolution.

Models used for assessing whether the risk of widespread extreme temperature events, with an eye toward heat-induced crop failure, has changed in the past century require three features. First, since the events of interest occur far in the tail of the temperature distribution, models must leverage extreme value theory. Second, because catastrophic crop failure occurs exactly when large areas of farmland fail during the same growing season, the nature of the problem is inherently spatial, and requires models to accurately represent spatial dependence. Finally, as the most relevant questions from a risk analysis perspective are most naturally assessed as probability statements about complicated functionals of unknown quantities, such analysis calls for a Bayesian solution. With these requirements in mind, we apply a hierarchical Bayesian spatial extreme value model that allows the distribution of extreme temperatures to change in time both marginally and in spatial coherence to examine whether the risk of widespread extremely high temperatures across agricultural lands in Europe has increased over the last century.

There is growing evidence that extremely high temperatures are detrimental to the quantity and quality of agricultural production. Schlenker and Roberts (2009) investigate the temperature effects on yields of corn, soybeans, and cotton in the U.S.. For corn and soy, they find that production drops precipitously when temperatures exceed 30°C. They apply their fitted model to conservative future warming scenarios and predict a drastic drop in
Lobell and Field (2007) find a decreasing linear trend with temperature for all of the six most important worldwide food crops. This study regressed global yields on maximum and minimum temperatures extracted from a gridded data product (Mitchell and Jones, 2005). They find that although yields have increased, the increase has been slowed by rising temperature. We note that the interpolation scheme that produced the gridded product used in this study (Mitchell and Jones, 2005) is quite ad hoc from a statistical perspective and uses no extreme value theory. Nevertheless, although we view the temperature dataset with some suspicion, the inferred relationship appears robust. For example, Lobell et al. (2011) produces similar results, using average temperatures to compute trends in production on a country-by-country level.

Despite a clearly-established relationship between very high temperatures and poor agricultural production, extreme value analysis of catastrophic heat events in the context of agriculture has not previously been attempted. We seek to fill this gap. It is immediately obvious that such an analysis must accurately characterize both the tail of the marginal temperature distributions at locations of interest and the spatial dependence of rare events. To this end, we look to the family of spatial max-stable process models (Davison et al., 2012, for a review). These models provide the tail behavior required by classical extreme value theory and explicitly represent spatial dependence among the extreme events. Previous applications of spatial max-stable processes include Smith and Stephenson (2009), Shang et al. (2011), and Westra and Sisson (2011).

While it is possible to model temperature fields with flexible models that can capture broad ranges of tail behaviors, spatial quantile regression for example (Reich, 2012; Tokdar and Kadane, 2011), an extreme value approach is preferred for this application. For very rare events, the scarcity of available data is likely insufficient to accurately fit the far tail of a non-parametric or semi-parametric model. In contrast, asymptotic theory allows us
to fit the tail using only three parameters. An important consequence is that high quantiles, such as 50-year return levels (e.g. Figure 6(a)), can be estimated from only a few decades worth of data.

With an appropriate class of models in hand, we crystallize our analysis into the following two questions. First, for given values $p$ and $T$, what is the probability a fraction $p$ of the agricultural land will experience a temperature higher than $T$ in a given year? And second, is that probability changing in time? These key questions are most naturally answered from within the Bayesian paradigm, which produces a joint posterior distribution from which one easily generates such probability statements about complicated transformations of model parameters. Toward this end, we employ the hierarchical spatial max-stable model of Reich and Shaby (2012), which is thus far unique in that it permits fully Bayesian analysis of a class of spatial extreme value processes. We will refer to this model as the hierarchical kernel extreme value process (HKEVP).

The strategy is to draw from the joint posterior predictive distribution at a set of prediction locations over several time points. We can then see, at each time point, how frequently and to what spatial extent these draws exceed some high temperature threshold. Since we are interested in questions related to crop failure, we only draw from the predictive distribution at locations that are currently used for agriculture.

While the HKEVP model has several attractive features, it does have a drawback for analysis motivated by crop failures. By only considering annual maximum daily temperatures, the HKEVP, like all max-stable process models, is unable to distinguish between different extreme events that occur in the same year. Since extreme high temperatures at different sites across Europe might occur at different times within a given summer, realizations from the HKEVP do not necessarily represent simultaneous phenomena. This limitation is mitigated somewhat by the large-scale structure of temperature fields, which often leads to annual maxima being attained simultaneously across wide areas. In addition, regardless of
whether extreme events occur strictly concurrently, widespread extreme high temperatures across agricultural lands occurring during the same growing season are likely to have significant societal impacts. A second limitation of our analysis is that agricultural activity within the study area is heterogeneous, with different crops (and different varieties of each crop) planted at different locations at different times. Any fixed threshold therefore does not have the same impact throughout all of Europe. Furthermore, the timing of extreme heat events within the growing season is likely important due to crop phenology, a feature that our block maxima approach cannot capture. However, despite these limitations, the max-stable process framework is the best available tool for studying spatial fields of extreme values.

2 A spatial max-stable process model

The spatial extreme value model of Reich and Shaby (2012), which we describe below, builds on the Gaussian extreme value process (GEVP) of Smith (1990). Extreme value processes arise as the limit of taking pointwise block maxima over independent stochastic processes. In the context of extreme temperatures, it is natural to define a block maximum at each location as the annual high temperature. Thus, considering the daily temperature as a spatial stochastic process, the annual maximum temperature at each point in the domain under consideration asymptotically approaches an extreme value process. The spatial process that we consider possesses the key properties of max-stability and extremal dependence. Max-stability implies that taking pointwise maxima of the spatial process results, after appropriate scaling, in a process that is identical in law to the original process. This property is important because it characterizes the class of extreme value processes, so it is necessary if we want to harness extreme value theory. Extremal dependence says that the joint probability of the temperature at two fixed locations exceeding some high threshold does not vanish as
the threshold increases to infinity. Extremal dependence would seem to be a requirement to model high temperatures, as events like heat waves have obvious spatial dependence. Exploratory analysis (Figure 4) confirms this feature in the data.

The HKEVP model for the field $Y(s)$ of annual maximum temperatures is defined hierarchically as follows. First, choose a large number $L$ of fixed spatial knots $v = \{v_1, \ldots, v_L\}$. Next, define the collection of scaled Gaussian densities $\varphi_{w,l}(s \mid v, \Sigma) = \varphi(s \mid v_l, \Sigma) / \sum_{j=1}^{L} \varphi(s \mid v_j, \Sigma)$. The scaling just enforces that the $\varphi_{w,l}(s \mid v, \Sigma)$ sum to 1 for all $s$, which is needed to produce unit Fréchet margins in the final model. The core of the hierarchical model is the quantity

$$\vartheta(s) = \left( \sum_{l=1}^{L} A_l \varphi_{w,l}(s \mid v, \Sigma)^{1/\alpha} \right)^{\alpha},$$

(1)

where $A_l \sim PS(\alpha)$ independently for $l = 1, \ldots, L$, and $PS(\alpha)$ denotes a positive stable distribution with parameter $\alpha \in (0, 1)$ (Samorodnitsky and Taqqu, 1994). The construction (1) is similar to that of Stephenson (2009) and Fougères et al. (2009), and induces spatial dependence through the (scaled) kernel functions $\varphi_{w,l}, l = 1, \ldots, L$. The hierarchical model for the observations at $s_1, \ldots, s_n$ of the process $Y(s)$ is

$$Y(s_i) \mid A_1, \ldots, A_L, \mu, \sigma, \xi, \alpha \overset{\text{iid}}{\sim} \text{GEV}(\mu^*(s_i), \sigma^*(s_i), \xi^*(s_i))$$

$$\mu^*(s) = \mu + \frac{\sigma}{\xi} [\vartheta(s)]^{\xi} - 1$$

$$\sigma^*(s) = \alpha \sigma \vartheta(s)^{\xi}$$

$$\xi^*(s) = \alpha \xi$$

$$A_l \mid \alpha \overset{\text{iid}}{\sim} PS(\alpha) \quad \text{for} \ l = 1, \ldots, L$$

(2)

where $\vartheta(s)$ is defined as in (1) and GEV($\mu, \sigma, \xi$) denotes the generalized extreme value distribution (see, e.g. Coles, 2001, page 47) with location parameter $\mu$, scale parameter $\alpha$, and shape parameter $\xi$. 

7
The HKEVP model defined by (1) and (2) is max-stable both marginally and conditionally on the spatial random effect (Reich and Shaby, 2012). Marginally over $A_1, \ldots, A_L$, $Y(s_0) \sim \text{GEV}(\mu, \sigma, \xi)$ at any location $s_0$. The parameter $\alpha$ controls the smoothness of the process in the following way. As $\alpha \to 1$, the latent variables $A_1, \ldots, A_L$, and thus $\vartheta(s)$, converge in distribution to the constant 1, and $Y(s)$ converges in distribution to $\text{GEV}(\mu, \sigma, \xi)$ independently for all $s$, resulting in no spatial dependence. Conversely, as $\alpha \to 0$, realizations of $Y(s)$ become continuous. The parameter $\alpha$ can therefore be seen as a small-scale error term analogous to the nugget variance in classical geostatistics. If both $\alpha \to 0$ and the number of knots $L \to \infty$, $Y(s)$ converges in distribution to the GEVP of Smith (1990). The effect of $\alpha$ on the hierarchically-defined process (2) is shown in Figure 1. When $\alpha = 0.2$, the process realization appears very smooth with obvious strong spatial dependence. The smoothness attenuates with increasing $\alpha$. When $\alpha = 0.8$, little trace of spatial dependence remains visible.

In this way, the HKEVP defined by (1) and (2) may be seen as an extension of the GEVP, improving upon the GEVP in two ways. First, the GEVP has been criticized as appearing implausibly smooth and has been diagnosed as fitting several datasets less satis-
factorily than more recently-proposed spatial max-stable processes (Smith and Stephenson, 2009; Davison et al., 2012). The HKEVP described above adds flexibility by including the parameter $\alpha \in (0, 1)$ which allows the resultant process to vary from very smooth process realizations to very rough ones (see Figure 1). Second, a major hurdle to more widespread application of spatial max-stable processes like the GEVP and related models is that joint densities corresponding to their finite-dimensional distributions are not known for more than a trivial number of locations. This precludes maximum likelihood and Bayesian estimation, although methods based on composite likelihoods have had some success (Padoan et al., 2010; Cooley et al., 2012; Shaby, 2012). In contrast, the HKEVP model (1) and (2) does not suffer this deficiency. It enables Bayesian formulation and attendant Markov chain Monte Carlo (MCMC) sampling (see, e.g. Robert and Casella, 2004) by relying on a random effects representation and considering a large but finite number of Gaussian densities.

2.1 Rounding out the model

Returning now from the generic case to the task of modeling extreme high temperatures in Europe, we begin to flesh out the model. Since we are considering such a large and heterogeneous geographical region, it is necessary to allow the marginal GEV parameters $\mu$, $\sigma$, and $\xi$ to vary spatially. In addition, since we are interested not only in the risk of widespread extremely high temperatures, but also how that risk might be changing in time, we must model temporal evolution. Thought of as a process in space and time, we denote the data generating extreme value process as $Y(s, t)$.

We model underlying spatial variation by assigning Gaussian process (GP) priors to the
scale and shape parameters, now denoted \(\sigma(s)\), and \(\xi(s)\):

\[
\sigma(s) \sim \text{GP}(X(s)\beta_{\sigma}, \Sigma_{\sigma}(\theta_{\sigma}))
\]

\[
\xi(s) \sim \text{GP}(X(s)\beta_{\xi}, \Sigma_{\xi}(\theta_{\xi})),
\]

where \(X(s)\) is a matrix of covariates consisting of the latitude, longitude, and elevation of each location. Temporal evolution in the distribution of extreme values is modeled as a linear trend in the GEV location parameter, now denoted \(\mu(s, t)\) and represented as \(\mu(s, t) = b_0(s) + b_1(s)t\), where \(b_0(s)\) and \(b_1(s)\) are both allowed to vary in space through Gaussian process priors of their own,

\[
b_j(s) \sim \text{GP}(X(s)\beta_{\mu,j}, \Sigma_{\mu,j}(\theta_{\mu,j})) \quad j = 0, 1.
\]

The covariance matrix for the \(\sigma(s)\) process is parametrized as \(\Sigma_{\sigma}(\theta_{\sigma})_{ij} = \tau^2_{\sigma}\exp(-\|s_i - s_j\|/\rho_{\sigma})\), where \(\theta_{\sigma} = (\tau^2_{\sigma}, \rho_{\sigma})^T\), and analogously for \(\Sigma_{\xi}(\theta_{\xi})\), \(\Sigma_{\mu,1}(\theta_{\mu,1})\), and \(\Sigma_{\mu,2}(\theta_{\mu,2})\). Marginally over \(A_1, \ldots, A_L\), \(Y(s_0, t_0) \sim \text{GEV}((\mu(s_0, t_0), \sigma(s_0), \xi(s_0))\) at any location \(s_0\) and time \(t_0\).

The range of spatial dependence of the hierarchical extreme value process is also allowed to change in time. The covariance matrix of the bivariate Gaussian kernel in (1), which we now call \(\Sigma(t)\), and which controls the shape and range of spatial dependence, is given the form \(\gamma^2(t)I\) for simplicity, restricting the process to be isotropic. As a result, a single parameter \(\gamma(t)\) controls the range of spatial dependence of the extreme events. Temporal dependence is modeled as a linear trend on the log scale, \(\log(\gamma(t)) = \beta_{0,\gamma} + \beta_{1,\gamma}t\). In this way, both the intensity of the extreme high temperatures, through \(\mu(s, t)\), and the range of their spatial dependence, through \(\gamma(t)\), are allowed to change in time.
Finally, vague prior distributions for the remaining parameters were specified as

\[
\alpha \sim \text{Unif}(0, 1)
\]

\[
\beta_{0,\gamma^2} \sim \mathcal{N}(0, 10^2), \quad \beta_{1,\gamma^2} \sim \mathcal{N}(0, 1^2)
\]

\[
\beta_{\mu,j} \sim \mathcal{N}(0, \sigma_{\beta_{\mu,j}}^2 \mathbf{I}), \quad j = 1, 2, \quad \beta_{\sigma} \sim \mathcal{N}(0, \sigma_{\beta_{\sigma}}^2), \quad \beta_{\xi} \sim \mathcal{N}(0, \sigma_{\beta_{\xi}}^2)
\]

\[
\tau_{\mu}^2, \tau_{\sigma}^2, \tau_{\xi}^2 \sim \text{InvGamma}(0.1, 0.1)
\]

\[
\rho_{\mu}^2, \rho_{\sigma}^2, \rho_{\xi}^2 \sim \text{InvGamma}(0.1, 0.1)
\]

\[
\sigma_{\beta_{\mu,1}}^2, \sigma_{\beta_{\mu,2}}^2, \sigma_{\beta_{\sigma}}^2, \sigma_{\beta_{\xi}}^2 \sim \text{InvGamma}(0.1, 0.1)
\]

### 2.2 Spatial prediction

Because we wish to investigate the behavior of temperature extremes occurring in large agricultural areas simultaneously, we must confront the problem of spatial prediction. That is, we are interested in the value of the random field at locations other than the observation locations. For Gaussian process models, spatial prediction is accomplished through the classical geostatistical technique Kriging. Under previous max-stable process models, spatial prediction is possible but complicated (see Wang and Stoev, 2010, e.g.). The HKEVP, in contrast, permits straightforward spatial prediction at unobserved sites.

Prediction at an unobserved location \( s_p \) and time \( t_p \) proceeds as follows. At each MCMC iteration \( i \)

1. Construct \( \varphi^{(i)}(s_p) \) given \( A_1^{(i)}, \ldots, A_L^{(i)} \) and \( \alpha^{(i)} \) as in (1).

2. Compute \( \mu^{*(i)}(s_p, t_p), \sigma^{*(i)}(s_p), \) and \( \xi^{*(i)}(s_p) \) given \( \varphi^{(i)}(s_p) \) as in (2).

3. Draw \( Y_p(s_p, t_p) \sim \text{GEV}(\mu^{*(i)}(s_p), \sigma^{*(i)}(s_p, t_p), \xi^{*(i)}(s_p)) \).

There is some subtlety in choosing between two types of predictive draws. The first, which we will call “Kriging-type” draws, conditions on \( A_1^{(i)}, \ldots, A_L^{(i)} \) at each MCMC iteration.
second, which we refer to as “climatological-type” draws, does not use the current value of 
$A_1, \ldots, A_L$, but rather conditions on $A_1^*, \ldots, A_L^* \overset{\text{iid}}{\sim} \text{PS}(\alpha^{(i)})$, draws from the prior given the 
current value of $\alpha$.

Kriging-type predictive draws interpolate the extreme temperature field as it was actually 
observed in a given year. Climatological-type draws produce a realization of stochastic 
process that has the same distribution as the extreme temperature field in a given year. In 
other words, Kriging-type draws estimate what really did happen at unobserved locations, 
whereas climatological-type draws sample from the distribution of what could have happened. 
For the purposes of assessing risk of widespread crop failure due to very high temperatures, 
our primary interest is in climatological-type spatial draws.

Having established a method for spatial prediction, we can now compute the posterior 
distribution of the fraction $p_T(t_p)$ of prediction sites that exceed some threshold $T$ at time 
t_p. At each MCMC iteration $i$, we compute climatology-type predictive draws $Y_p^{(i)}(s_{p,j}, t_p)$ 
for every prediction site $s_{p,j}$, $j = 1, \ldots, N_p$. We then compute

$$p_T^{(i)}(t) = \frac{1}{N_p} \sum_{j=1}^{N_p} 1\{Y_p^{(i)}(s_{p,j}, t_p) > T\}, \quad (3)$$

the fraction of sites for which climatology-type draws exceed $T$. In this way, we obtain a 
sample of the posterior distribution of $p_T(t)$ for any desired threshold $T$ and time $t$.

All model parameters are sampled using fairly standard Metropolis-within-Gibbs tech-
niques (Robert and Casella, 2004). One challenging aspect is sampling the parameters 
$A_1, \ldots, A_L$, which have heavy-tailed positive stable priors, and some of which may be poorly 
identified. Extensive computational details, including a scheme to sample $A_1, \ldots, A_L$ using 
an auxiliary variables technique, are contained in Reich and Shaby (2012).
3 The data

Extreme temperature data was downloaded from the European Climate Assessment website (http://eca.knmi.nl/). This repository collects and collates daily data from well over a thousand meteorological stations and makes available a set of variables, such as daily temperature, precipitation, and cloud cover, as well as several indices of extremes, such as yearly high temperatures, number of frost days, and drought index, on a station-by-station basis (Tank et al., 2009). This dataset has undergone rigorous quality control to flag, for example, inhomogeneities due to station relocation or changing urban heating effects. Annual maximum temperatures for the years 1900–2011 were selected.

Landcover information was extracted from the MODIS MOD12 data product (http://www.modis.bu.edu). This 1-km resolution global landcover classification map was produced by passing retrievals from the MODIS instrument aboard the Terra satellite through a classification tree algorithm (Friedl et al., 2002). Obviously land use changes throughout the years. For our purposes, focusing attention on locations that are currently used for agriculture seems most natural, since our primary interest is in risk to agriculture as it is currently practiced.

Elevation data is from the Shuttle Radar Topography Mission (Farr et al., 2007) dataset SRTM30v2.1, downloaded from USGS at http://dds.cr.usgs.gov/srtm/. This is a 30 arc-second resolution data product derived from instruments flown aboard the space shuttle Endeavour in 1994.

The prediction grid was generated by transforming the vertices of a square grid into geographical longitude/latitude coordinates using the inverted Hammer projection (Snyder, 1987). Since this is an equal area projection, each grid point in geographical coordinates represents the same land area. The resultant coordinates were then checked against the MODIS classification map and retained only if they occurred within a pixel that was classified as agricultural land. This process resulted in 3696 prediction locations. Finally, elevations at the
prediction sites were obtained by cross-referencing their coordinates with the SRTM30v2.1 dataset. Prediction locations are shown as black dots in Figure 2. It is evident from Figure 2 that the geographic distribution of the temperature monitoring locations, which were used to fit the model, differs substantially from the geographic distribution of the agricultural prediction sites.

Figure 2: Prediction locations are indicated with black dots. Monitoring sites used to fit the model are shown in red. Knots are indicated with green ×’s

Spatial coverage of the monitoring stations included in the European Climate Assessment database is quite sparse in the beginning of the 20th century and extremely dense in some parts of the continent more recently. We thinned the set of monitoring stations by eliminating those with very short records, merging (i.e. retaining the max of the reported maxes) those with identical or nearly identical locations, and discarding those sited more than 250km from the nearest prediction point. The thinning process retained 985 monitoring sites. Locations of the stations used to fit the model are shown in red in Figure 2. Note that many of these sites appeared mid-century or later.
Knots were specified on a grid in geographical space and then removed if they were farther than 200km away from any observation or prediction point. The remaining \( L = 297 \) knots are shown as green \( \times \)'s on Figure 2. The distance between adjacent knots is roughly equal to the kernel bandwidth \( \gamma \), as estimated from preliminary model fitting. An extensive study of knot selection for the hierarchical kernel extreme value model may be found in Reich and Shaby (2012). The basic finding is that increasing the number of knots such that they are spaced at distances smaller than \( \gamma \) does not help model fitting. Conversely, placing knots farther apart than \( \gamma \) degrades performance in estimating both the GEV marginal parameters and the dependence parameters \( \gamma \) and \( \alpha \). The rule of thumb is therefore to try to place knots at distances roughly equal to \( \gamma \).

The form of the temporal dependence for the European dataset was investigated thoroughly through exploratory analysis. Maximum likelihood estimates of the GEV parameters \( \mu, \sigma, \) and \( \xi \) were fit to 30-year and 50-year moving windows of data on a station-by-station basis. These estimates were generally very noisy (demonstrating the importance of borrowing strength across space). Examples of 30-year moving window fits of \( \mu \) at 16 representative sites (see Figure S-1) are shown in Figure 3. There was no discernible evidence against keeping \( \sigma \) and \( \xi \) constant in time (see Figures S-2 and S-3 in the online supplement), but the plots in Figure 3 hinted at a possible linear time trend for \( \mu \). Similarly, F-madogram analysis (Cooley et al., 2006), which showed clear evidence of extremal dependence, revealed no obvious form for the time evolution of \( \gamma \), but suggested a slight increase in time. Figure 4 shows empirical F-madogram estimates of extremal coefficients (Smith, 1990) for 30 years of data around 4 reference years in the study period. The curves are clearly below 2 at moderate spatial lags indicating the presence of extremal dependence. In addition, it appears that the curves are stretched slightly to the right as time increases, suggesting an increase in bandwidth \( \gamma \).
Figure 3: 30-year sliding window MLEs for GEV location. A spatially-varying linear trend in time appears to be appropriate.

4 Results

Analysis is based on 10,000 MCMC iterations that remained after 10,000 samples from a burn-in period were discarded. Convergence was assessed by examining trace plots of several representative parameters. To assess model fit, we follow Davison et al. (2012) and examine QQ-plots of various empirical versus modeled groupwise statistics for a set of 15 evaluation sites. The first row of Figure 5 shows pairwise maxima for pairs of sites situated 100km (left), 300km (center), and 800km (right) apart. The pairwise maxima show good agreement at all spatial lags. The second row shows groupwise minima (left), means (center), and maxima (right) for observed and modeled yearly maxima at 5 randomly-chosen sites. This
Figure 4: Empirical F-madogram estimates of extremal coefficient, using 5,000 bins. There is strong evidence of extremal dependence, as well as a suggestion that the spatial range of dependence has increased with time. The panel for 2010, curiously, seems to indicate extremal dependence (extremal coefficient < 2.0) even at very large spatial lags.

also shows good agreement, with the fit breaking down somewhat for the left-hand column, the groupwise minima of maxima. The third row shows groupwise minima (left), means (center), and maxima (right) at all 15 evaluation sites. Here, with the exception of the maxima of maxima, model fit does not fit as well. The model seems not to accurately represent high-order dependencies. This could be due to inadequate flexibility of the spatial model, or it could be that dependencies exist that are a function something other than space.

A useful place to begin interrogating the posterior sample is the question of whether the structure of the field of extreme temperatures in Europe has changed from 1900 to 2011.
Figure 5: Model diagnostics for max-stable model using 15 evaluation sites. The top row shows pairwise maxima for sites 100km (left), 300km (center), and 800km (right) apart. The middle row shows the groupwise minima (left), mean (center), and maxima (right) for 5 randomly-selected sites. The bottom row shows groupwise minima (left), mean (center), and maxima (right) for all 15 sites. As in Davison et al. (2012), margins have been transformed to unit Gumbel for clarity.

Changes in time are represented in the model as linear trends. The two parameters for which linear trends were specified are the GEV location field $\mu(s,t)$, whose trend is allowed to vary spatially, and the log of the kernel standard deviation $\gamma(t)$, which determines the range of spatial dependence in the extreme temperature fields. The linear trend of $\mu(s,t)$ induces a corresponding trend in $m$-year return level. The $m$-year return level is defined as the value one would expect to see on average once every $m$ years, and is a quantity of considerable interest in extreme value analysis. The posterior mean decadal change in 50-year return level
is plotted in Figure 6(a).

From Figure 6(a) we can see that the distribution of extreme high temperatures seems to have shifted to the right (indicating warmer temperatures) in central Europe and Italy, while the distribution seems to have shifted to the left in Denmark, large swaths of Eastern Europe, and small pockets of Western Europe. Figure 6(b) shows that in much of Western and Central Europe, the posterior probability that the 50-year return level has increased is near 1.

Figure 6: Time trends. Panel (a) shows the posterior mean of the decadal change in 50-year return level, in degrees C. Panel (B) shows the posterior probability that the 50-year return level has increased from 1900–2011. Panel (c) shows the time trend of the scaling parameter $\gamma$, which determines the range, in km, of spatial dependence in the extreme temperature field. The posterior mean is shown as a dashed line and pointwise 90\% credible intervals are shown as a gray region.
Figure 6(c) shows the posterior mean and pointwise 90% credible intervals of the time trend in the kernel standard deviation parameter $\gamma$. The spatial dependence range parameter is about 150km at the beginning of the 20th century and increases to about 175km by 2011. The interpretation is that spatial extent of extreme high temperature events seems to have expanded over the past century.

The heart of our risk analysis is summarized in Figures 7 and 8. Each panel in Figure 7 represents a temperature threshold. For each threshold $T$, the fraction of cropland in exceedance, $p_T$, is plotted as a function of year. Each boxplot summarizes the posterior distribution of $p_T$, the fraction of cropland exceeding threshold $T$ during each year. These boxplots are derived from Kriging-type predictive distributions for each year, computed using (3). The dashed lines and gray regions are the posterior mean and 90% pointwise credible intervals for $p_T$, computed from the climatology-type predictive draws. In this representation, the boxplots represent predictive distributions of what the model estimates actually happened in each year, while the dashed lines and gray regions represent the distribution of what the model expected to have happened. For each threshold shown, the fraction of cropland in exceedance gradually increases throughout the century, but the credible intervals show that this increase small compared to year-to-year variation. In addition, the boxplots produced from the Kriging-type predictive distributions show noticeable nonlinearity in time, seeming to indicate that our model for temporal evolution on which the Kriging-type predictive distribution depends might not be flexible enough. A possible remedy is to use a quadratic or spline-based time trend for GEV location. This apparent nonlinearity warrants further study.

Another way of summarizing temporal changes in risk is shown in Figure 8. Filled boxplots summarize climatology-type predictive distributions for the percentage of cropland exceeding several different thresholds. These predictive distributions are shown for three years: 1900, 1955, and 2011. For the threshold of 37°C, for example, the median percentage of
Figure 7: Posterior summary of $p_T(t)$, the fraction of cropland that exceeds the given threshold $T$ as a function of year $t$. Boxplots for each year summarize the posterior Kriging-type predictive distribution of $p_T(t)$. The posterior mean of the climatology-type predictive distribution of $p_T(t)$ is shown as a dashed line, and pointwise 90% credible intervals are shown as a gray region.
Figure 8: Fraction of cropland that exceeds a given threshold (in °C) as a function of threshold, for three different years. Boxplots summarize the climatology-type predictive distributions for each threshold, for three years. Solid boxplots are from the HKEVP, and dashed boxplots are from the Gaussian model (with whiskers omitted to decrease visual clutter).
cropland in exceedance was about 38% in 1900, 40% in 1955, and 42% in 2011. Just as we saw in Figure 7, Figure 8 shows, for each temperature threshold, a slight but inconclusive increase in agricultural land in exceedance throughout the century. Although it is not recommended for extreme value analysis, we also fit a spatially-varying Gaussian model for comparison. This model is similar to the model for $\mu(s)$ in Section 2.1, but treats the spatially-varying coefficient model as a model for the data itself (see Section S-2 in the online supplement for details). Results from the Gaussian model are shown as dashed boxplots in Figure 8. This model also shows an increase across years, with exceedance fractions systematically lower than those reported by the max-stable model. It is instructive that the models diverge more at higher temperatures, revealing the deficiency of the Gaussian model at fitting the far tails of the distribution.

5 Discussion

We have used a spatial extreme value model to examine how the risk of widespread extreme high temperatures, which have been linked to crop failure, has changed across European agricultural land from 1900 to 2011. Our results show that the risk of large percentages of cropland exceeding a high temperature threshold has probably increased in the last century, but only slightly so.

Our analysis excluded the possibility that the GEV shape parameter $\xi$ has changed in time. This modeling choice was made because exploratory analysis showed no evidence of a temporal trend. However, if $\xi$ has in fact increased, this could have a large impact on the frequency of exceeding biologically-important temperature thresholds. Further investigation is necessary to explore the possible dynamics of this important parameter.

Another modeling aspect that warrants further consideration is the linear time trend of GEV location. Although this linear trend varies spatially, the boxplots of Kriging-type
predictive distributions (Figure 8) seem to call for either additional flexibility in the temporal structure or restriction of the analysis to a shorter time window. Furthermore, the boxplots in Figure 8 do not seem to register important extreme heat events like the European heat wave of 2003. Examination of predictive surfaces for this year do show extremely high temperatures, concentrated mostly in France. Because such a large fraction of the total prediction points are located in Eastern Europe, which was relatively cool in 2003, the catastrophic heat wave in France did not register greatly in the aggregate summaries. Perhaps a more detailed country-by-country analysis would provide additional insight.

One limitation of the max-stable approach is that it only considers annual high temperatures. Thus, it does not enable one to take duration of high temperature events into account. There is some suggestion that the deleterious effects of very high temperatures are cumulative, and some work has been done to estimate this effect. For example, Schlenker and Roberts (2009) regress crop yields on heat input, with heat input measured using degree-days, so time of exposure is taken into account. A potential way to finesse this limitation would be to define the block maximum as the largest $m$-day average temperature of the year, where $m$ is something like 5, rather than the largest daily high temperature as we have done here. This approach would consider duration of high temperatures to some degree, but it would risk limiting the applicability of the max-stable asymptotics by inducing additional temporal dependence in the data. Beyond considering multi-day averages, a new extreme value approach that considers the duration in addition to the severity of high temperature events would be highly desirable.

Additional information and supplementary material for this article is available online at the journal’s website.
Acknowledgements

The authors would like to thank David Lobell for helpful critiques, guidance on framing the question, and especially for introducing us to the issue of susceptibility of crops to extreme temperatures.

References


