The Extremogram in Space (and Time):

Richard A. Davis*
Columbia University

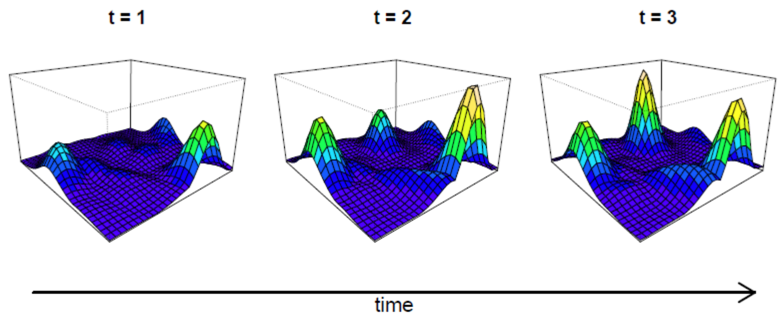
Collaborators:
Yongbum Cho, Columbia University
Souvik Ghosh, LinkedIn
Claudia Klüppelberg, Technical University of Munich
Christina Steinkohl, Technical University of Munich

* Support of the Villum Kann Rasmussen Foundation is gratefully acknowledged.

Plan

- Extremogram in space
  - lattice vs continuous space
- Estimating extremogram—random pattern
- Limit theory for empirical extremogram
- Simulation examples

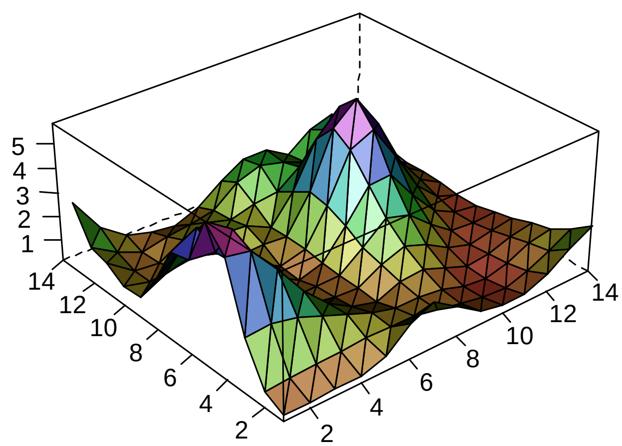
Extremal Dependence in Space and Time



Space-time domain: $\{(\boldsymbol{s}, t) \in \mathbb{R}^d \times [0, \infty)\}$

Extremogram in Space

Setup: Let $X(s)$ be a stationary (isotropic?) spatial process defined on $s \in \mathbb{R}^2$ (or on a regular lattice $s \in \mathbb{Z}^2$).

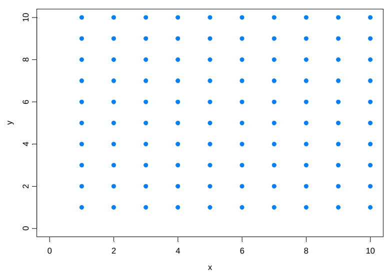


Lattice vs cont space

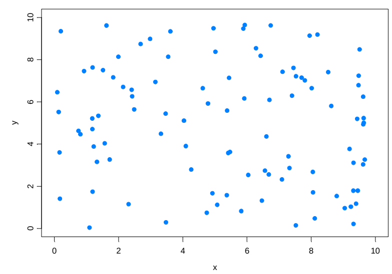
Setup: Let $X(s)$ be a RV stationary (isotropic?) spatial process defined on $s \in \mathbb{R}^2$ (or on a regular lattice $s \in \mathbb{Z}^2$). Consider the former—latter is more straightforward.

$$\rho_{A,B}(h) = \lim_{x \to \infty} P(X(s+h) \,\epsilon\, xB \mid X(s) \,\epsilon\, xA), \qquad h \in \mathbb{R}^2$$

regular grid

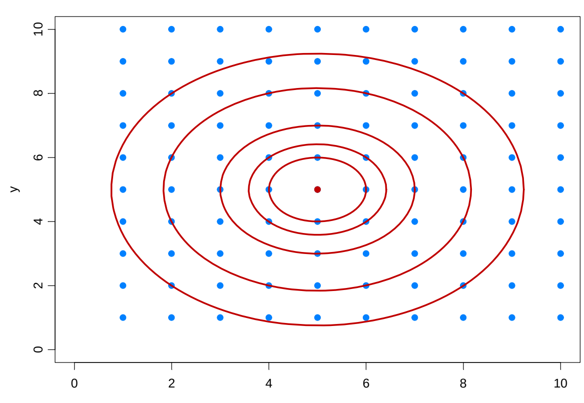


random pattern



Regular grid

regular grid



h = 1; # of pairs = 4

h = sqrt(2); # of pairs = 4

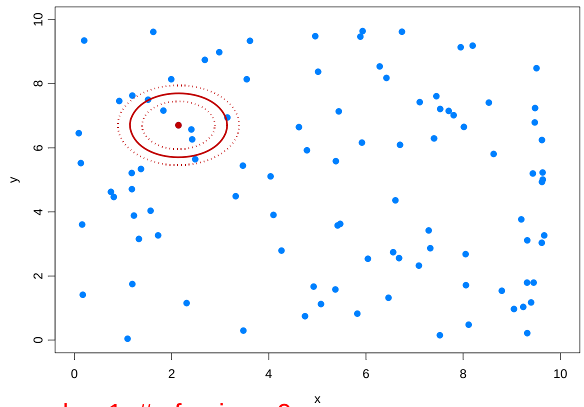
h = 2; # of pairs = 4

h = sqrt(10); # of pairs = 8

h = sqrt(18); # of pairs = 4

Random pattern

random pattern

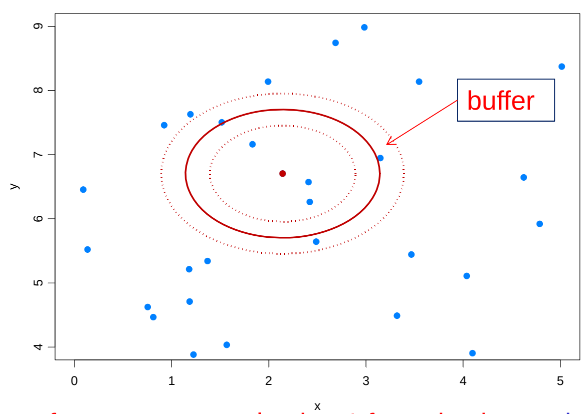


h = 1; # of pairs = 0

h = 1 $\pm$ .25

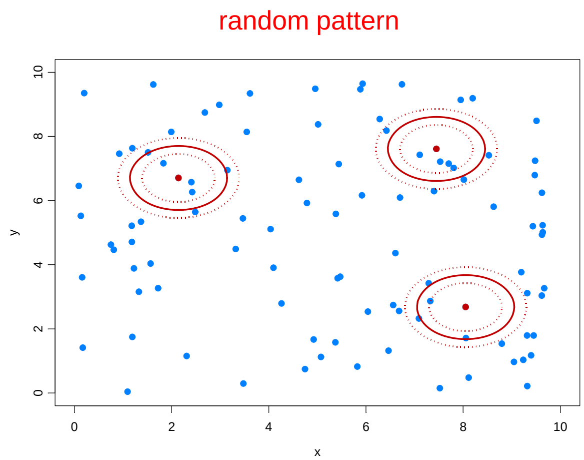
Random pattern

Zoom-in



Estimate of extremogram at lag h = 1 for red point: weight "indicators of points" in the buffer.
Bandwidth: half the width of the buffer.

Random pattern



Note:

- Expanding domain asymptotics: domain is getting bigger.
- Not infill asymptotics: insert more points in fixed domain.

Estimating extremogram--random pattern

Setup: Suppose we have observations, $X(s_1), \ldots, X(s_N)$ at locations $s_1, \ldots, s_{N_n}$ of some Poisson process N with rate ν in a domain $S_n \uparrow \mathbb{R}^2$.

Here, $N_n = N(S_n) =$ number of Poisson points in S_n, $N_n \sim Pois(\nu|S_n|)$.

Weight function $w_n(x)$: Let $w(\cdot)$ be a bounded pdf and set

$$w_n(x) = \frac{1}{\lambda_n^2} w\left(\frac{x}{\lambda_n}\right),$$

where the bandwidth $\lambda_n \to 0$ and $\lambda_n^2|S_n| \to \infty$.

Estimating extremogram--random pattern

$$\rho_{A,B}(h) = \lim_{x \to \infty} P(X(s+h) \epsilon\, xB, X(s) \epsilon\, xA)/PX(s) \epsilon\, xA), \qquad h \in \mathbb{R}^2$$

Kernel estimate of ρ:

$$\hat{\rho}_{A,B}(h) = \frac{\frac{m_n}{\nu^2 |S_n|} \sum_{i \neq j=1}^{N_n} w_n(h - s_i + s_j) I(X(s_i) \in a_m B) I(X(s_j) \in a_m A)}{\frac{m_n}{\nu |S_n|} \sum_{j=1}^{N_n} I(X(s_j) \in a_m A)}$$

$$\hat{\rho}_{A,B}(h) = \frac{\frac{m_n}{\nu^2 |S_n|} \int_{S_n} \int_{S_n} w_n(h + s_1 - s_2) I(X(s_1) \in a_m B) I(X(s_2) \in a_m A)\, N^2(ds_1, ds_2)}{\frac{m_n}{\nu |S_n|} \int_{S_n} I(X(s) \in a_m A) N(ds)}$$

Note: $N^2(ds_1, ds_2) = N(ds_1)N(ds_2)I(s_1 \neq s_2)$ is product measure off the diagonal.

Limit Theory

Theorem: Under suitable conditions on $(X(s))$, (i.e., regularly varying, mixing, local uniform negligibility, etc.), then

$$\left(\frac{|S_n| \lambda_n^2}{m_n}\right)^{\frac{1}{2}} \left(\hat{\rho}_{A,B}(h) - \rho_{A,B,m}(h)\right) \to N(0, \Sigma),$$

where $\rho_{A,B,m}(h)$ is the pre-asymptotic extremogram,

$$\rho_{A,B,m}(h) = P(X(s+h) \epsilon\, a_m B, X(s) \epsilon\, a_m A)/PX(s) \epsilon\, a_m A), \qquad h \in \mathbb{R}^2,$$

(a_m is the $1 - 1/m$ quantile of $|X(s)|$).

Remark: The formulation of this estimate and its proof follow the ideas of Karr (1986) and Li, Genton, and Sherman (2008).

Limit theory

Asymptotic "unbiasedness": $\hat{\rho}_{A,B}(h)$ is a ratio of two terms;

$$\hat{\rho}_{A,B}(h) = \frac{\hat{\tau}_{A,B,m}(h)}{\hat{\tau}_{A,m}}$$

will show that both are asymptotically unbiased.

Denominator: By RV, stationarity, and independence of N and $(X(s))$,

$$E\hat{\tau}_{A,m} = E\left(\frac{m_n}{\nu|S_n|}\int_{S_n} I(X(s) \in a_m A)N(ds)\right)$$

$$= \frac{m_n}{\nu|S_n|}P(X(0) \in a_m A)E\big(N(S_n)\big)$$

$$= m_n P(X(0) \in a_m A)$$

$$\to \mu(A)$$

Limit Theory

Numerator:

$$E\left(\frac{m_n}{\nu^2|S_n|}\int_{S_n}\int_{S_n} w_n(h+s_1-s_2)I(X(s_1) \in a_m B)I(X(s_2) \in a_m A)\, N^2(ds_1, ds_2)\right)$$

$$=\frac{m_n}{\nu^2|S_n|}\int_{S_n}\int_{S_n} w_n(h+s_1-s_2)P(X(0) \in a_m B, X(s_2-s_1) \in a_m A)\,\nu^2 ds_1 ds_2$$

$$=\frac{1}{|S_n|}\int_{S_n}\int_{S_n}\frac{1}{\lambda_n^2}w\left(\frac{h+s_1-s_2}{\lambda_n}\right)\tau_m(s_2-s_1)\,ds_1 ds_2$$

where $\tau_m(h) = mP(X(0) \in a_m B, X(h) \in a_m A)$. Making the change of variables $y = \frac{h+s_1-s_2}{\lambda_n}$ and $u = s_2$, the expected value is

$$\frac{1}{|S_n|}\int_{\frac{S_n-S_n+h}{\lambda_n}}\int_{S_n\cap(S_n-\lambda_n y+h)} w(y)\tau_m(h-\lambda_n y))\,du dy$$

$$=\int_{\frac{S_n-S_n+h}{\lambda_n}} w(y)\tau_m\big(h-\lambda_{ny}\big)\,dy|S_n \cap (S_n - \lambda_n y + h)|/|S_n|$$

Limit Theory

$$\int_{\frac{S_n - S_n + h}{\lambda_n}} w(y)\tau_m(h-\lambda_n y)\,dy\,\frac{|S_n \cap (S_n - \lambda_n y + h)|}{|S_n|}$$

$$\to \int_{\mathbb{R}^2} w(y)\,\tau_{A,B}(h)dy = \tau_{A,B}(h).$$

Remark: We used the following in this proof.

- $\frac{|S_n \cap (S_n - \lambda_n y + h)|}{|S_n|} \to 1$ and $\frac{S_n - S_n + h}{\lambda_n} \to \mathbb{R}^2$.
- $\tau_m(h-\lambda_n y) = mP(X(0) \in a_m B, X(h-\lambda_n y) \in a_m A) \to \tau_{A,B}(h)$.

Need a condition for the latter.

Limit theory

Local uniform negligibility condition (LUNC): For any $\epsilon, \delta > 0$, there exists a δ' such that

$$\limsup_n nP\left(\sup_{|s|<\delta'} \frac{|X_s - X_0|}{a_n} > \delta\right) < \epsilon.$$

Proposition: If $(X(s))$ is a strictly stationary regularly varying random field satisfying LUNC, then for $\lambda_m \to 0$,

$$mP\left(\frac{X(0)}{a_m} \in A, \frac{X(s+\lambda_m)}{a_m} \in B\right) \to \tau_{A,B}(s)$$

This result generalizes to space points, $0, s_1 + \lambda_m, \ldots, s_k + \lambda_m$.

Limit Theory

Outline of argument:

- Under LUNC already shown asymptotic unbiasedness of numerator and denominator.
  - $E\hat{\tau}_{A,m} \to \mu(A)$
  - $E\hat{\tau}_{A,B,m}(h) \to \tau_{A,B}(h)$

with $\rho_{A,B}(h) = \tau_{A,B}/\mu_A(h)$.

Strategy: Show joint asymptotic normality of $\hat{\tau}_{A,m}$ and $\hat{\tau}_{A,B,m}(h)$

$$\frac{|S_n|}{m_n}\mathrm{var}(\hat{\tau}_{A,m}) \to \frac{\mu(A)}{\nu} + \int_{\mathbb{R}^2} \tau_{A,A}(y)dy \Longrightarrow \hat{\tau}_{A,m}(h) \to_p \mu_A(h)$$

Limit Theory

Step 1: Compute asymptotic variances and covariances.

i. $\frac{|S_n|}{m_n}\mathrm{var}(\hat{\tau}_{A,m}) \to \frac{\mu(A)}{\nu} + \int_{\mathbb{R}^2} \tau_{A,A}(y)dy$

ii. $\left(\frac{|S_n|\lambda_n^2}{m_n}\right)\mathrm{var}(\hat{\tau}_{AB,m}(h)) \to \frac{1}{\nu^2}\tau_{AB}(h)\int_{\mathbb{R}^2} w^2(y)dy$

Proof of (i): Sum of variances + sum of covariances

$$\frac{|S_n|}{m_n}E(\hat{\tau}_{A,m}^2) = \frac{m_n}{\nu^2|S_n|}E\left[\int_{S_n} I(X(s_1) \in a_m A)N(ds_1)\right]$$

$$+\frac{m_n}{\nu^2|S_n|}E\left[\int_{S_n}\int_{S_n} I(X(s_1) \in a_m A, X(s_2) \in a_m A)\, dN^2(ds_1, ds_2)\right]$$

$$\to \frac{\mu(A)}{\nu} + \int_{\mathbb{R}^2} \tau_{A,A}(y)dy$$

Limit Theory

Step 2: Show joint CLT for $\hat{\tau}_{A,m}$ and $\hat{\tau}_{A,B,m}(h)$ using a blocking argument.

Idea: Focus on $\hat{\tau}_{A,B,m}(h)$. Set

$A_n =$

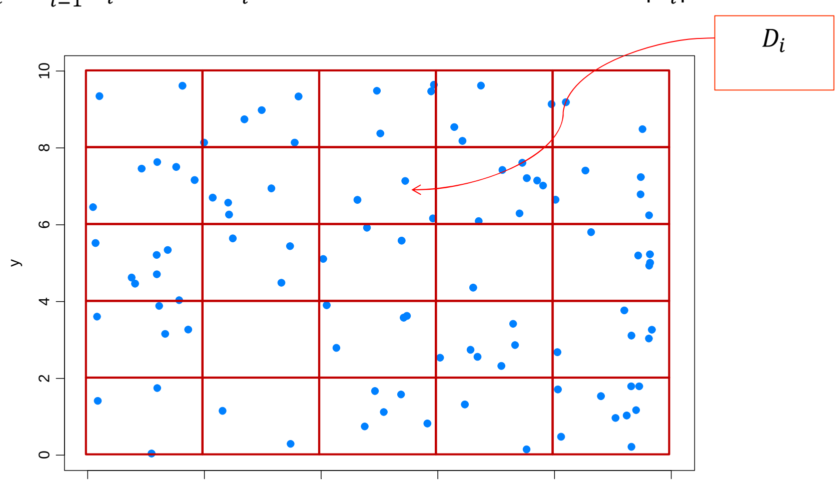
$$\left(\frac{m_n \lambda_n^2}{|S_n|}\right)^{\frac{1}{2}} \frac{1}{\nu^2} \int_{S_n} \int_{S_n} w_n(h + s_1 - s_2) I(X(s_1) \in a_m A) I(X(s_2) \in a_m B)\, N^2(ds_1, ds_2)$$

and put $\tilde{A}_n = A_n - E(A_n)$. We will show $\tilde{A}_n$ is asymptotically normal.

Limit Theory

Subdivide $S_n = [0, n]^2$ into big blocks and small blocks.

$S_n = \cup_{i=1}^{k_n} D_i$ where D_i has diminesions $n^\alpha \times n^\alpha$ and size $|D_i| = n^{2\alpha}$



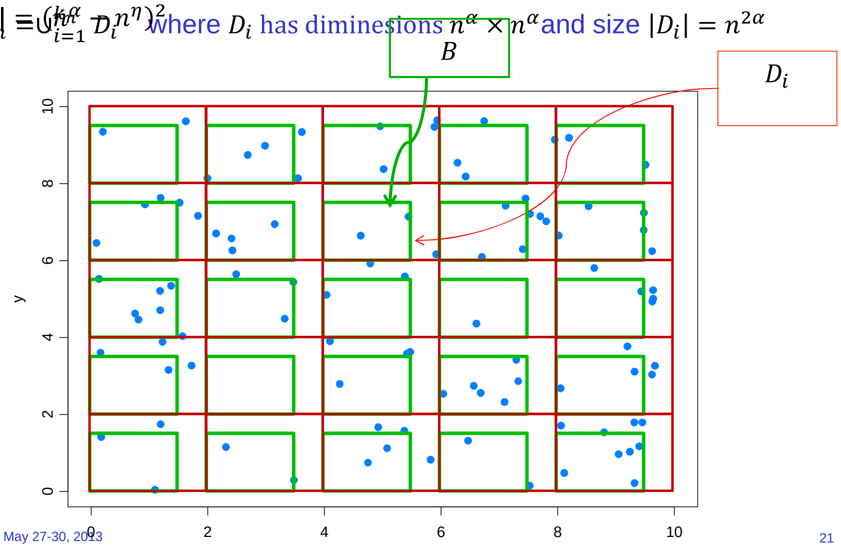
Limit Theory

Subdivide $S_n = [0,n]^2$ into big blocks and small blocks:

$S_n = \cup_{i=1}^{k_n} D_i$ where D_i has diminesions $n^\alpha \times n^\alpha$ and size $|D_i| = n^{2\alpha}$

Inset block B_i inside D_i with dimension $(n^\alpha - n^\eta) \times (n^\alpha - n^\eta)$:

$|B_i| = (n^\alpha - n^\eta)^2$



Limit Theory

Recall that $\tilde{A}_n$ is a (mean-corrected) double integral over $S_n \times S_n$, $i.e.$,

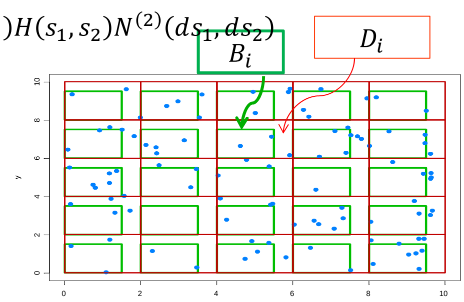
$$\tilde{A}_n = \int_{S_n \times S_n} w_n(h + s_1 - s_2) H(s_1, s_2) N^{(2)}(ds_1, ds_2)$$

$$= \sum_{i=1}^{k_n} \int_{D_i \times D_i} w_n(h + s_1 - s_2) H(s_1, s_2) N^{(2)}(ds_1, ds_2)$$

$$= \sum_{i=1}^{k_n} \int_{B_i \times B_i} w_n(h + s_1 - s_2) H(s_1, s_2) N^{(2)}(ds_1, ds_2)$$

$$+ \ R_n$$

$$= \sum_{i=1}^{k_n} \tilde{a}_{ni} + \tilde{A}_n - \sum_{i=1}^{k_n} \tilde{a}_{ni}$$



Limit Theory

Remaining steps: $\tilde{a}_{ni} = \int_{B_i \times B_i} w_n(h + s_1 - s_2) H(s_1, s_2) N^{(2)}(ds_1, ds_2)$

i. Show $\mathrm{var}\left(\tilde{A}_n - \sum_{i=1}^{k_n} \tilde{a}_{ni}\right) \to 0.$

ii. Let $(\tilde{c}_{ni})$ be an iid sequence with $\tilde{c}_{ni} =_d \tilde{a}_{ni}$ whose sum has characteristic function $\phi_n^c(t)$. Show $\phi_n^c(t) \to \exp\left(-\frac{\sigma^2}{2} t^2\right)$.

iii. $\phi_n^c(t) - \phi_n(t) \to 0.$

Intuition.

(i) The sets $D_i \setminus B_i$ are small by proper choice of α and η.

(ii) Use a Lynapounov CLT (have a triangular array).

(iii) Use a Bernstein argument (see next page).

Limit Theory

Useful identity: $\prod_{i=1}^{k} a_i - \prod_{i=1}^{k} b_i = \sum_{i=1}^{k} a_1 \cdots a_{i-1}(a_i - b_i) b_{i+1} \cdots b_k$

$$|\phi_n(t) - \phi_n^c(t)| = |E \prod_{i=1}^{} e^{it\tilde{a}_{ni}} \quad - E \prod_{i=1}^{k} e^{it\tilde{c}_{ni}}|$$

$$= |E \sum_{i=1}^{k_n} \prod_{j=1}^{i-1} e^{it\tilde{a}_{nj}} (e^{it\tilde{a}_{ni}} - e^{it\tilde{c}_{ni}}) \prod_{j=i+1}^{k_n} e^{it\tilde{c}_{ni}}|$$

$$\le \sum_{i=1}^{k_n} |\mathrm{cov}(\textstyle\prod_{j=1}^{i-1} e^{it\tilde{a}_{nj}}, e^{it\tilde{a}_{ni}})| \quad \text{(by indep of } \tilde{c}_{ni})$$

$$\le \sum_{i=1}^{k_n} |E(\mathrm{cov}(\textstyle\prod_{j=1}^{i-1} e^{it\tilde{a}_{nj}}, e^{it\tilde{a}_{ni}}) | N)|$$

$$\le \sum_{i=1}^{k_n} 4E\alpha_{\left(N(\cup_{j=1}^{i-1} B_j), N(B_i)\right)}(n^{\eta})$$

where $\alpha_{(r,s)}(h)$ is a strong mixing bounding function that is based on the separation h between two sets U and V with cardinality r and s.

Strong mixing coefficients

Strong mixing coefficients: Let $X(s)$ be a stationary random field on $\mathbb{R}^2$. Then the mixing coefficients are defined by

$$\alpha_{j,k}(h) = \sup_{E_1,E_2} |P(A \cap B) - P(A)P(B)|,$$

where the sup is taking over all sets $A \in \sigma(E_1), B \in \sigma(E_2)$, with $card(E_1) \leq j, card(E_2) \leq k$, and $d(E_1, E_2) \geq h$.

Proposition (Li, Genton, Sherman (2008), Ibragimov and Linnik (1971)): Let U and V be closed and connected sets such that $|U| \leq s, |V| \leq t$ and $d(U,V) \geq h$. If X and Y are rvs measurable wrt $\sigma(U)$ and $\sigma(V)$, respectively, and bded by 1, then

$$\mathrm{cov}(X,Y) \leq 4\alpha_{s,t}(h)$$

$\left(16\alpha_{s,t}(h)\ \ if\ X, Y\ complex\right)$.

Limit Theory

Mixing condition: $\sup_s \alpha_{ss}(h)/s = O\,(h^{-\epsilon})$ for some $\epsilon > 2$.

Returning to calculations:

$$|\phi_n(t) - \phi_n^c(t)| = \sum_{i=1}^{k_n} |\mathrm{cov}(\prod_{j=1}^{i-1} e^{it\tilde{a}_{nj}}, e^{it\tilde{a}_{ni}})|$$

$$\leq \sum_{i=1}^{k_n} 16E\alpha_{\left(N(\cup_{j=1}^{i-1}B_j), N(B_i)\right)}(n^\eta)$$

$$\leq \sum_{i=1}^{k_n} 16E\,N\left(\cup_{j=1}^{i} B_j\right)n^{-\epsilon\eta}$$

$$\leq \sum_{i=1}^{k_n} 16in^{2\alpha}n^{-\epsilon\eta} \leq Ck_n^2 n^{2\alpha}n^{-\epsilon\eta}$$

$$= Cn^{4-2\alpha-\epsilon\eta}$$

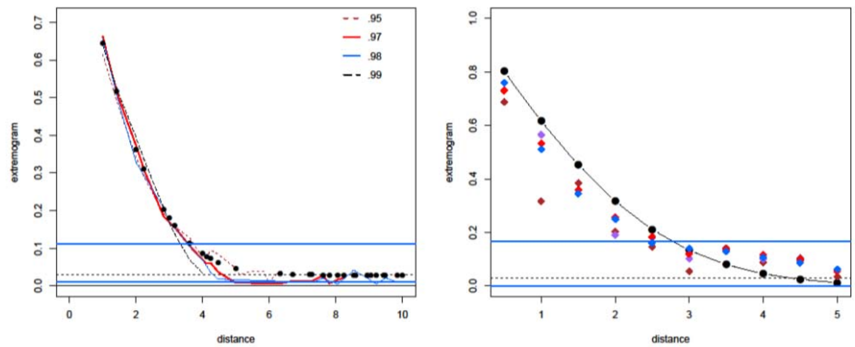
$$\to 0\ \ if\ (4 - 2\alpha - \epsilon\eta < 0).$$

Simulations of spatial extremogram

Extremogram for one realization of B-R process
(function of level)

Note: black dots = true; blue bands are permutation bounds

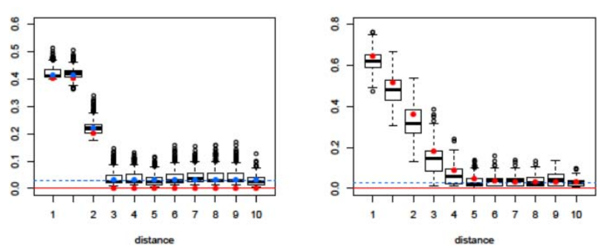
Lattice | Non-Lattice



Simulations of spatial extremogram

Box-plots based on 1000 (100) replications of MMA(1) (left) and BR (right)

Lattice



Non-lattice;
$\lambda_n = 1/\log n$ (left)
$\lambda_n = 5/\log n$ (right)