Spatial Extremes: the Smith Model

Robert L. Wolpert
Department of Statistical Science
Duke University, Durham, NC, USA

May 5, 2017

1 Introduction

2 Model

We are concerned with the so-called "Smith model" (Smith, 1990), one of the earliest max-stable models proposed for studying dependent spatial extremes. Let $\nu_u(dv)$ be a σ-finite positive measure on $\mathbb{R}_+$ and let $\mathcal{S} = \mathbb{R}^2$ be "space", with elements denoted s or σ (for later space-time applications we may need to consider a product space $\Omega = \mathcal{S} \times \mathcal{T}$ and, for other applications, spaces that are larger or more abstract). Let $\mathcal{N}(dv\, d\sigma) \sim \mathsf{Po}\big(\nu(dv\, d\sigma)\big)$ be a Poisson random measure on the Borel sets of $\mathbb{R}_+ \times \mathcal{S}$ with product mean measure

$$\nu(dv\, d\sigma) := \nu_u(dv)\, d\sigma$$

Fix a kernel function $k(s;\sigma) \geq 0$ on $\mathcal{S}^2$ and define a stochastic process indexed by $s \in \mathcal{S}$ by

$$X_s := \sup_j v_j\, k(s;\, \sigma_j),$$

where $\{(v_j, \sigma_j)\} \subset \mathbb{R}_+ \times \mathcal{S}$ comprise the random countably-infinite support of $\mathcal{N}$. The CDF for X_s is

$$\begin{aligned} \mathsf{P}[X_s \leq x] &= \mathsf{P}[\text{ No Poisson points in } A\ ] \\ &= \exp\big(-\nu(A)\big), \text{ where} \\ A &:= \{(v, \sigma):\ v\, k(s;\, \sigma) > x\}. \end{aligned}$$

If $\nu(A) < \infty$ then X_s will be finite almost-surely; if $k(s;\sigma) = k(s - \sigma)$ depends only on the difference $s - \sigma$, then X_s will be stationary, and isotropic if $k(s;\sigma) = k\big(|s-\sigma|\big)$ depends only on the Euclidean distance. We use the same notation $k(\cdot)$ for any of these three functions, distinguishing them by their argument(s). Below we will explore the marginal distribution of X_s and the joint distribution of $\{X_{s_i}\}$ for a finite set $S = \{s_i\} \subset \mathcal{S}$.

2.1 Dependence

First consider the value of X_s in one-dimensional setting, with $\mathcal{S} = [0, 1]$ the unit interval. Figure (1) illustrates that the value of X_s is, by definition, the maximum value of $\upsilon_j k(s-\sigma_j)$. But, X_s is also the smallest value y such that

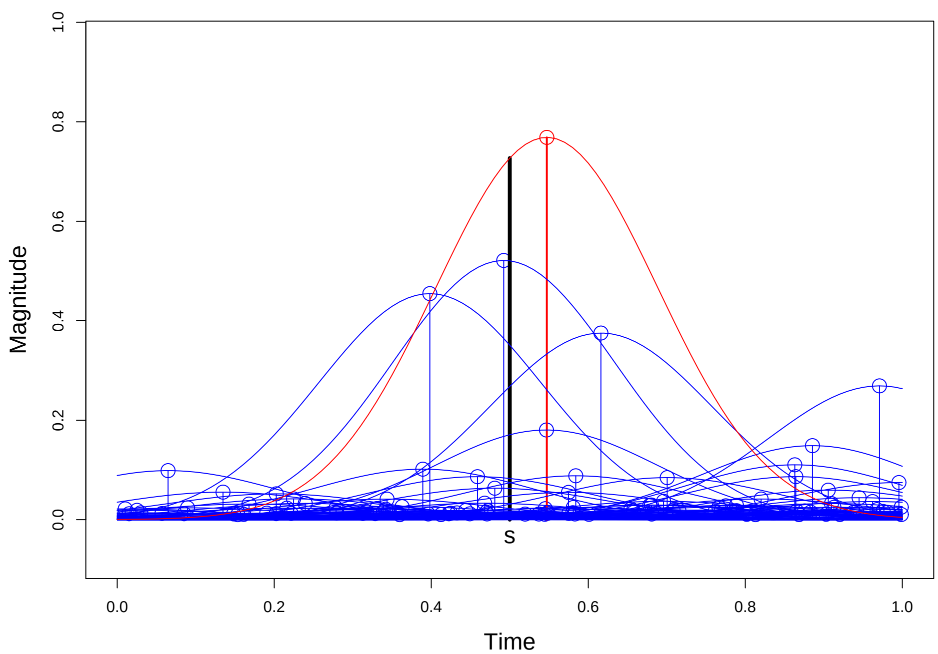

Figure 1: X_s is the maximum value of $\upsilon_j k(s-\sigma_j)$

some (σ_j, υ_j) lies in $A := \{(\sigma, \upsilon):\ \upsilon > y/k(s-\sigma)\}$, as shown in Figure (2). It follows that the probability $\mathsf{P}[X_s \le y]$ is the probability $\exp(-\nu(A))$ that $\mathcal{N}(A) = \emptyset$.

In this example $\nu_u(d\upsilon) = \upsilon^{-2}\mathbf{1}_{\{\upsilon>0\}}d\upsilon$ and $\int_{\mathcal{S}} k(s)\,ds = 1$. It follows that each X_s has the unit Fréchet distribution with CDF $\mathsf{P}[X_s \le y] = \exp(-1/y)$,

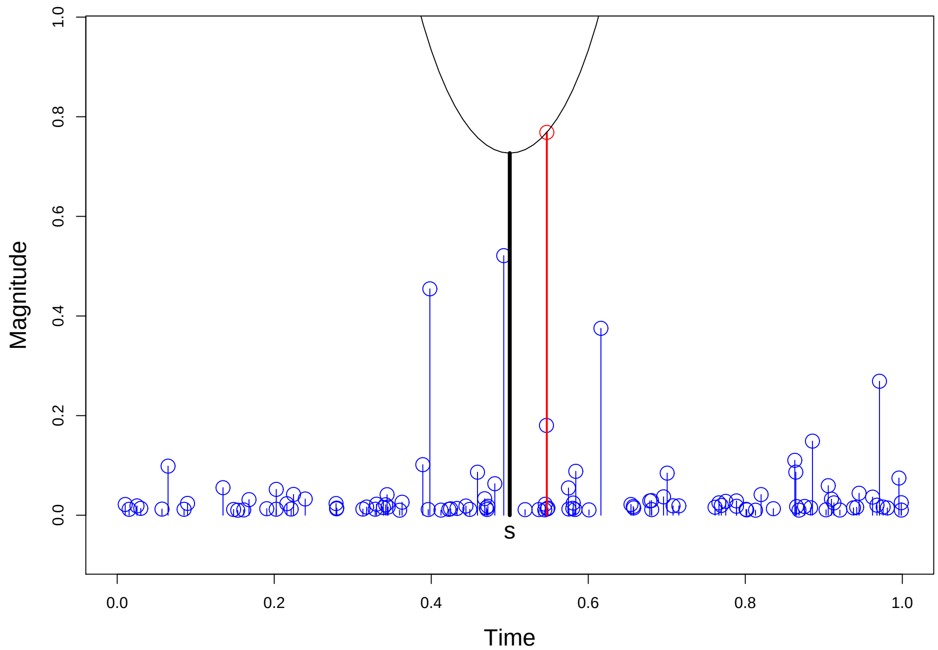

Figure 2: X_s is also the smallest value y such that some (σ_j, υ_j) lies in $\{(\sigma, \upsilon) : \ \upsilon > y/k(s - \sigma)\}$

a distribution with infinite mean and variance. The values X_{s_1} and X_{s_2} of X_s at two different locations won't be independent, but the concepts of "correlation" or "covariance" aren't adequate to describe their dependence since they're not in L_2. If $\text{dist}(s_1, s_2)$ is large then they will be *nearly* independent, however, while if $\text{dist}(s_1, s_2)$ is small they won't be. To see this geometrically, compare Figures (3) and (4). These are different realizations of the same random field, evaluated at $s_1 = 0.5$ and $s_2 = 0.75$. In Figure (3) the mass points $\{(\sigma_j, \upsilon_j)\}$, $\{(\sigma_{j'}, \upsilon_{j'})\}$ at which the maxima are attained differ, the usual case when s_1, s_2 are distant; in Figure (4), the maxima are both attained at a single point (σ_j, υ_j), the usual case when s_1, s_2 are close. The point is illustrated in Figures (5) and (6) in a way that offers an avenue to computing the joint CDF $\mathsf{P}[X_{s_1} \le y_1, \ X_{s_2} \le y_2]$.

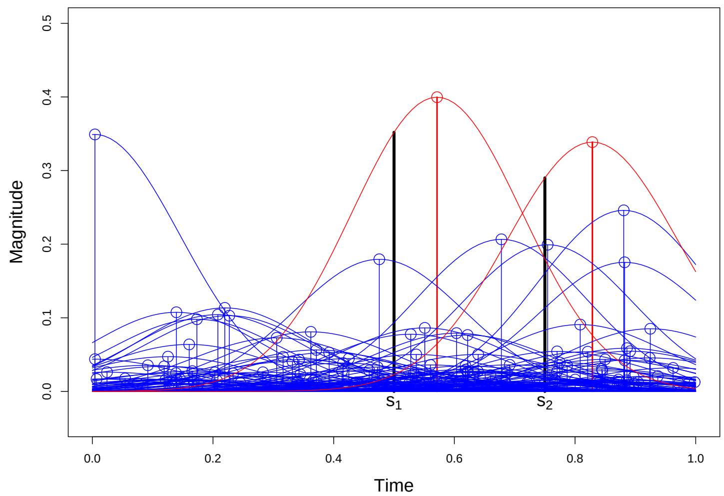

Figure 3: The maximum values of $\upsilon_j k(s_1 - \sigma_j)$ and $\upsilon_{j'} k(s_2 - \sigma_{j'})$ with different extremal (σ_j, υ_j), $(\sigma_{j'}, \upsilon_{j'})$

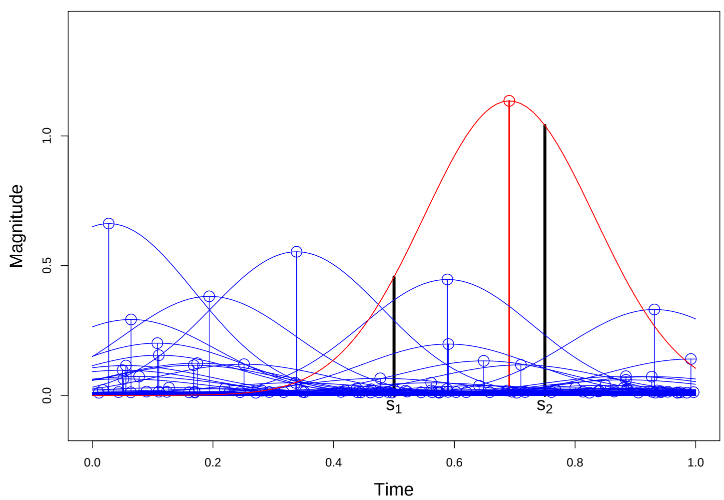

Figure 4: The maximum values of $\upsilon_j k(s_1 - \sigma_j)$ and $\upsilon_j k(s_2 - \sigma_j)$ with the same extremal (σ_j, υ_j)

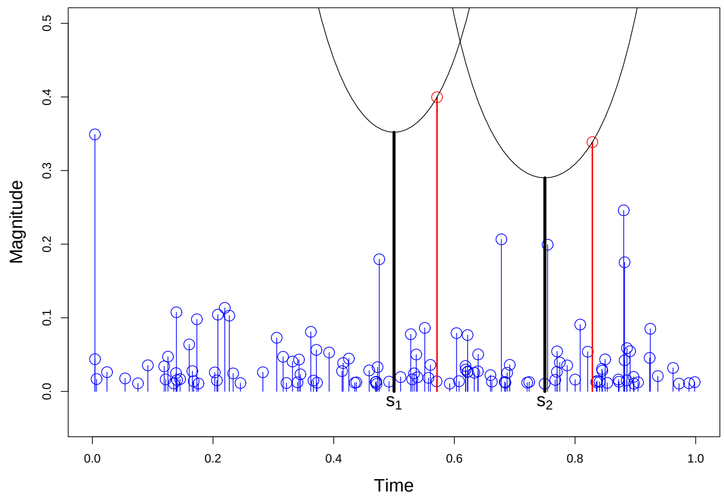

Figure 5: The maximum values of $\upsilon_j k(s_1 - \sigma_j)$ and $\upsilon_{j'} k(s_2 - \sigma_{j'})$ with different extremal (σ_j, υ_j), $(\sigma_{j'}, \upsilon_{j'})$

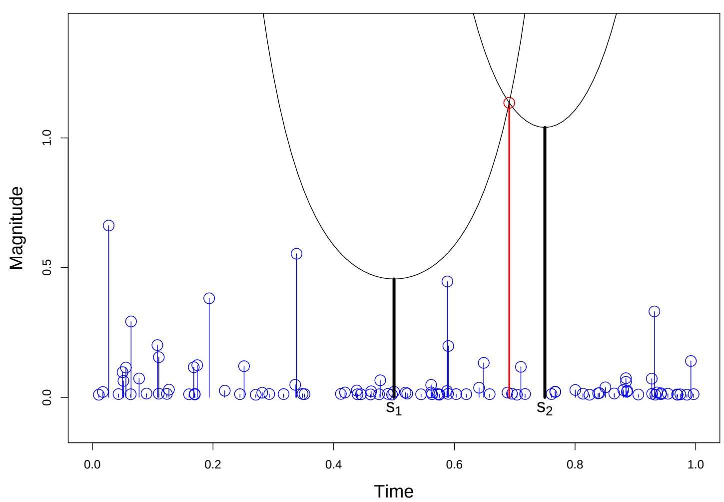

Figure 6: The maximum values of $k(s_1 - \sigma_j)$ and $k(s_2 - \sigma_j)$ with the same extremal (σ_j, υ_j)

3 Joint Distribution at Two Points

Let $k(s, \sigma)$ be isotropic and monotone decreasing in the Euclidean distance $|s - \sigma|$. We will illustrate in $d = 1$ dimension using the specific choice of $k(s, \sigma) = \exp(-\lambda(s - \sigma)^2)$, but the same approach works more generally. For any two locations s_1, s_2 and levels y_1, y_2 the joint CDF

$$F(y_1, y_2) = \mathsf{P}\big[X_{s_1} \le y_1,\ X_{s_2} \le y_2\big]$$

can be evaluated as $\exp(-\nu(A))$ for the union A of the sets $B_i := \{(v, \sigma) : vk(s_i, \sigma) > y_i\}$ or, equivalently, of the disjoint sets $A_1 := B_1 \cap \{\sigma < \tilde{\sigma}\}$ and $A_2 := B_2 \cap \{\sigma > \tilde{\sigma}\}$ to the left and right of the red line in Figure (7) at $\sigma = \tilde{\sigma}$, respectively, where

$$\tilde{\sigma} = (s_1 + s_2)/2 + \log(y_2/y_1)/\lambda(s_1 - s_2).$$

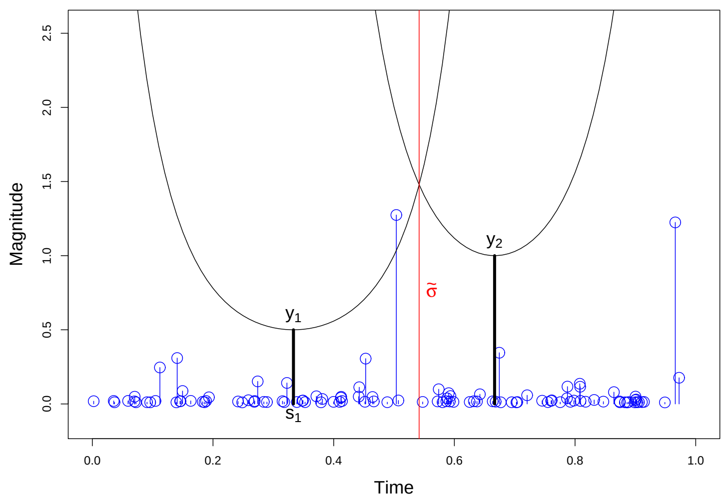

Figure 7: The joint CDF of X_s at s_1 and s_2, evaluated at y_1 and y_2, is $\exp(-\nu(A))$ for the set A of points (v, σ) above the two black curves. We evaluate $\nu(A)$ as the sum $\nu(A_1) + \nu(A_2)$ of the portions to the left and right of the red line at $\sigma = \tilde{\sigma}$, where $y_1/k(s_1, \tilde{\sigma}) = y_2/k(s_2, \tilde{\sigma})$.

In one dimension with $s_1 < s_2$, these measures are

$$\begin{aligned}
\nu(A_1) &= \int_{-\infty}^{\tilde{\sigma}} \int_{y_1/k(s_1,\sigma)}^{\infty} \gamma\alpha v^{-\alpha-1}\, dv\, d\sigma \\
&= \gamma {y_1}^{-\alpha} \int_{-\infty}^{\tilde{\sigma}} k^{\alpha}(\sigma - s_1) d\sigma \\
&= \gamma {y_1}^{-\alpha} \int_{-\infty}^{\tilde{\sigma}} \exp(-\lambda\alpha(\sigma - s_1)^2)\, d\sigma \\
&= \frac{\gamma\sqrt{2\pi}}{\sqrt{\lambda\alpha}} {y_1}^{-\alpha} \Phi\big(\sqrt{\lambda\alpha}(\frac{s_1}{2} + \log(y_2/y_1)/\lambda s_1)\big) \\
\nu(A_2) &= \frac{\gamma\sqrt{2\pi}}{\sqrt{\lambda\alpha}} {y_2}^{-\alpha} \Phi\big(\sqrt{\lambda\alpha}(\frac{s_2}{2} + \log(y_1/y_2)/\lambda s_2)\big)
\end{aligned}$$

leading to closed-form expressions for the joint CDF $F(y_1, y_2) = \exp\big(-\nu(A_1) - \nu(A_2)\big)$ and hence the joint pdf (and likelihood function) $f(y_1, y_2)$.

4 Convex Regions and Tessellations

Let $C \subset \mathcal{S}$ be a convex polygon and define

$$\begin{aligned}
X_C &:= \sup_{s \in C} X_s \\
&= \sup_{j \in \mathbb{N}, s \in C} v_j\, k(s;\, \sigma_j); \qquad (1)
\end{aligned}$$

we will also address the distribution of X_C and the joint distribution of $\{X_{C_i}\}$ for disjoint collections of convex sets, like a rectangular grid or hexagonal tiling or a triangulation of a spatial area.

4.1 Example: Fréchet/Square Exponential

Fix $\mathcal{S} = \mathbb{R}^2$ and numbers $\alpha > 0$, $\gamma > 0$, and $\lambda > 0$ and consider the specific example

$$k(s; \sigma) := e^{-\lambda|s-\sigma|^2/2}$$

where $|s-\sigma|$ denotes the Euclidean distance separating s and σ (the so-called *square exponential* kernel), and

$$\nu(dv\, d\sigma) := \gamma\alpha v^{-\alpha-1} \mathbf{1}_{\{v>0\}}\, dv\, d^2\sigma$$

where $d^2\sigma$ denotes Lebesgue measure in $\mathcal{S}$. For $0 < \alpha < 2$, this is the Lévy measure for the stationary α-Stable random field. For this kernel and measure,

$$\begin{aligned}
\nu(A) &= \nu\left\{(\upsilon,\sigma):\ \upsilon > x\exp(\lambda|s-\sigma|^2/2)\right\} \\
&= \iint_{\mathcal{S}} \gamma\left[x\exp(\lambda|s-\sigma|^2/2)\right]^{-\alpha} d^2\sigma \\
&= x^{-\alpha} 2\pi\gamma \int_0^\infty e^{-\alpha\lambda r^2/2}\, r\, dr \\
&= x^{-\alpha}\frac{2\pi\gamma}{\alpha\lambda}, \quad \text{so} \qquad (2)\\
\Pr[X_s \le x] &= \exp\left(-x^{-\alpha} 2\pi\gamma/\alpha\lambda\right), \qquad x > 0 \\
f(x \mid \alpha,\gamma,\lambda) &= (2\pi\gamma/\lambda)\exp\left(-x^{-\alpha}2\pi\gamma/\alpha\lambda\right) x^{-\alpha-1}\mathbf{1}_{\{x>0\}} \qquad (3)
\end{aligned}$$

and X_s has a Fréchet $\mathsf{Fr}\,(\alpha, 2\pi\gamma/\alpha\lambda)$ distribution. For $|s_i - s_j| \gg (\gamma/\alpha\lambda)^{1/\alpha}$ the random variables X_{s_i} will be nearly independent, while for $|s_i - s_j| \ll (\gamma/\alpha\lambda)^{1/\alpha}$ they will nearly coincide. Their pth moments are infinite for $p \ge \alpha$, so in the interesting cases of $0 < \alpha < 2$ the variance is infinite and covariance undefined (and for $\alpha \le 1$ even the mean is infinite).

4.2 Fréchet on Convex Sets

Here we generalize (3) from a single point $s \in \mathcal{S}$ to a convex set $C \subset \mathcal{S}$ and from the squared exponential to a wider class of kernel functions.

Proposition 1. *Let $\alpha > 0$ and let $k(s;\sigma) = k\big(|s-\sigma|\big)$ be an isotropic kernel with finite monotonically-decreasing nonnegative distance function $k(r)$. Set*

$$c_0 := k^\alpha(0) \qquad c_1 := \int_0^\infty k^\alpha(r)\, dr \qquad c_2 := \int_0^\infty k^\alpha(r)\, 2\pi\, r\, dr. \qquad (4)$$

If $k \in L_\alpha(\mathbb{R}^2)$ and if k is bounded then each $c_j < \infty$ and for any convex set $C \subset \mathcal{S}$,

$$X_C^* := \sup_{s\in C} X_s$$

has a Fréchet distribution $X_C^* \sim \mathsf{Fr}\,(\alpha,\ c)$, i.e.,

$$\mathsf{P}[X_C^* \le x] = e^{-c\, x^{-\alpha}}, \qquad x > 0$$

with rate

$$c = \gamma\,[c_2 + c_1 P + c_0 A]$$

for the perimeter P and area A of C. Conversely, if $k \notin L_\alpha(\mathbb{R}^2)$ or if k is unbounded, then $X_C^* = \infty$ almost-surely.

Proof. First consider a rectangular region C; without loss of generality we may rotate and translate so that the lower-left corner is the origin, and write $C = [0, a] \times [0, b]$.

Now from (1) we find

$$\begin{aligned}
\mathsf{P}[X_C^* \le x] &= \exp(-\nu(A)), \text{ where} \\
A &= \big\{(\upsilon, \sigma) :\ \upsilon > x/k\big(|s-\sigma|\big) \text{ for any } s \in C\ \big\} \\
\nu(A) &= \int_{\mathcal{S}} \gamma\,\big[x/k\big(\mathrm{dist}(\sigma,\,C)\big)\big]^{-\alpha}\ d^2\sigma \\
&= x^{-\alpha}\gamma \int_{\mathcal{S}} k^\alpha\big(\mathrm{dist}(\sigma,\,C)\big)\,d^2\sigma \\
&= x^{-\alpha}\,\gamma\,[c_2 + c_1\,P + c_0\,A] \qquad (5)
\end{aligned}$$

where $P = 2(a + b)$ is the perimeter of C and $A = a \cdot b$ the area. To see this, divide the integral over $\mathcal{S} = \mathbb{R}^2$ into the nine regions determined by restricting σ_1 to $(-\infty, 0)$, $[0, a]$, (a, ∞) and by restricting σ_2 to $(-\infty, 0)$, $[0, b]$, (b, ∞). The four "corner" pieces (with both σ_1 and σ_2 unbounded) sum to the first term in (5); the four semi-bounded strips sum to the second term; and the one bounded term (C itself) the last term.

The same argument shows that (5) holds for *any* convex polygon with $n \ge 3$ sides; now partition $\mathcal{S}$ into $2n+1$ regions by drawing two half-lines from each corner orthogonal to the edges that meet there to form n wedges (whose total contribution is the first term), n semi-infinite strips (whose total contribution is the second), and C itself (the third term). Finally, any convex set in $\mathbb{R}^2$ can be written as the increasing union of a sequence of convex polygons whose perimeters and areas converge. ∎

Equation (5) also holds for the limiting cases of line segments (where $A = 0$ and P is twice the segment length) and points (where $A = P = 0$), where for the squared-exponential kernel it reduces to (2).

For any $\alpha > 0$, the moments $\{c_j\}$ for the squared-exponential kernel $k(s;\sigma) := \exp\left(-\lambda(s-\sigma)^2\right)$ are all finite:

$$\begin{aligned} c_0 &:= k^\alpha(0) &&= e^{-0} &&= 1 \\ c_1 &:= \int_0^\infty k^\alpha(r)\,dr &&= \int_0^\infty e^{-\alpha\lambda\,r^2/2}\,dr &&= \sqrt{\frac{\pi}{2\alpha\lambda}} \\ c_2 &:= \int_0^\infty k^\alpha(r)\,2\pi\,r\,dr &&= \int_0^\infty e^{-\alpha\lambda\,r^2/2}\,2\pi\,r\,dr &&= \frac{2\pi}{\alpha\lambda} \end{aligned}$$

and hence (5) holds for all $\alpha > 0$.

Note this gives a way to predict maxima for *regions* if we can infer the values of α, γ, and whatever parameters determine $k(s;\sigma)$ from observations at *points*— for example, we might hope to learn the parameters governing precipitation from data at weather stations, then predict extremes over counties or watersheds.

5 Perfect Simulation

It is possible to simulate $\{X_{s_i}\}_{i\in I}$ perfectly for any specified parameters α, γ, λ and any collection $\{s_i\}$ of locations ("monitoring stations") in $\mathbb{R}^2$, using a variation of the Inverse Lévy Measure algorithm of Wolpert and Ickstadt (1998a,b). Here's how. (Note: recent work by Wang and Stoev (2010) may enable *conditional* sampling given observations, and hence posterior sampling).

Begin with a ball of some radius $R > 0$ containing the set $\{s_i\}_{i\in I}$. Let $\mathcal{R}$ be a ball with the same center but larger radius $(R+\Delta)$ for some number $\Delta > 0$ (we'll choose Δ below— something like $3/\sqrt{\alpha\lambda}$ or so should work). Draw points $\sigma_j \overset{\text{iid}}{\sim} \mathsf{Un}(\mathcal{R})$, and draw the event times τ_j from a standard unit-rate Poisson process (*i.e.*, partial sums of independent standard exponential random variables). Set

$$\upsilon_j := \left[\ \frac{\gamma\ \pi(R+\Delta)^2}{\tau_j}\ \right]^{1/\alpha}.$$

Then $\{(\upsilon_j,\sigma_j)\}$ are distributed exactly as the support of $\mathcal{N}$ restricted to $\mathbb{R}_+\times\mathcal{R}$, sorted in order of decreasing υ_j. With each successive j, reset each

$$x_i = \sup_{k\le j}\left\{\upsilon_k\ e^{-\lambda|s_i-\sigma_k|^2/2}\right\}$$

until finally $\upsilon_j < \underline{x}_I := \min\{x_i\}_{i \in I}$, whereupon stop.

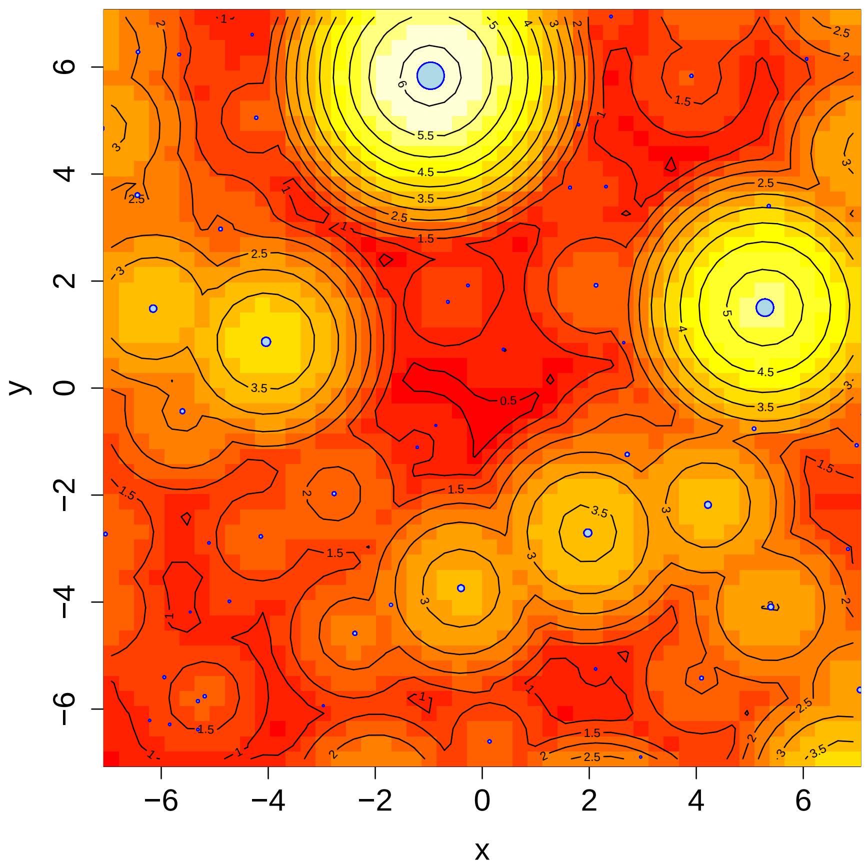

Figure 8: Simulation of X_s at 50×50 grid of points $\{s_i\}$. Blue disks represent support points, with area proportional to magnitude. $J = 484$ draws (υ_j, σ_j) were required to reach $\upsilon_J = 1.1687 < \underline{x}_I = 1.1690$, determining 68 support points each supporting from one to 305 points x_i.

It is possible that $\mathcal{R}$ wasn't large enough— that if we could have drawn points from all of $\mathbb{R}^2$, some very-distant point with huge mass would have increased the value of some x_i. That event is for some support point of $\mathcal{N}$

to lie in the set

$$A = \left\{ (\upsilon, \sigma) : \ \upsilon > \inf_{i \in I} x_i \, e^{\lambda |s_i - \sigma|^2}, \quad \sigma \notin \mathcal{R} \right\},$$

whose probability is $1 - e^{-\nu(A)}$. This is bounded above by $1 - e^{-\nu(A^*)}$ for the set

$$A^* := \left\{ (\upsilon, \sigma) : \ \upsilon > \underline{x}_I \, e^{\lambda \left(|\sigma| - R \right)^2}, |\sigma| \geq R + \Delta \right\}$$

since $A \subset A^*$ (after translating), with

$$\begin{aligned}
\nu(A^*) &= \int_{\mathcal{R}^c} \gamma \left[\underline{x}_I \, e^{\lambda (|\sigma| - R)^2/2} \right]^{-\alpha} \, d\sigma \\
&= 2\pi\gamma \int_{R+\Delta}^{\infty} \exp\left\{ -\alpha\lambda (r - R)^2/2 \right\} \, r \, dr \, (\underline{x}_I)^{-\alpha} \\
&= \frac{\gamma\sqrt{2\pi}}{\alpha\lambda} \left\{ \phi\left(\Delta\sqrt{\alpha\lambda} \right) + R\sqrt{\alpha\lambda} \, \Phi\left(-\Delta\sqrt{\alpha\lambda} \right) \right\} \, (\underline{x}_I)^{-\alpha},
\end{aligned}$$

where $\phi(z)$ and $\Phi(z)$ are the pdf and CDF for the standard Normal distribution.

For sufficiently large Δ the probability $1 - e^{-\nu(A^*)} \leq \nu(A^*)$ will be small enough that $\{x_i\}_{i \in I}$ is acceptably close to (and identical to with high probability) a sample draw from the $\{X_{s_i}\}$ (in examples it's easy to attain $\nu(A^*) < 10^{-10}$). For perfect sampling with no approximation error, simply draw (probably zero) a Poisson $\mathsf{Po}(\nu(A^*))$ number of points from A^* (easy to do, since it's a radial set); remove any points that are outside A; and, if any support points remain, use them to update $\{x_i\}$. Figure (8) shows an example of a simulation, with a 50×50 regular array of points $\{s_i\}$, using $\alpha = \lambda = \gamma = 1$ on the ball of radius $R = 10$ centered at the origin. For this example $\Delta = 4$ and the error bound was $\nu(A^*) = 0.001347$.

6 Inference

OK, here's a speculative but interesting (to me) idea. Suppose we know α, λ, and γ.

6.1 X_s Observed at One Point

Suppose we are able to observe

$$X_s := \sup_j \{\upsilon_j \, k(s;\, \sigma_j)\}$$

exactly, without measurement error. Then the conditional distribution of $x = X_s$ given the location $\sigma_j = \sigma$ of the supporting point (υ_j, σ_j) (the one attaining the supremum above) is:

$$\mathsf{P}[X_s \le x \mid \sigma] = \exp\left\{-\gamma x^{-\alpha} e^{-\alpha\lambda|s-\sigma|^2/2}\right\}$$

so the conditional pdf for X_s given the location is

$$f(x \mid \sigma) = \alpha\gamma x^{-\alpha-1} \exp\left\{-\alpha\lambda|s-\sigma|^2/2 - \gamma x^{-\alpha} e^{-\alpha\lambda|s-\sigma|^2/2}\right\}.$$

By Bayes' rule,

$$f(\sigma \mid X_s = x) = c\ \exp\left\{-\alpha\lambda|s-\sigma|^2/2 - \gamma x^{-\alpha} e^{-\alpha\lambda|s-\sigma|^2/2}\right\}, \qquad (6)$$
$$c = \frac{\alpha\gamma\lambda\ x^{-\alpha}}{2\pi[1 - e^{-\gamma x^{-\alpha}}]}$$

and of course $\upsilon = x \exp(\lambda|s-\sigma|^2/2)$ is determined by x and σ because, geometrically, the support point for X_s lies on the bowl-shaped set

$$\mathcal{M}_1 := \left\{(\upsilon, \sigma) :\ \upsilon = x\ e^{\lambda|s-\sigma|^2/2}\right\}.$$

For large x we have $\sigma \approx \mathsf{No}(s, \frac{1}{\alpha\lambda} I)$, while for small x we have the same asymptotic behavior as $|\sigma - s| \to \infty$ but the density falls off to zero as $\sigma \to s$. The marginal density for X_s (and hence the likelihood for α, γ, λ) is given in (3); together that and (6) determine the joint distribution for $\alpha, \gamma, \lambda, \sigma, \upsilon$ and x for any specified prior distribution.

6.2 Two Points

Suppose we know α, λ, and γ and observe

$$X_{s_1} := \sup_j \{\upsilon_j \, k(s_1;\, \sigma_j)\} = x_1 \quad \text{and} \quad X_{s_2} := \sup_j \{\upsilon_j \, k(s_2;\, \sigma_j)\} = x_2.$$

Then two cases arise— either both suprema are attained at the *same* $j \in \mathbb{N}$ (*i.e.*, at the same element (υ, σ) in the support of $\mathcal{N}$), or not. Any dependence between X_{s_1} and X_{s_2} arises from the possibility of both maxima arising from the same support point— a point necessarily on the manifold

$$\mathcal{M}_{12} := \mathcal{M}_1 \cap \mathcal{M}_2 = \{(\upsilon, \sigma) : \ \upsilon = x_1/k(s_1;\,\sigma) = x_2/k(s_2;\,\sigma)\} \subset \mathbb{R}_+ \times \mathcal{S}.$$

For example, with $k(s;\,\sigma) = \exp\big(-\lambda|s-\sigma|^2\big)$ and $s_1 = (0,0)$, $s_2 = (d,0)$,

$$\mathcal{M}_{12} := \left\{(\upsilon,\sigma) : \ \sigma_x = \frac{d}{2} + \frac{1}{2\lambda d}\log\frac{x_2}{x_1},\ \sigma_y \in \mathbb{R},\ \upsilon = x_1\, e^{\lambda|s_1-\sigma|^2/2}\right\}, \quad (7)$$

whose projection $\mathcal{M}_{12}^{\mathcal{S}}$ onto $\mathcal{S}$ is a line perpendicular to the segment connecting s_1 and s_2, offset from the midpoint toward the point with the larger value of X_s by a distance that shrinks as λ or $d := |s_1 - s_2|$ grows.

If X_{s_1} and X_{s_2} do *not* share a common support point, then each must have its own point in the appropriate component of the set $\mathcal{M}_{1\wedge 2} := \mathcal{M} \setminus \mathcal{M}_{12}$, where

$$\mathcal{M} := \left\{(\upsilon,\sigma) : \ \upsilon = \min\left\{x_1\, e^{\lambda|s_1-\sigma|^2},\ x_2\, e^{\lambda|s_2-\sigma|^2}\right\}\right\}.$$

Notice the projection $\mathcal{M}_{1\wedge 2}^{\mathcal{S}}$ of $\mathcal{M}_{1\wedge 2}$ onto $\mathcal{S}$ partitions $\mathcal{S}$ into two parts, separated by line $\mathcal{M}_{12}^{\mathcal{S}}$ described below (7).

SO— a posteriori, the point(s) (υ, σ) leading to the observed x_1, x_2 are *either* a pair of points, one in each component of $\mathcal{M}_{1\wedge 2}$, *or* a single point in $\mathcal{M}_{12}$. The likelihood might be available for each of these; whether it is or not, it may be possible to draw MCMC samples of the support point(s) and parameters. I don't know yet how to find, *e.g.*, the probability of one support point *vs.* two, or the distribution of the latent support points, or the likelihood function.

One idea: Use (6) to get cond'l dist'n of σ, υ for the support point for X_{s_1}... then for each υ and σ, find the probability that this is ALSO the support point for X_{s_2} as $e^{-\nu(B)}$ where B is the set of $\{(\upsilon,\sigma)\}$ that are *not* in the (already known to be empty) bowl above s_1, but *are* in the cone above s_2.

6.3 Three Points

Now suppose we observe $x_i = X_{s_i}$ at three locations s_1, s_2, s_3. Now there are three possibilities— either each supremum X_s arises from a different point $(\upsilon,\sigma) \in \mathbb{R}_+ \times \mathcal{S}$ in the support of $\mathcal{N}$; or one of the three pairs s_i, s_j

share a single support point in one of three manifolds $\mathcal{M}_{ij}$ similar to that of (7); or all three points share the same support point, necessarily at the intersection $\mathcal{M}_{123} = \cap \mathcal{M}_i$ of the three manifolds.

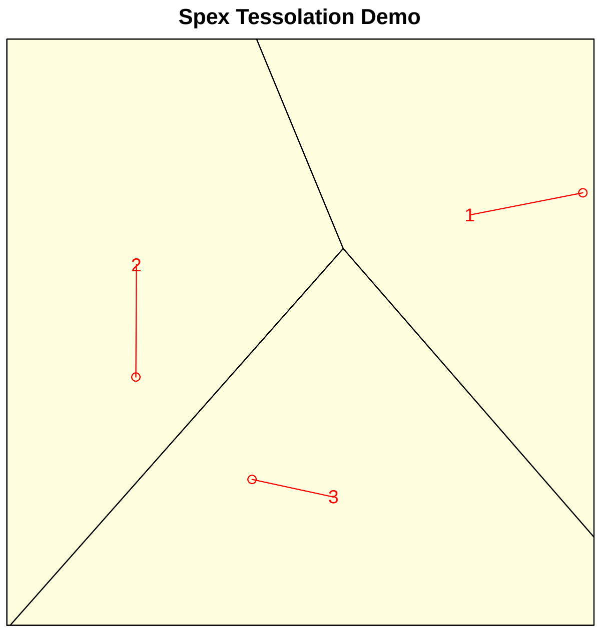

Figure 9: Example of $\{\mathcal{M}_i\}$ for $n = 3$ points s_i. Points s_i are indicated by small circles, connected by line segments to a label i at the center of $\mathcal{M}_i$.

Geometrically, the three lines $\mathcal{M}_{12}^{\mathcal{S}}$, $\mathcal{M}_{13}^{\mathcal{S}}$, and $\mathcal{M}_{23}^{\mathcal{S}}$ all intersect in a single point $\mathcal{M}_{123}^{\mathcal{S}}$. Each pair $\mathcal{M}_{ij}^{\mathcal{S}}$, $\mathcal{M}_{ik}^{\mathcal{S}}$ bounds a wedge-shaped region, the intersection of two half-spaces, which often (but not always, if X_{s_i} is far from X_{s_j} or X_{s_k}) contains the associated point s_i. Those three wedges partition $\mathcal{S}$. The posterior distribution for the support of $\{X_{s_i}\}$ is concentrated on three sets: either each open wedge contains a single point (υ_i, σ_i); or one of the three open wedges contains such a point supporting one of the three X_{s_i}, and the half-line bounding the other two wedges contains a point supporting the other two; or all three points are supported by a single point at $\sigma = \mathcal{M}_{123}^{\mathcal{S}}$.

Again it's possible that the likelihood function would be available. And, again, even if the likelihood is unavailable it may be possible to draw MCMC samples; at worst, we could introduce a measurement-error model and draw

MCMC samples of *all* the parameters *and* the support points $\{(\upsilon_j, \sigma_j)\}$.

6.4 Four Points

Something interesting happens beginning with $n = 4$ points $\{x_i\}$: different topological possibilities arise. Two are shown in Figure (10)— one with a bounded component, and one without. A third possibility is for four lines $\mathcal{M}_{ij}$ to intersect in a single point $\mathcal{M}_{ijk\ell}$; in this latter case the likelihood will be concentrated on a single support point for all four X_{s_i}.

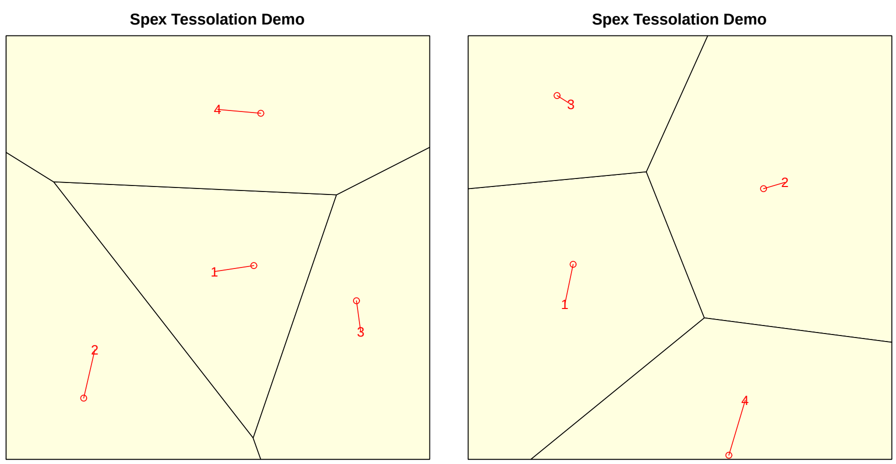

Figure 10: Two possible patterns for $n = 4$ points s_i, one with a bounded component and one without.

6.5 n Points

The same approach can be explored for n points $\{s_i\}$, perhaps with RJMCMC, but the combinatorial problems in evaluating a likelihood will become unwieldy quickly as n grows. Given the observed values $x_i = X_{s_i}$ the support point(s) must lie on the lower boundary of a collection of bowls, the manifold

$$\mathcal{M} := \left\{ (\upsilon, \sigma) : \ \upsilon = \min_i \left\{ x_i \, e^{\lambda |\sigma - s_i|^2/2} \right\} \right\}.$$

Each s_i lies in a (possibly unbounded) polygonal region formed by the intersection of all the half-spaces $\mathcal{M}_{ij}^{\mathcal{S}}$. Those polygons and their boundaries

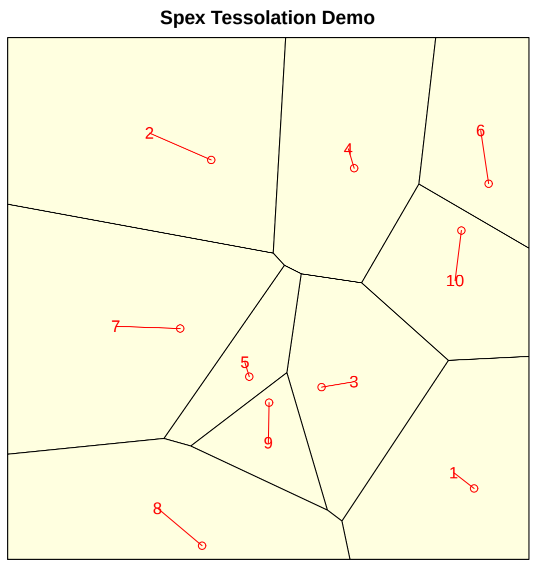

Figure 11: Example of $\{\mathcal{M}_i\}$ for $n = 10$ points s_i.

are the only places where support points can possibly lie. Any polygon corner where four or more edges intersect *must* be a support point for all the associated $\{s_i\}$ whose polygons meet at that corner. Almost surely, no polygon will have more than one such corner. The combinatorial challenge comes from adjacent polygons without such a corner— which might and might not share a support point, necessarily on the edge they share.

If we *don't* know α, λ, and γ then the likelihood surface will be even more complicated for $n \geq 3$, since the topology of the possible support sets will vary with the values of α and λ, but a measurement-error approach (using Metropolis/Hastings MCMC) might work.

7 Extensions

As yet undeveloped but (to me) promising ideas include:

1. Let $\mathcal{S}$ be $\mathbb{R}^p$ instead of $\mathbb{R}^2$;

2. Replace space $\mathcal{S}$ with space-time $\mathcal{S} \times \mathcal{T}$;

3. Replace λ with a positive-definite matrix Λ, so
$$k(s;\sigma) := e^{-(s-\sigma)'\Lambda(s-\sigma)/2}$$

4. Add "attributes" to points $\sigma \in \mathcal{S}$— altitude, aspect, *etc.* Even λ (or Λ) could be attributes, hence locus-specific;

5. Replace Fréchet marginal distribution with Gumbel, by replacing density $\nu_u(u) = \gamma\alpha u^{-\alpha-1}$ on $\mathbb{R}_+$ with $\nu_u(u) = \gamma\alpha e^{-\alpha u}$ on all of $\mathbb{R}$.

6. Replace *points* $\sigma \in \mathcal{S} = \mathbb{R}^p$ with translated subspaces $\ell \in \mathcal{S} = \mathbb{R}^{p-k} \times \mathfrak{g}_{p,k}$, where $\mathfrak{g}_{p,k}$ is the Grassmann manifold of k-dimensional subspaces of $\mathbb{R}^p$ (*e.g.*, lines through the origin if $k = 1$); we build a Poisson field on $\mathbb{R}_+ \times \mathbb{R}^{p-k} \times \mathfrak{g}_{p,k}$ and construct X_s from it.

8 End Notes

- If $k(s;\sigma) := \exp(-\lambda(s-\sigma)^2/2)$ is replaced with the t density $k(s;\sigma) := (1 + (s-\sigma)^2/\nu)^{-(p+\nu)/2}$ in $\mathcal{S} = \mathbb{R}^p$ rescaled to have maximum value one, (2) changes only slightly to
$$\nu(A) = x^{-\alpha}\ \gamma(\pi\nu)^{p/2}\Gamma\Big(\frac{\alpha(p+\nu)-p}{2}\Big)\Big/\Gamma\big(\alpha(p+\nu)\big),$$
so again X_s has an α-Fréchet distribution with intensity available in closed form, and each c_j of (4) is available explicitly.

References

Smith, R. L. (1990), "Max-stable processes and spatial extremes," Online at `http://www.stat.unc.edu/postscript/rs/spatex.pdf`.

Wang, Y. and Stoev, S. A. (2010), "Conditional Sampling for Max-Stable Random Fields," Technical Report 509, Department of Statistics, the University of Michigan, Ann Arbor, MI, USA.

Wolpert, R. L. and Ickstadt, K. (1998a), "Poisson/gamma random field models for spatial statistics," *Biometrika*, 85, 251–267.

Wolpert, R. L. and Ickstadt, K. (1998b), "Simulation of Lévy Random Fields," in *Practical Nonparametric and Semiparametric Bayesian Statistics*, eds. D. K. Dey, P. Müller, and D. Sinha, New York, NY: Springer-Verlag, *Lecture Notes in Statistics*, volume 133, pp. 227–242.