Applications of the Extremogram to Time Series and Spatial Processes

Richard A. Davis
Columbia University

Collaborators:
Thomas Mikosch, University of Copenhagen
Ivor Cribben, University of Alberta
Yongbum Cho, AIG
Souvik Ghosh, LinkedIn
Claudia Klüppelberg, Technical University of Munich
Christina Steinkohl, Alliance
Yuwei Zhao, University of Ulm

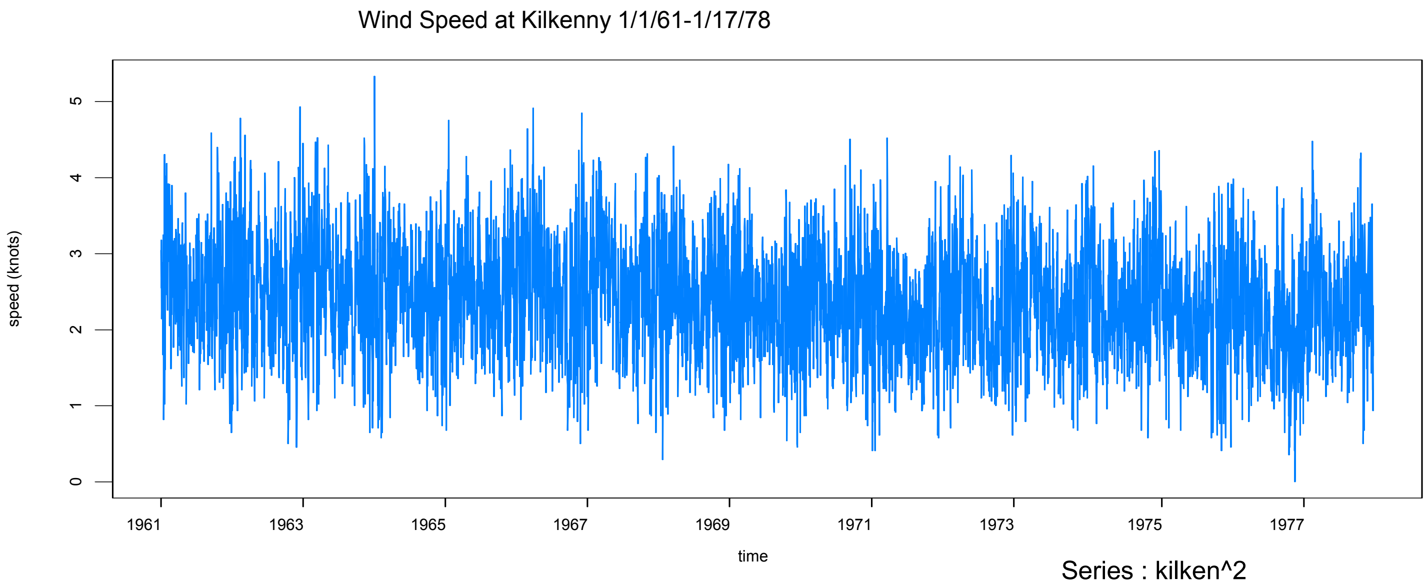
Extremes and Time Series Modeling

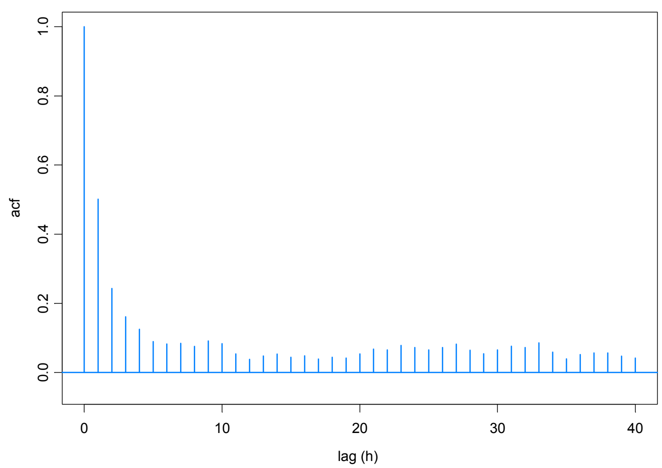
Do fitted models actually capture the desired (*extremal)* characteristics of the data?

- How do we assess “fitted” (expected) with “observed”?
- Need a mechanism for measuring extremal dependence.

Goal of this talk: Describe the extremogram which may be useful as a tool for addressing this question. That is, how well does the “sample” extremogram match up with the “population” extremogram?

Illustration (Windspeed at Kilkenny)





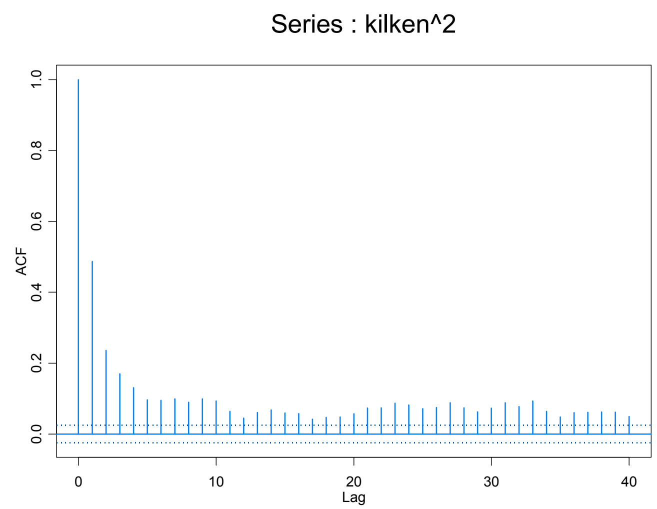
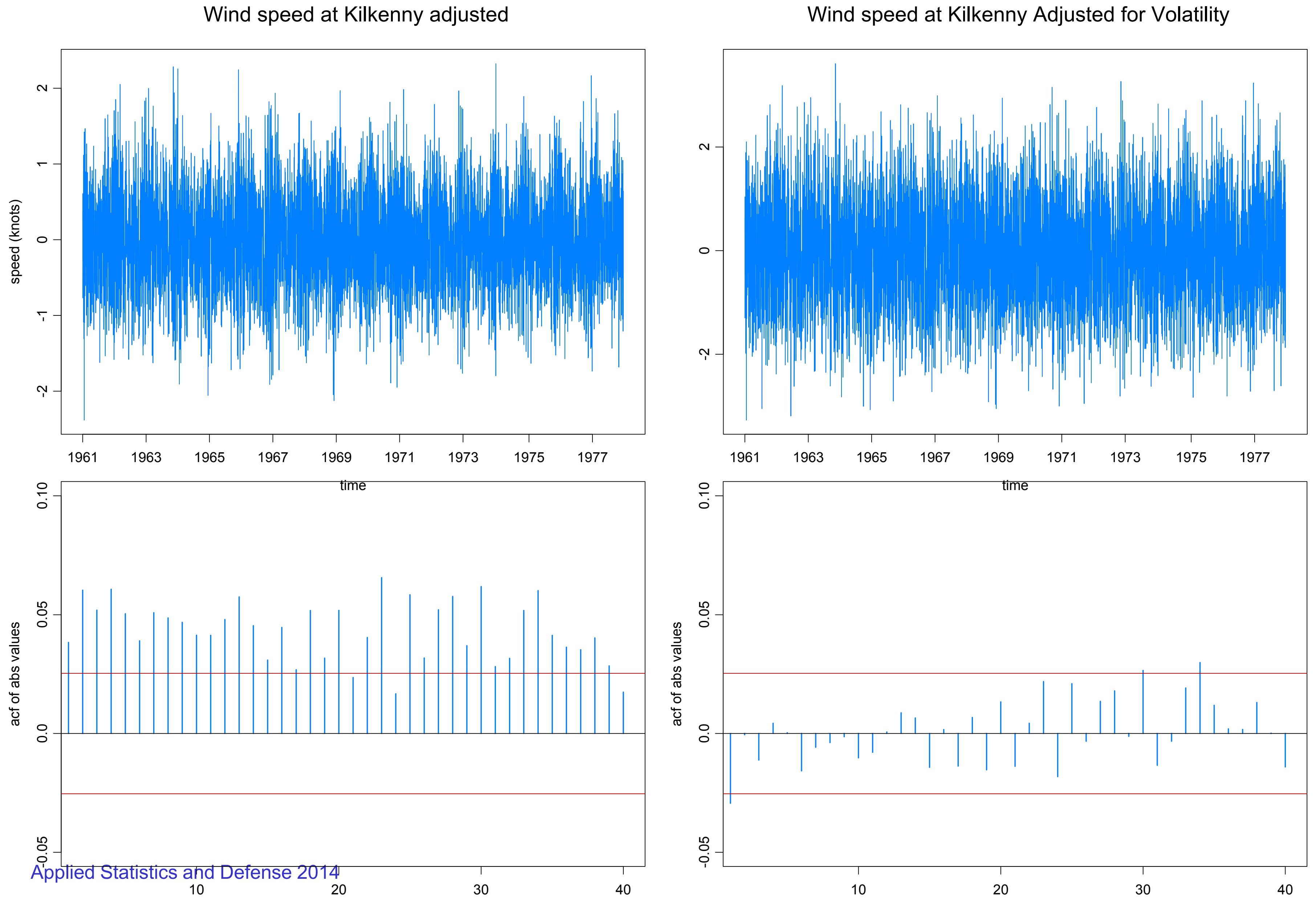
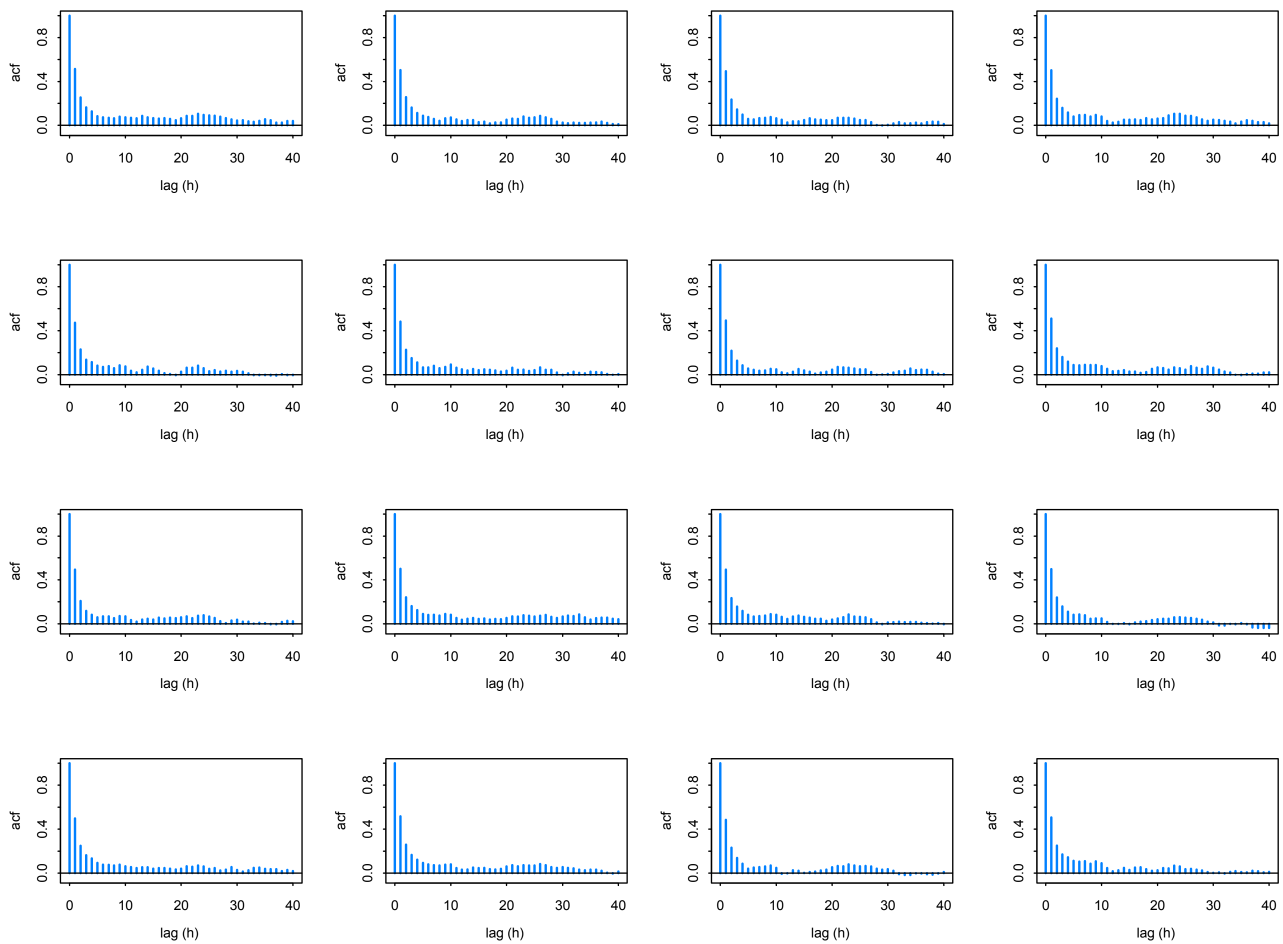


Illustration with ACF



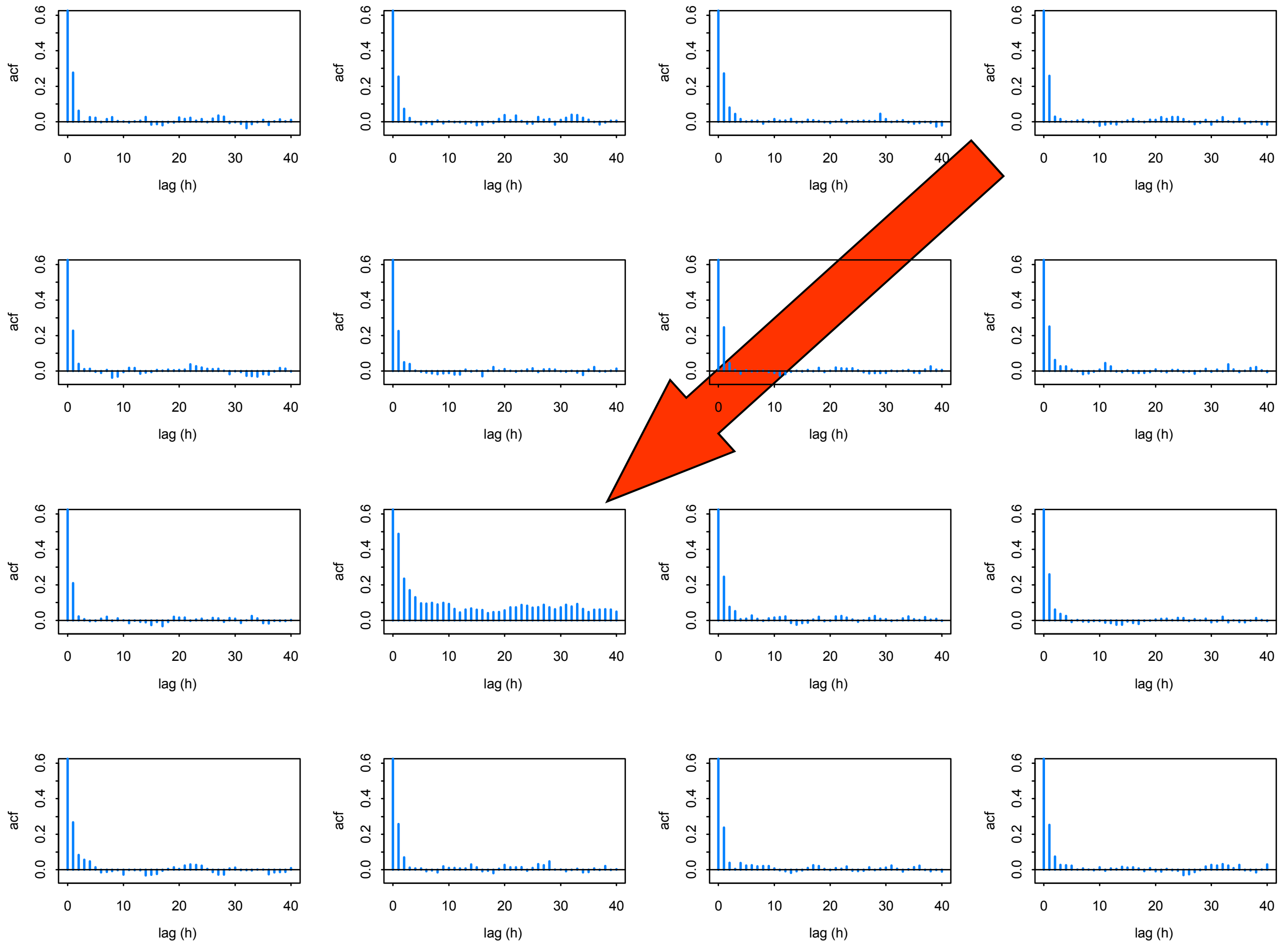
ACF Plots for Kilkenny

ACF of the from 15 simulated realizations from fitted AR model + real data.



ACF Plots for Kilkenny

ACF of the ***squares*** from 15 simulated realizations from fitted AR model + real data.



Game Plan

☞ Extremes and time series modeling

- A motivating example
- Starting point: GARCH vs SV

☞ The Extremogram

- Examples
- Sufficient conditions for existence: regular variation
- Empirical extremogram
- Illustrations (permutation procedures)
- Cross-extremogram (devolatilizing/deGARCHing)

☞ Application to spatial processes

- Kernel estimate of extremogram
- Rainfall data

Characteristics of financial time series

Define $X_t = \ln (P_t) - \ln (P_{t-1})$ (log returns)

- heavy tailed

$$P(|X_1| > x) \sim RV(-\alpha), \quad 0 < \alpha < 4.$$

- uncorrelated

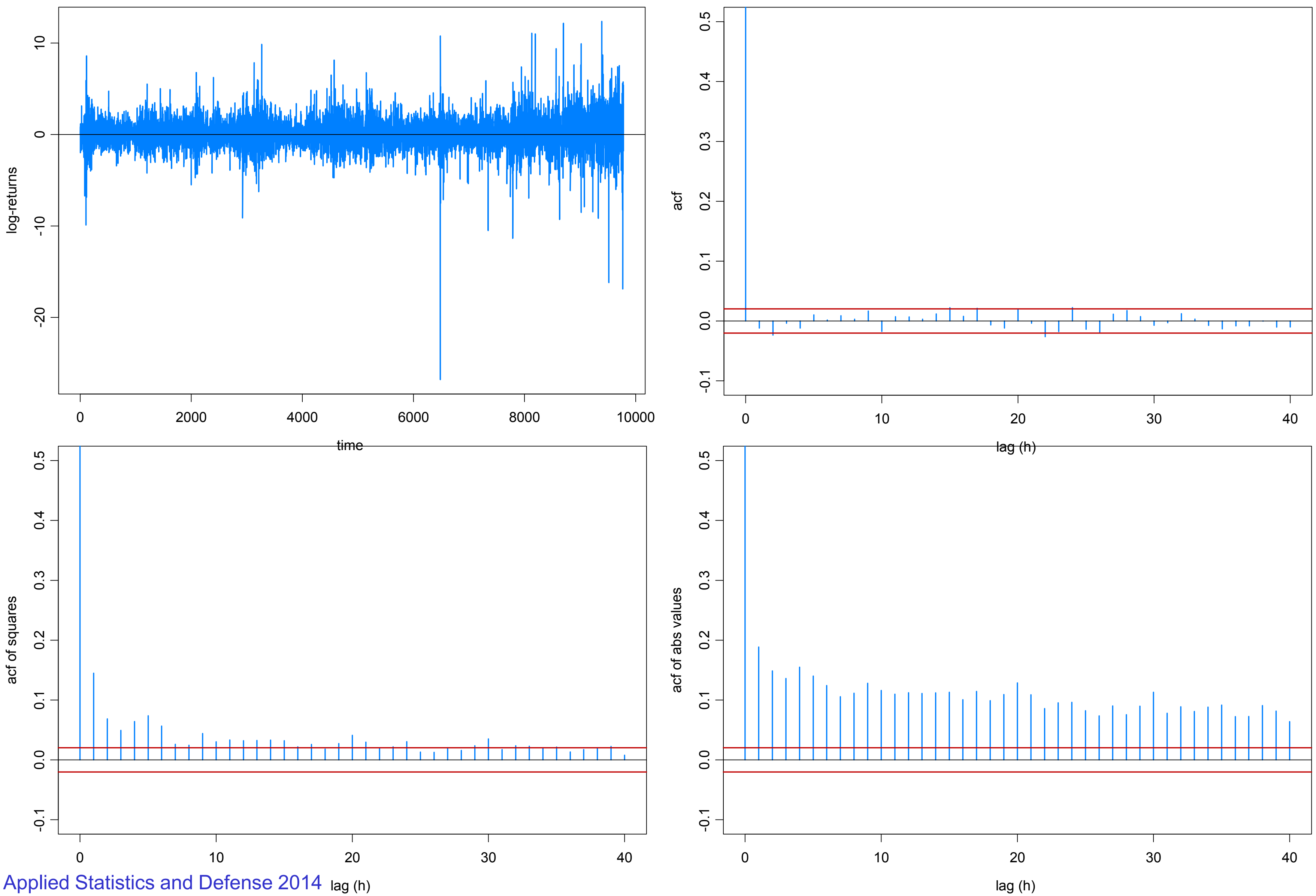
$\hat{\rho}_X(h)$ near 0 for all lags $h > 0$

- $|X_t|$ and X_t^2 have slowly decaying autocorrelations

$\hat{\rho}_{|X|}(h)$ and $\hat{\rho}_{X^2}(h)$ converge to 0 *slowly* as h increases.

- process exhibits 'volatility clustering'.

Example: Log returns for IBM 1/3/62-11/3/00



Starting point: GARCH vs SV

$X_t = \sigma_t Z_t$ (observation eqn in state-space formulation)

(i) GARCH(1,1)

$$X_t = \sigma_t Z_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 X_{t\text{-}1}^2 + \beta_1 \sigma_{t\text{-}1}^2, \ \{Z_t\} \sim \mathrm{IID}(0,1)$$

(ii) Stochastic Volatility

$$X_t = \sigma_t Z_t, \quad \log \sigma_t^2 = \phi_0 + \phi_1 \log \sigma_{t-1}^2 + \varepsilon_t, \quad \{\varepsilon_t\} \sim \mathrm{IID\,N}(0,\sigma^2)$$

Key question:

What intrinsic (extremal?) features in the data (*if any*) can be used to discriminate between these two models?

The Extremogram

The extremogram of a stationary time series $\{X_t\}$ can be viewed as the analogue of the correlogram in time series for measuring dependence in extremes (see Davis and Mikosch (2009)).

Definition: For two sets A & B ***bounded away from 0***, the ***extremogram*** is defined as

$$\rho_{A,B}(h) = \lim_{x\to\infty} P(\mathbf{X}_h \in xB \mid \mathbf{X}_0 \in xA)$$

$$= \lim_{x\to\infty} P(\mathbf{X}_0 \in xA, \mathbf{X}_h \in xB)/P(\mathbf{X}_0 \in xA),$$

for $h = 0, 1, \ldots$, provided the limit exists, where $\mathbf{X}_h=(X_h,X_{h+1},\ldots,X_{h+k})$.

Remark: This definition requires that the limit exists.

a) exists for heavy-tailed time series (see forthcoming slide)

b) exists for some light-tailed time series w/ special choices of A and B.

c) extremal dependence ***depends*** on the choice of sets A & B.

The Extremogram (cont)

If one takes $A=B=(1,\infty)$ and $k = 0$, then

$$\rho_{A,B}(h) = \lim_{x\to\infty} P(X_h > x \mid X_0 > x) = \lambda(X_0, X_h)$$

often called the ***extremal dependence coefficient*** ($\lambda = 0$ means independence or asymptotic independence).

Other choices of A and B can lead to interesting extremograms:

- $P(X_1 < -x \mid X_0 < -x)$ (negative return followed by a neg return)
- $P(X_1 > x \mid X_0 < -x)$ (neg return followed by a pos return)
- $P(X_1 + \cdots + X_4 > 2x \mid X_0 < -x)$ (neg return followed by a big pos return aggregated over next 4 days)
- $P(X_1 > x, \ldots, X_4 > x \mid X_0 > x)$ (pos return followed by a pos return in next 4 days)

The Extremogram: examples

Let $A = B = (1,\infty)$, then

$$\rho_{A,B}(h) = \lim_{x\to\infty} P(X_0 > x, X_h > x)/P(X_0 > x)$$

Gaussian Processes: In this case,

$$\rho_{A,B}(h) = 0 \text{ for all } h > 0 \quad \text{(asymptotic independence).}$$

connected to the Gaussian copula.

GARCH: In this case

$$\rho_{A,B}(h) > 0 \text{ for all } h > 0,$$

but decays to 0 geometrically fast.

SV process: $X_t = \sigma_t Z_t, \quad \log \sigma_t^2 = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \ \{\varepsilon_t\} \sim \text{IIDN}(0,\sigma^2)$

In this case,

$$\rho_{A,B}(h) = 0 \text{ for all } h > 0.$$

The Extremogram: examples

Let $A = B = (1,\infty)$, then

$$\rho_{A,B}(h) = \lim_{x\to\infty} P(X_0 > x, X_h > x)/P(X_0 > x)$$

AR(1): $X_t = \phi X_{t-1} + Z_t$, $\{Z_t\}$~IID Cauchy. Then

$$\rho_{A,B}(h) = \max(0, \phi^h).$$

Note if $\phi < 0$, then extremogram alternates between positive #'s and 0

MaxMA(2): Let (Z_t) be iid with Pareto distribution, i.e., $P(Z_1 > x) = x^{-\alpha}$ for $x \geq 1$, and set $X_t = \max(Z_t, Z_{t-1}, Z_{t-2})$. Then

$$\rho_{A,B}(h) = 1 \quad \text{for } h = 0.$$
$$= 2/3 \quad \text{for } h = 1$$
$$= 1/3 \quad \text{for } h = 2$$
$$= 0, \quad \text{for } h > 2.$$

Regular Variation — multivariate case

Regular variation of $\mathbf{X}=(X_1, \ldots, X_k)$: (heavy-tailed analogue of multivariate Gaussian)

(i) The radial part $|\mathbf{X}|$ is heavy-tailed, i.e.,

$$P(|\mathbf{X}|> tx)/P(|\mathbf{X}|>t) \to x^{-\alpha}.$$

(ii) The angular part $\mathbf{X} / |\mathbf{X}|$ is asymptotically independent of the radial part $|\mathbf{X}|$, i.e., there exists a random vector $\theta \in S^{k-1}$ such that

$$P(\mathbf{X}/|\mathbf{X}| \in \bullet \mid |\mathbf{X}|>t) \to_w P(\theta \in \bullet) \quad \text{as } t \to \infty.$$

($\to_w$ weak convergence on S^{k-1} = unit sphere in R^k) .

- $P(\theta \in \bullet)$ is called the spectral measure
- α is the index of $\mathbf{X}$.

Definition: A time series $\{X_t\}$ is ***regularly varying*** if all the finite dimensional distributions are regularly varying.

Regular Variation and the Extremogram

Facts

1. The extremogram of a RV stationary time series $\{X_t\}$ exists for all choices of sets A & B bounded away from the origin.

2. Many heavy-tailed time series (GARCH and SV) are regularly varying.

The Empirical Extremogram

A natural estimator of the extremogram,

$$\rho_{A,B}(h) = \lim_{x\to\infty} P(X_h \in xB \mid X_0 \in xA)$$

based on observations, $X_1,\ldots,X_n$, is the empirical extremogram defined by

$$\hat{\rho}_{A,B}(h) = \frac{\frac{m}{n}\sum_{t=1}^{n-h} I_{\{a_m^{-1}X_t \in A,\, a_m^{-1}X_{t+h} \in B\}}}{\frac{m}{n}\sum_{t=1}^{n} I_{\{a_m^{-1}X_t \in A\}}},$$

where a_m is the $1 - 1/m$ quantile of $|X_t|$. For the theory to work, need

$$m_n \to \infty \text{ with } m/n \to 0.$$

Under suitable mixing conditions, a CLT for the empirical estimate is established in D&M (2009).

The Empirical Extremogram — central limit theorem

$$\hat{\rho}_{A,B}(h) = \frac{\frac{m}{n}\sum_{t=1}^{n-h} I_{\{a_m^{-1}X_t \in A,\, a_m^{-1}X_{t+h} \in B\}}}{\frac{m}{n}\sum_{t=1}^{n} I_{\{a_m^{-1}X_t \in A\}}}$$

After first establishing a joint CLT for the numerator and denominator, we obtain the limit result

$$(n/m)^{1/2}(\hat{\rho}_{A,B}(h) - \rho_m(h)) \to_d N(0, \sigma^2(A,B)),$$

where $\rho_m(h)$ is the ratio of expectations (***pre-asymptotic extremogram***),

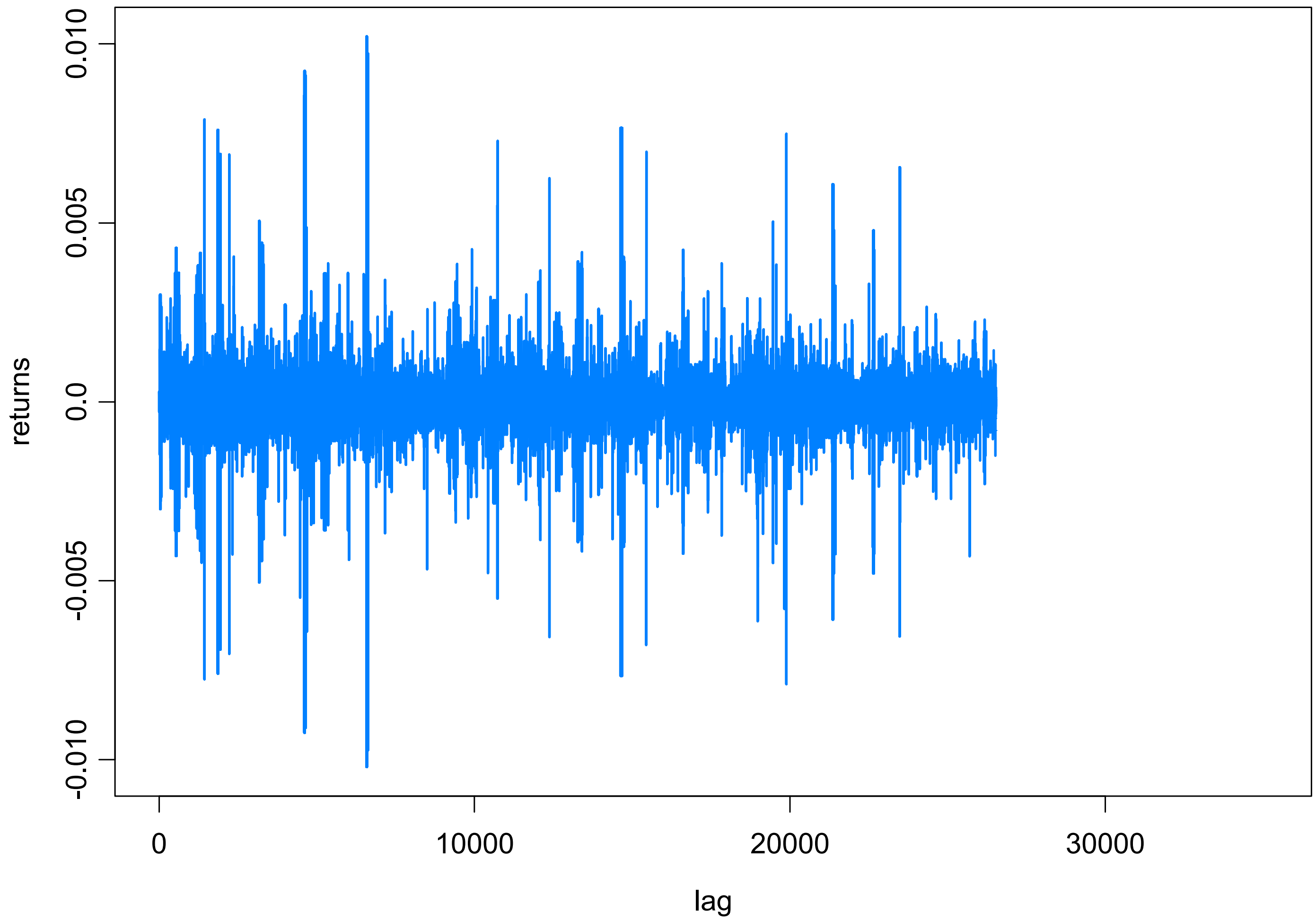
$$P(a_m^{-1}X_0 \in A,\ a_m^{-1}X_h \in B) / P(a_m^{-1}X_0 \in A).$$

Now provided a bias condition, such as

$$(n/m)^{1/2}\left(mP(a_m^{-1}X_0 \in A,\ a_m^{-1}X_h \in B) - \mu_h(A \times B)\right) \to 0,$$

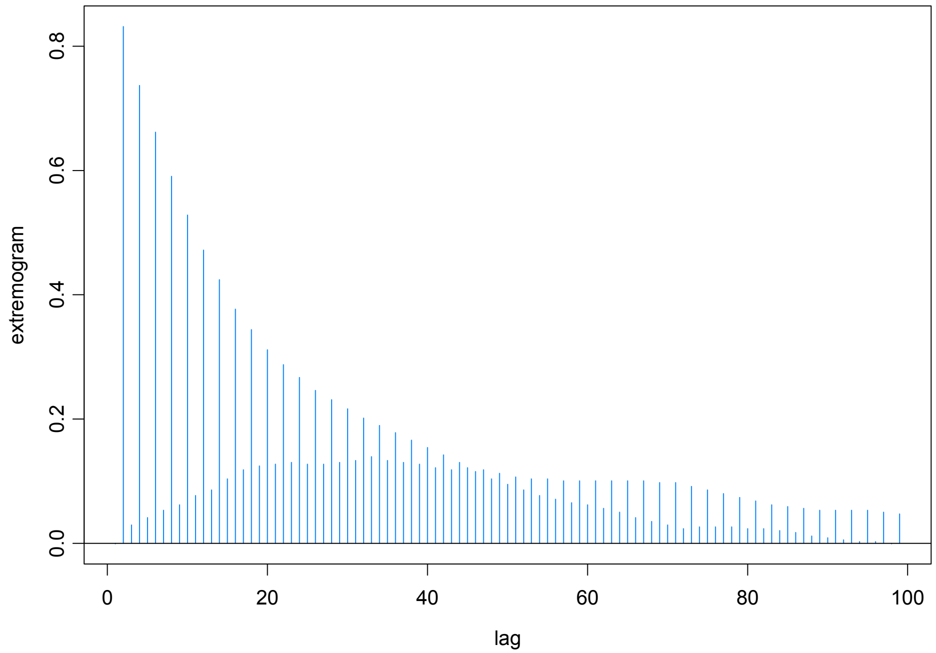
holds, then $\rho_m(h)$ can be replaced with $\rho_{A,B}(h)$. This condition can often be difficult to check.

Application to Five-Minute Return Data (US/DM) exchange



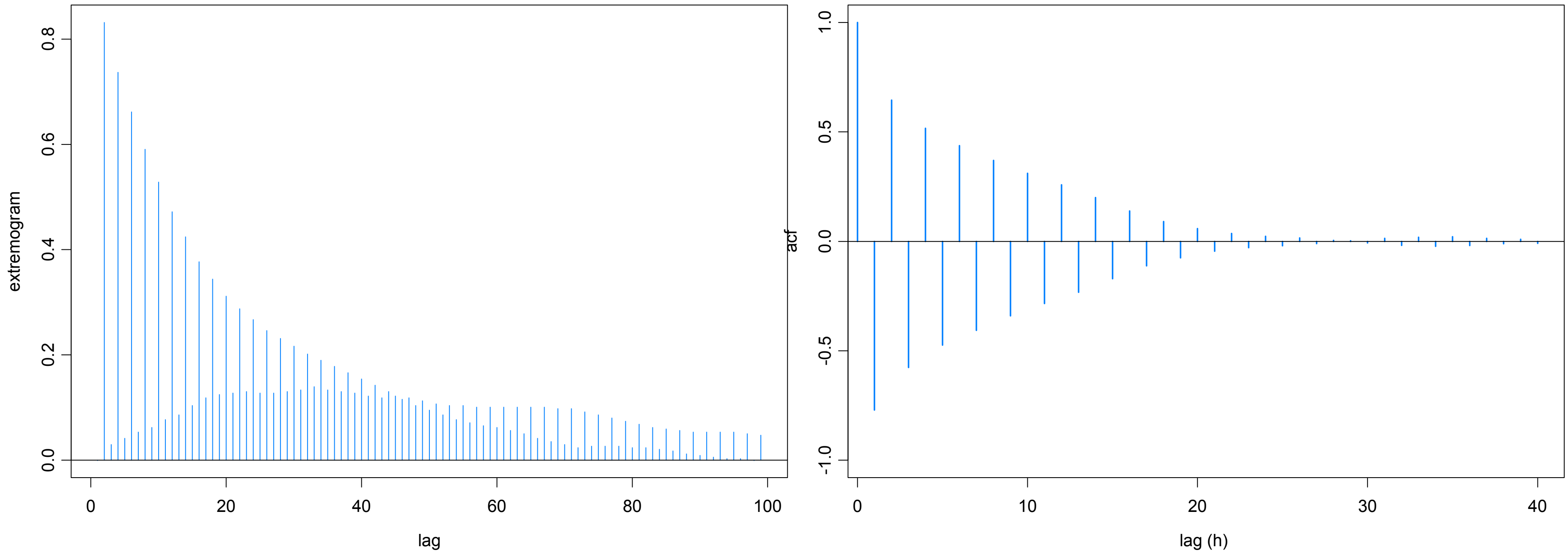
Application to Five-Minute Return Data (US/DM) exchange

Extremogram $A=B=(1,\infty)$



Application to Five-Minute Return Data (US/DM) exchange

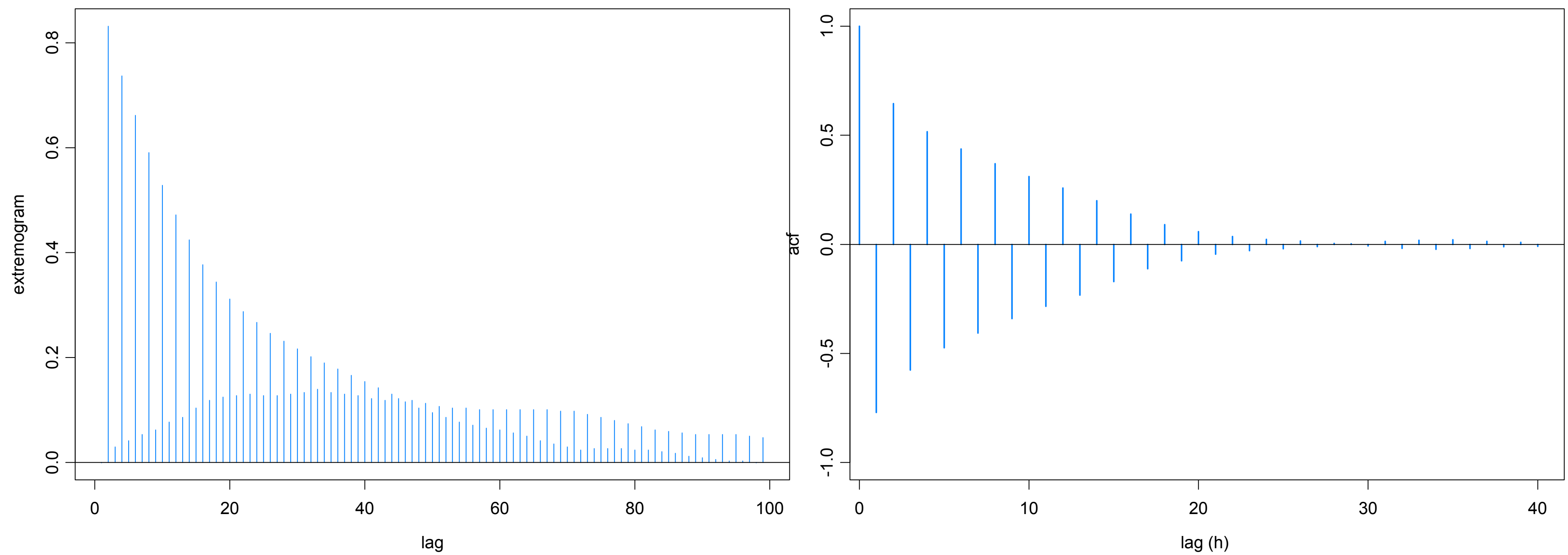
Extremogram A=B=$(1,\infty)$



Best fitting AR model is of order 18; refine with nonzero coefficients at lags 1, 2, 3, 5, 6, 7, 11, 13, 14, 16, and18.

Application to Five-Minute Return Data (US/DM) exchange

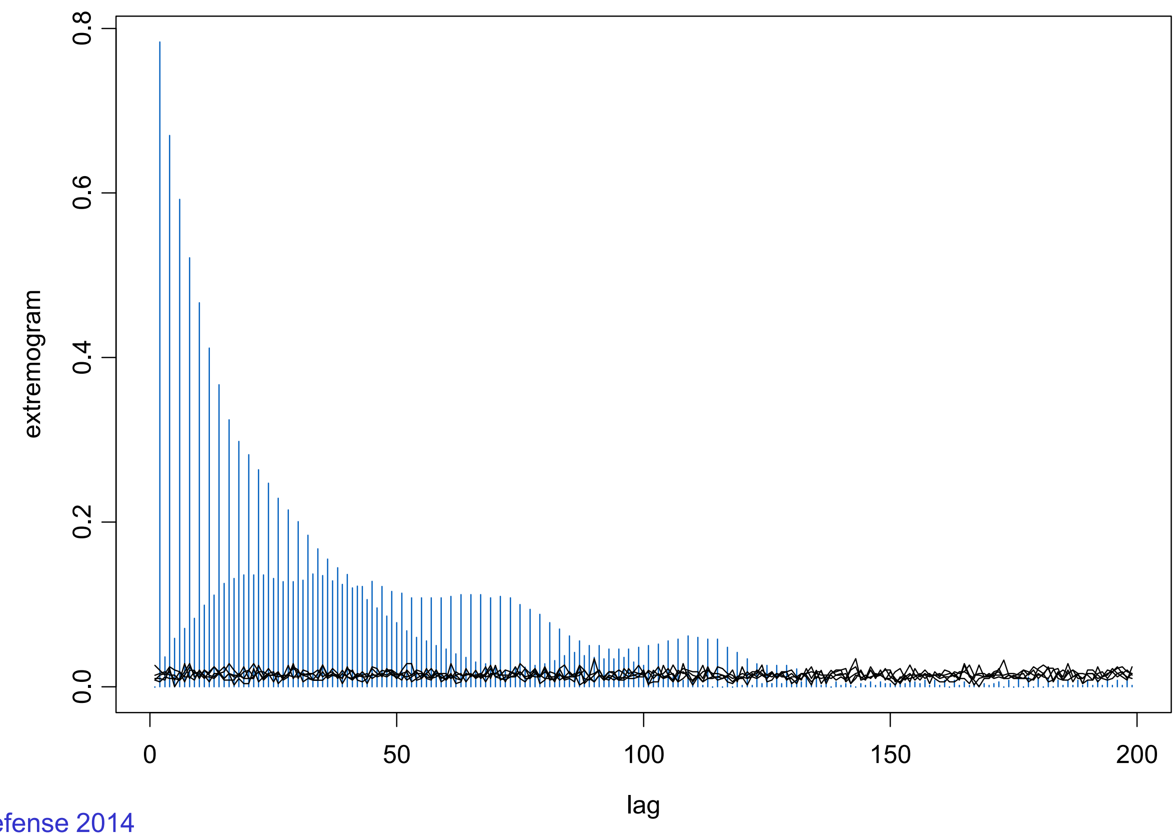
Extremogram A=B=(1,∞)



Best fitting AR model is of order 18; refine with nonzero coefficients at lags 1, 2, 3, 5, 6, 7, 11, 13, 14, 16, and18.

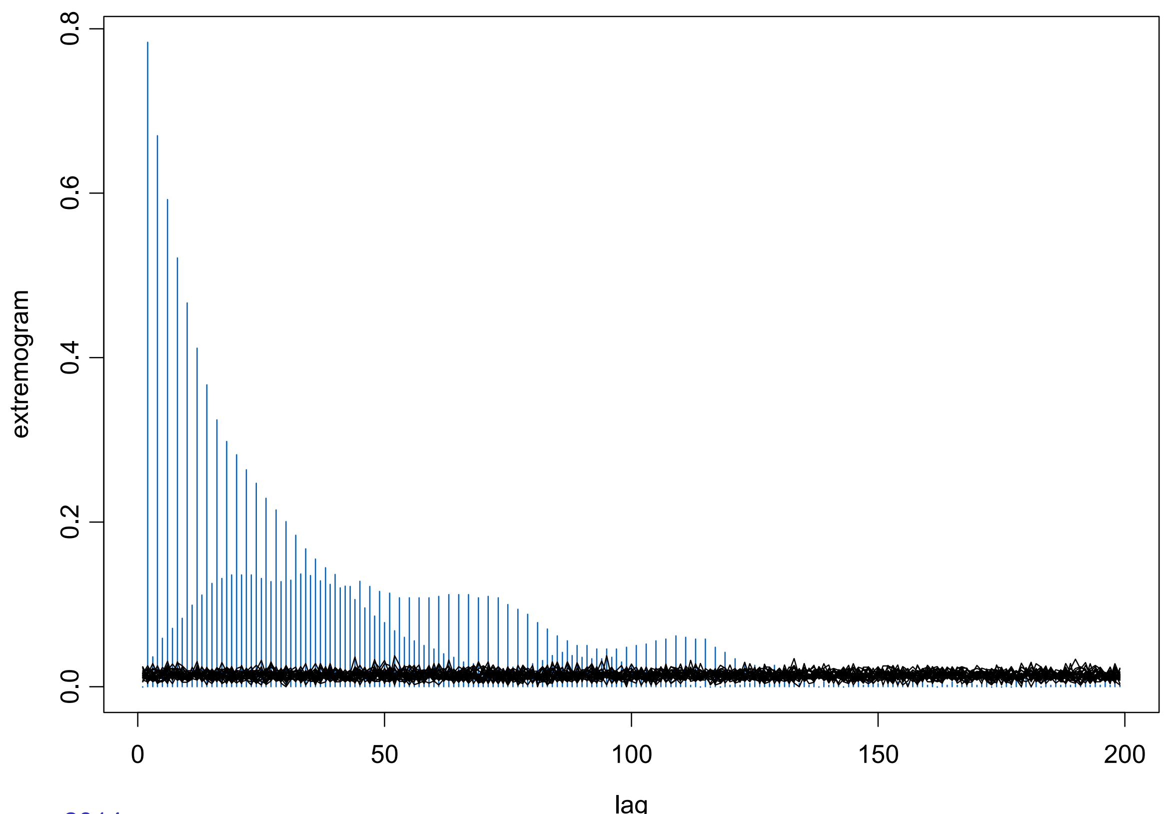
Resampling and Testing for Serial Dependence

A natural way (*not often used in time series*) for testing serial correlation is to compute the ACF for random permutations of the data. If the sample ACF appears more ***extreme*** than the ACFs based on random permutations, then there is evidence of serial correlation. We apply the same idea to the extremogram.



Resampling and Testing for Serial Dependence

A natural way (*not often used in time series*) for testing serial correlation is to compute the ACF for random permutations of the data. If the sample ACF appears more ***extreme*** than the ACFs based on random permutations, then there is evidence of serial correlation. We apply the same idea to the extremogram.



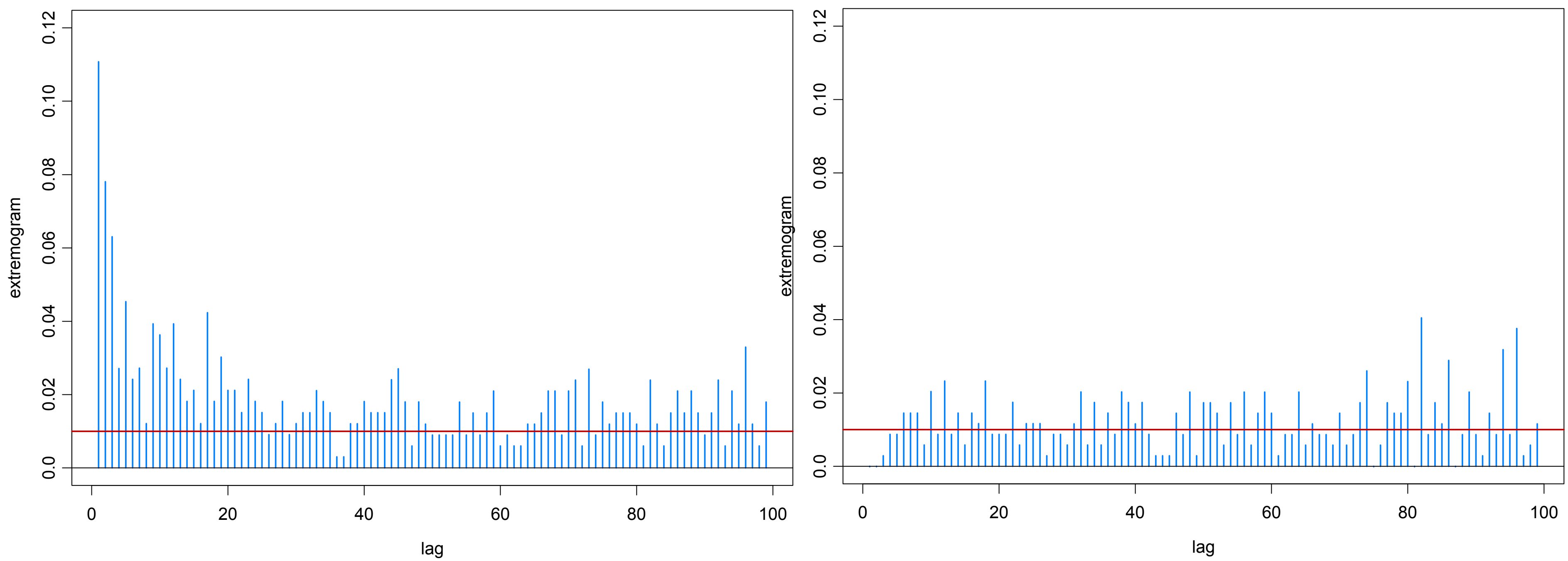
Application to Five-Minute Return Data (US/DM) exchange

Extremogram for residuals from subset AR(18) and from GARCH

A=B=$(1,\infty)$

Residuals from AR

Residuals from GARCH

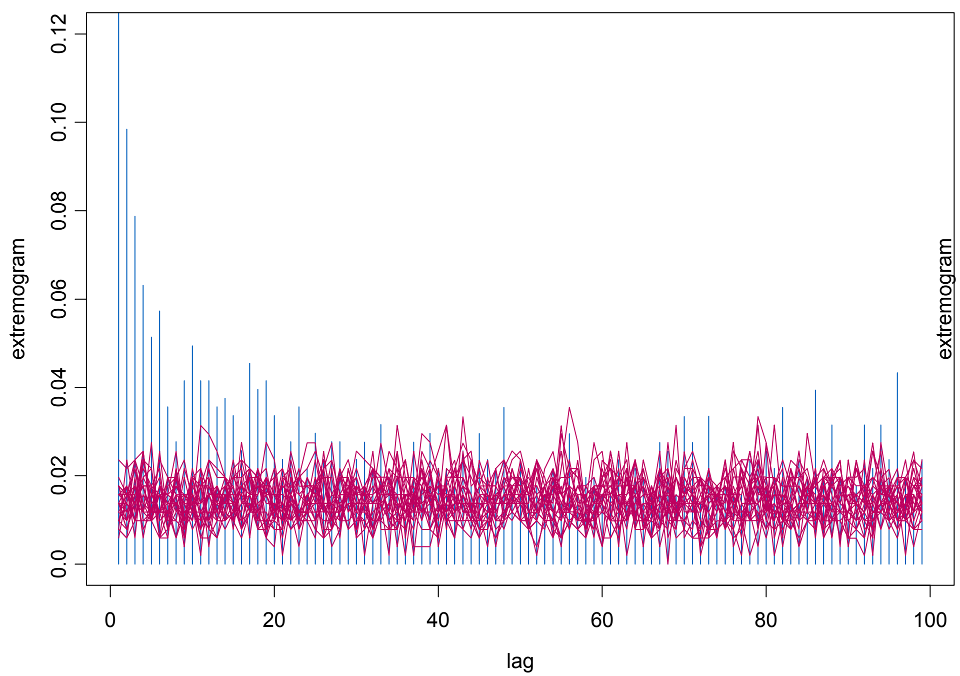


Application to Five-Minute Return Data (US/DM) exchange

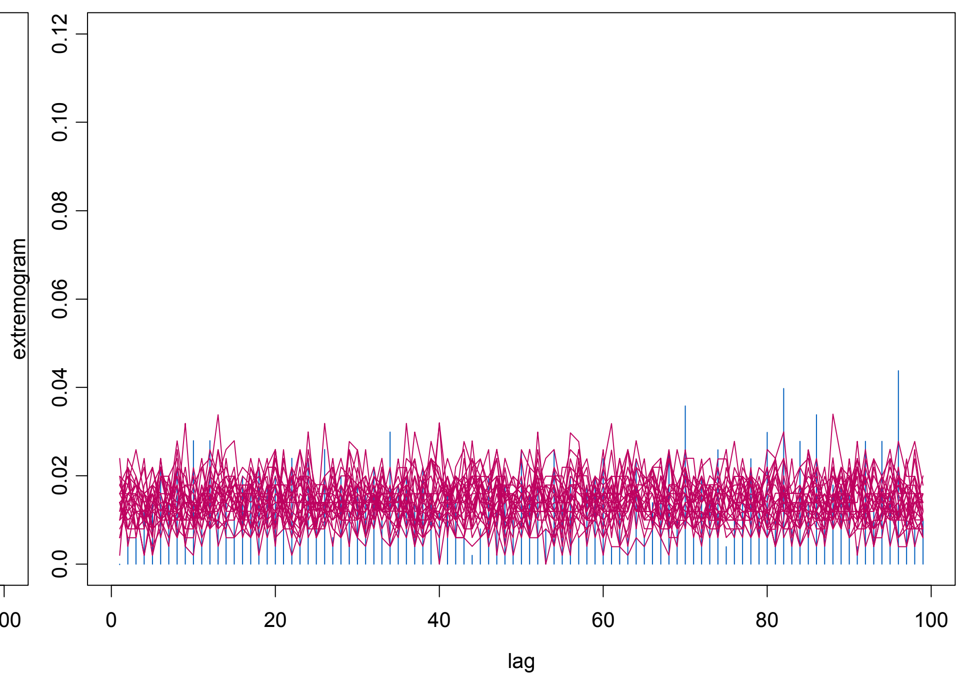
Extremogram for residuals from subset AR(18) and from GARCH

A=B=(1,∞)

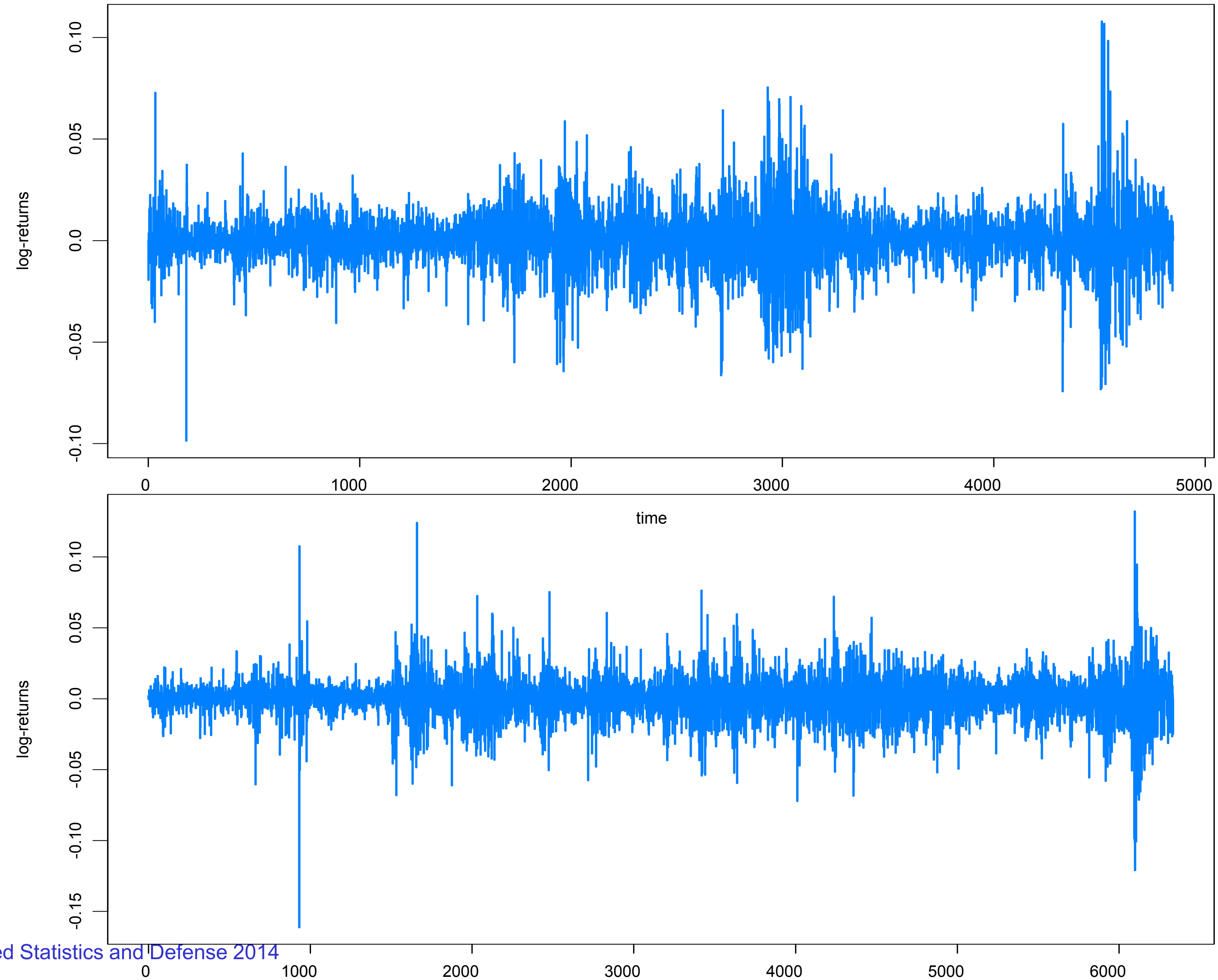
Residuals from AR



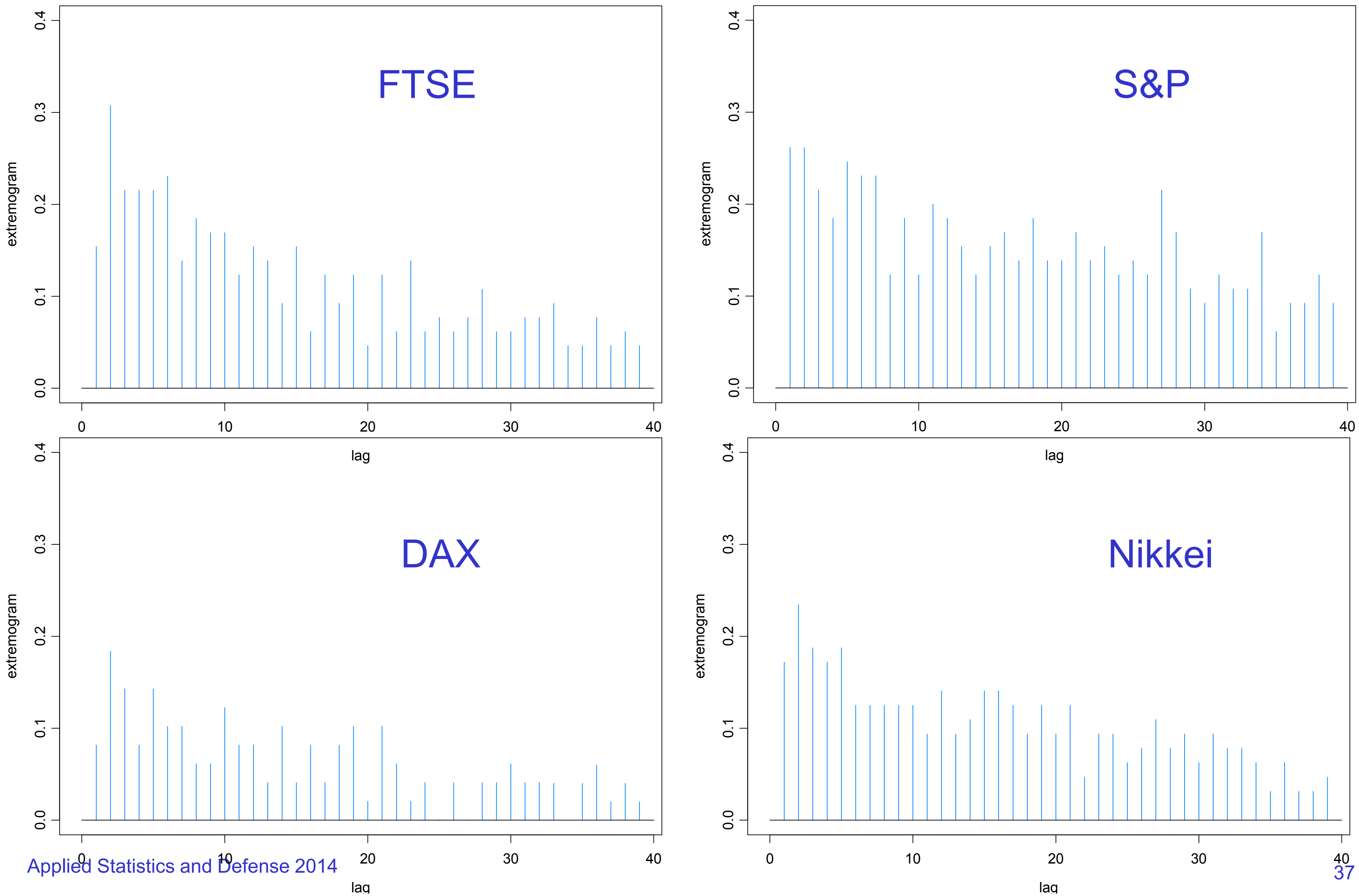
Residuals from GARCH



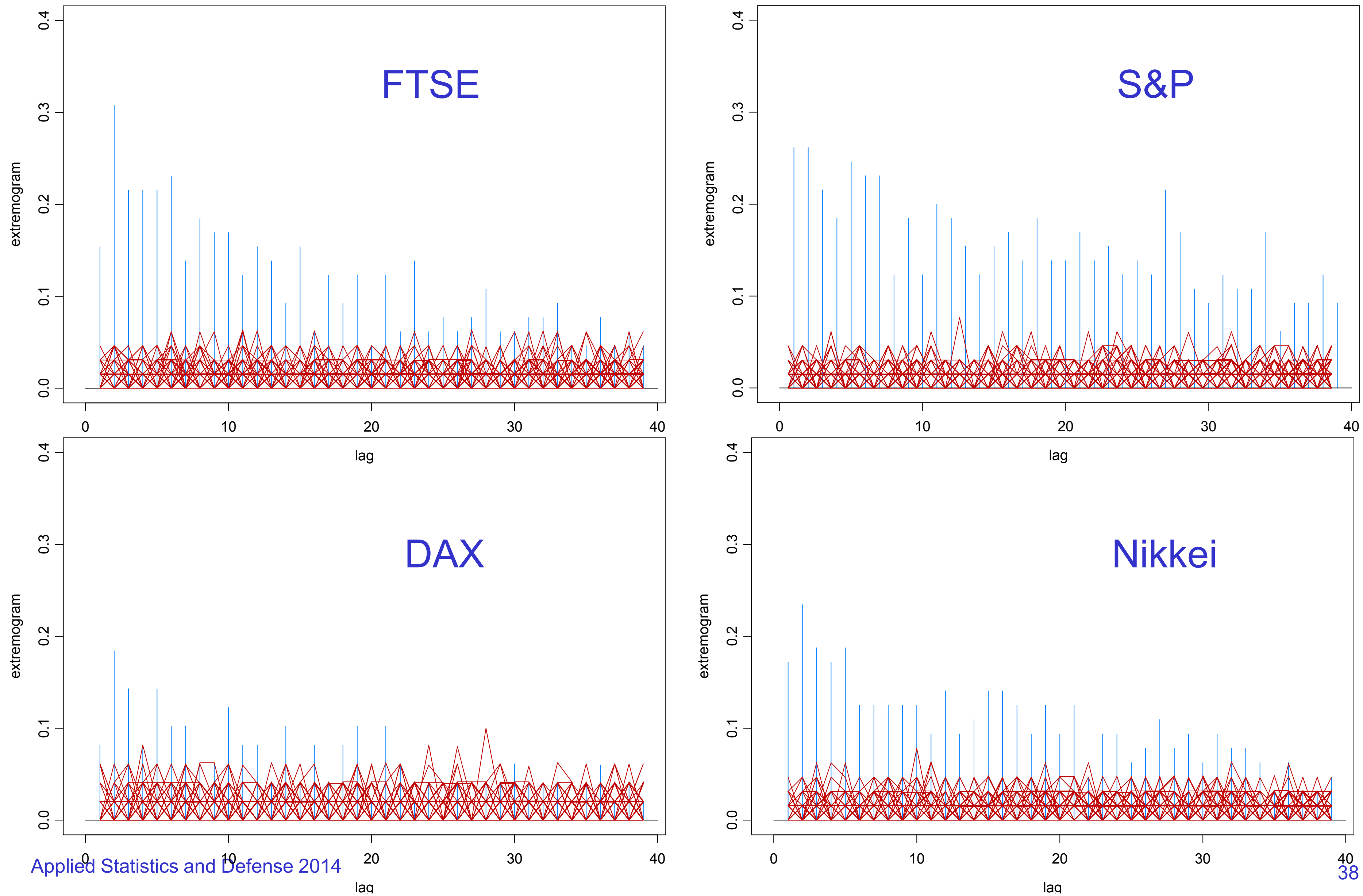
Log-returns for DAX and Nikkei (Apr 4, `84-Oct 2, `09



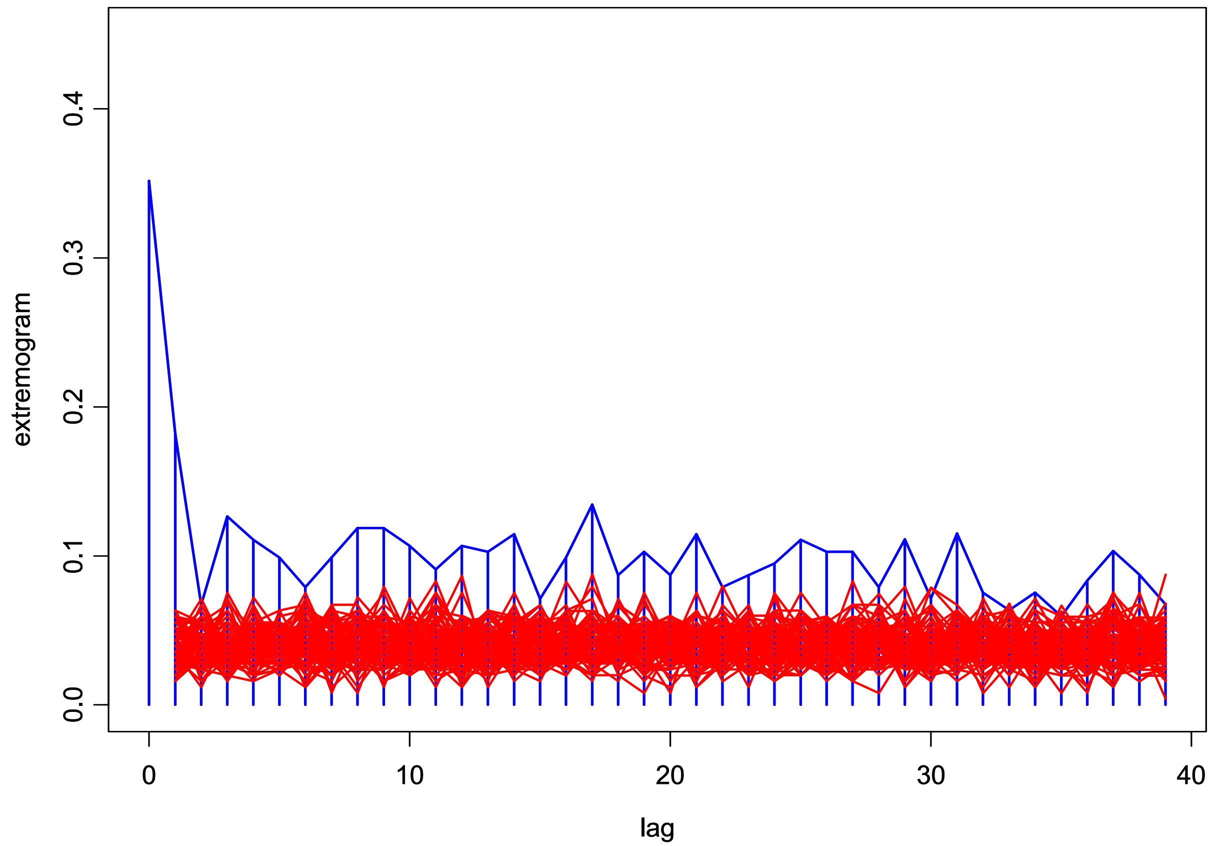
Extremogram for FTSE, S&P, DAX, Nikkei



Extremogram for FTSE, S&P, DAX, Nikkei



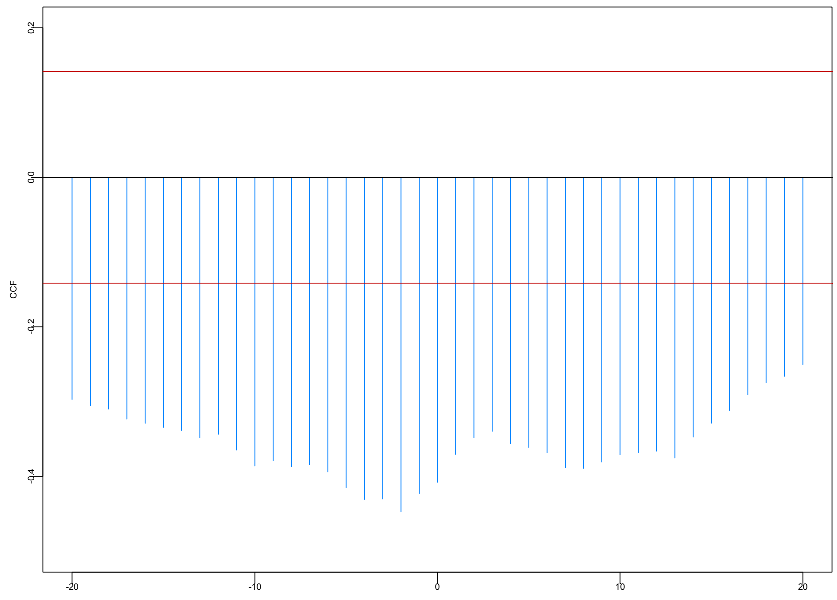
Cross-Extremogram FTSE and SP



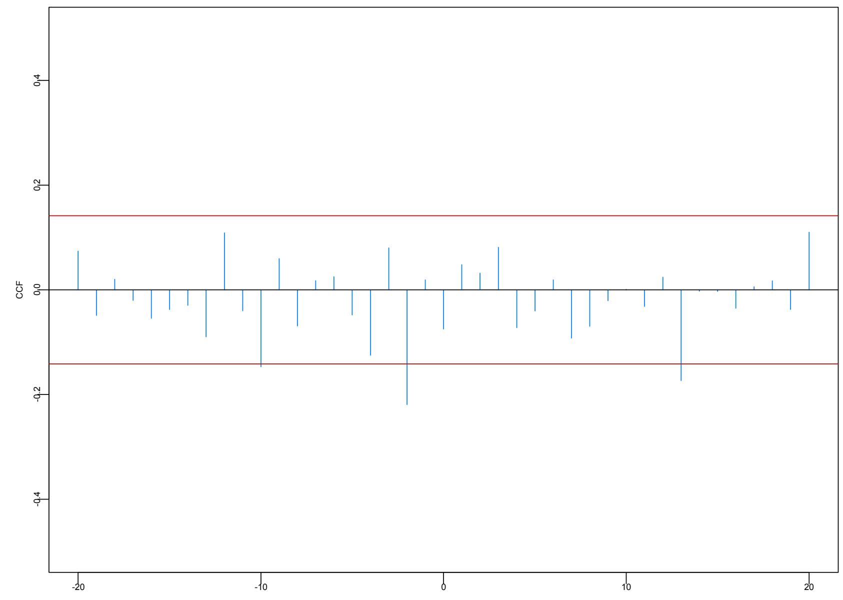
Cross-Extremogram

Strategy: Devolatilize the component series before computing the extremogram. This is ***analogous*** to the issue of spurious cross-correlations in a time series without whitening the series first.

Cross-correlation between two “independent” AR(1)’s



Cross-correlation between the ***whitened*** series’



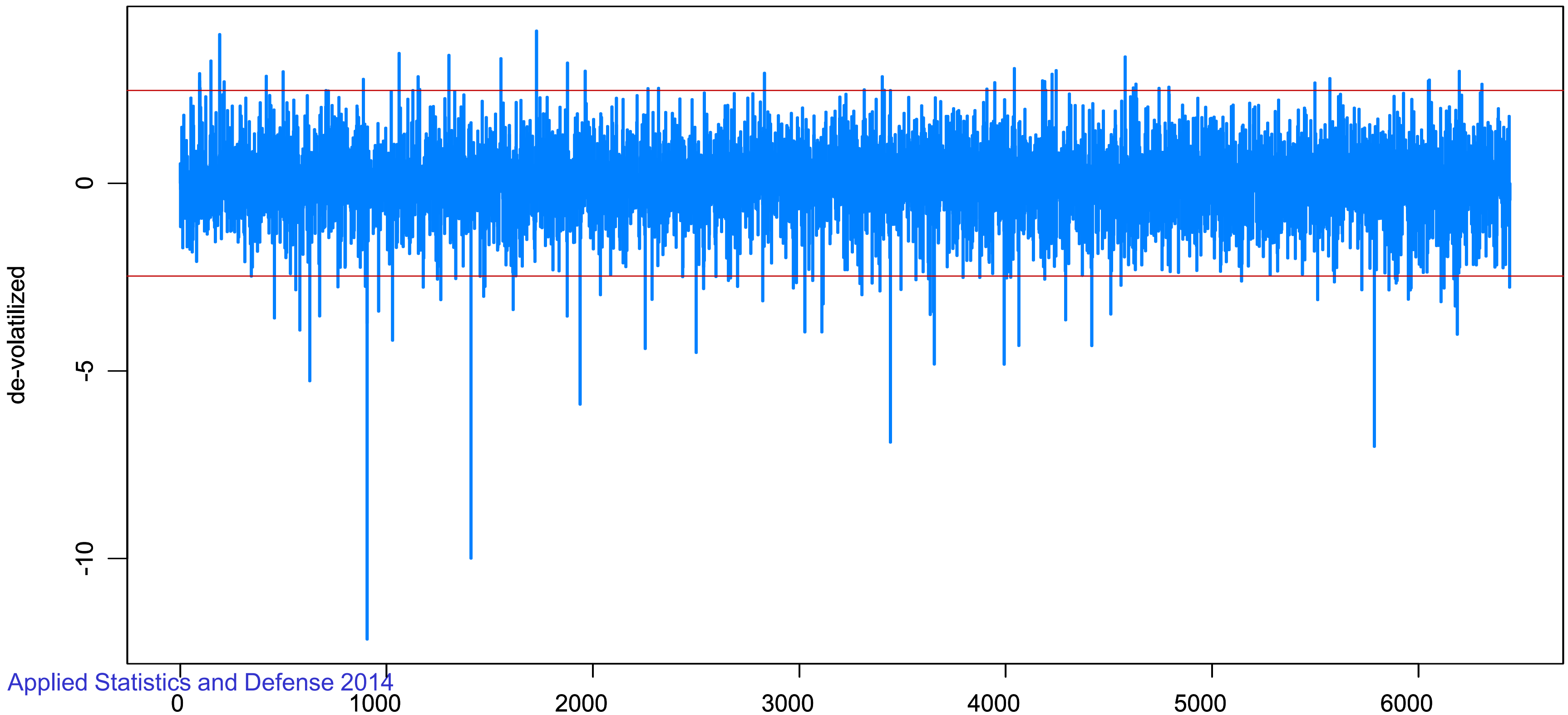
Devolatilizing (deGARCHing) S&P

Extremogram for S&P: significant for large number of lags ~40+

Devolatilize S&P by fitting GARCH(1,1):

$$X_t = (6.28e - 7 + .057X_{t-1}^2 + .939\sigma_{t-1}^2)^{1/2} Z_t,$$

$\{Z_t\} \sim IID\ t(6.14),$



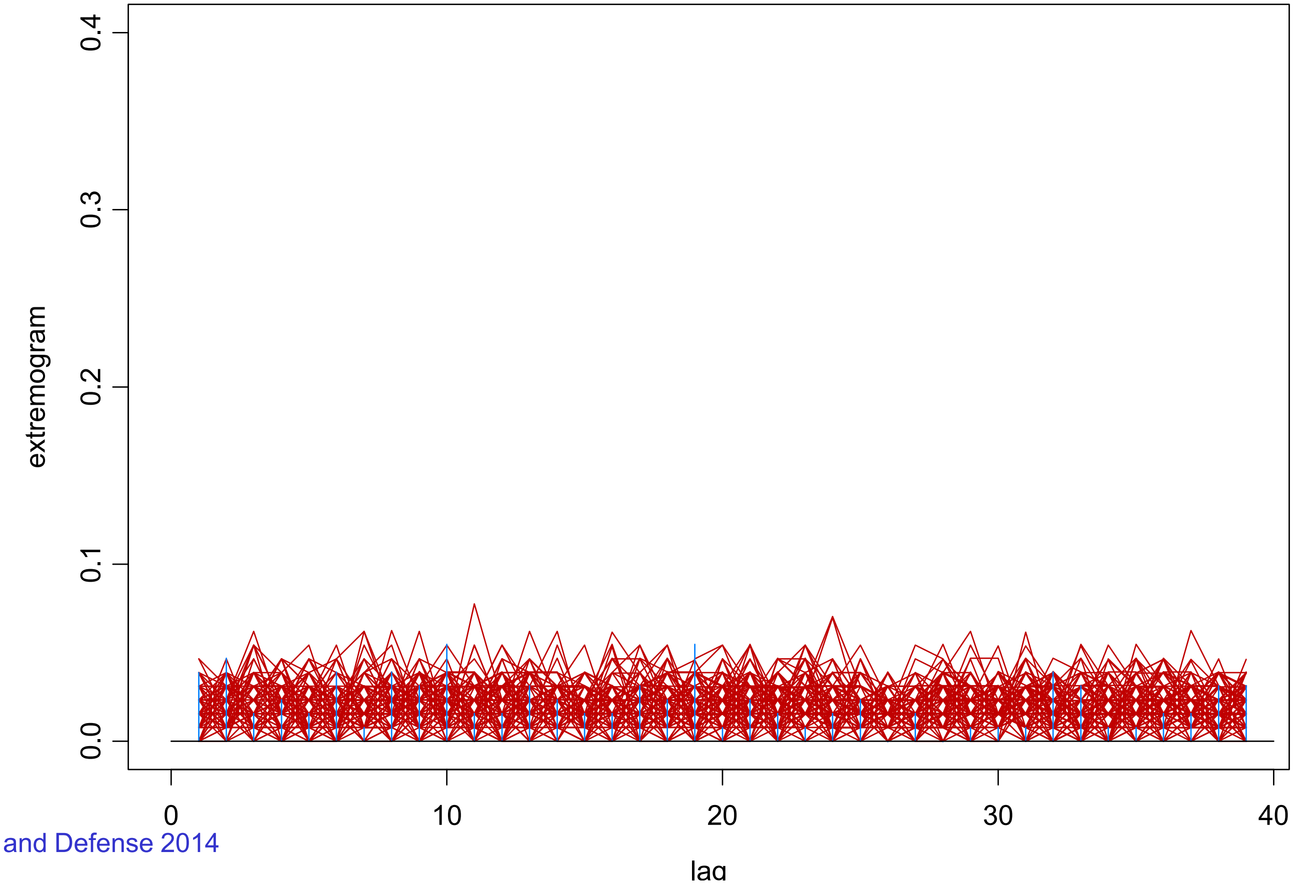
Devolatilizing S&P

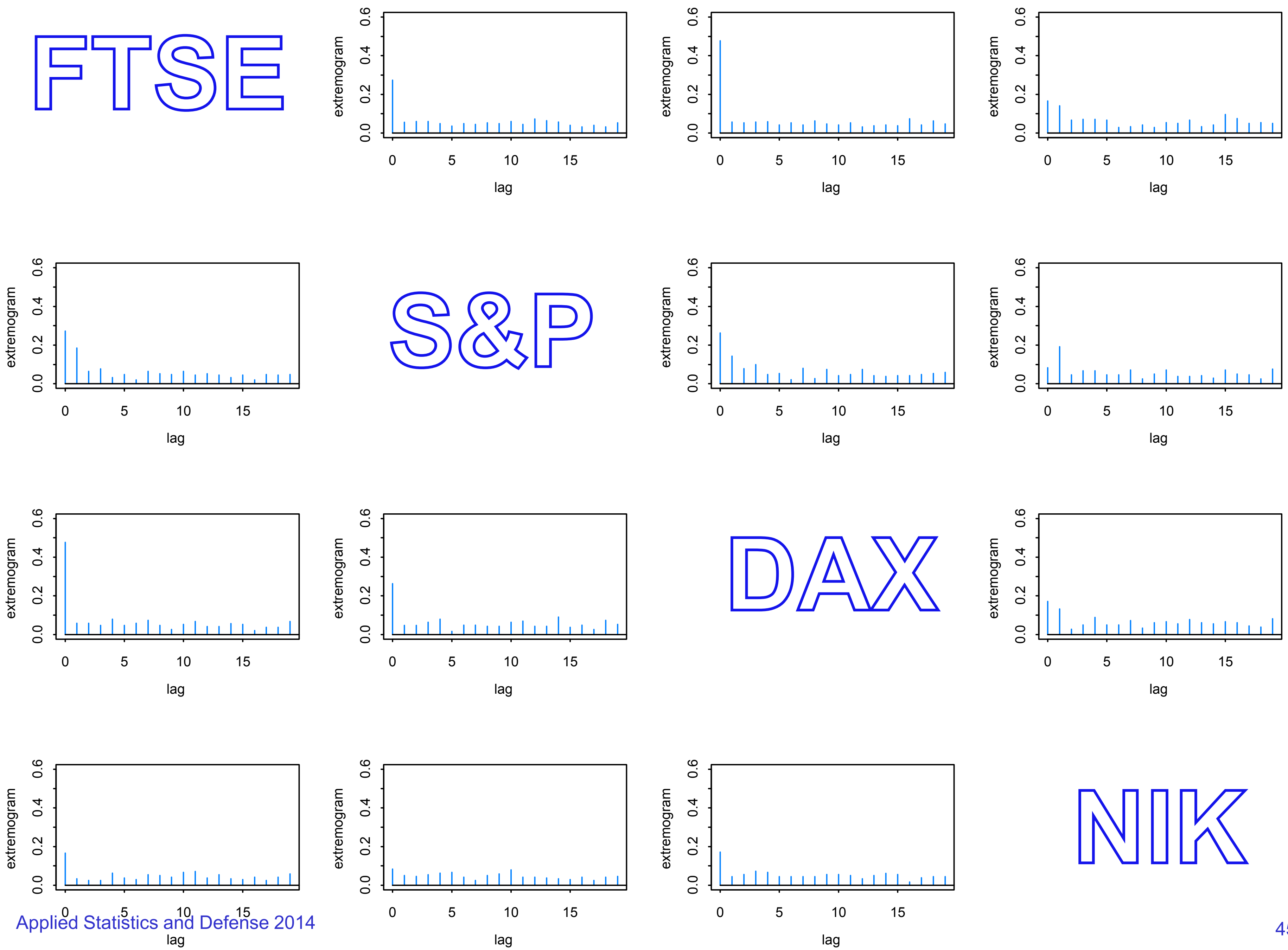
Extremogram for S&P: significant for large number of lags ~40+

Devolatilize S&P by fitting GARCH(1,1):

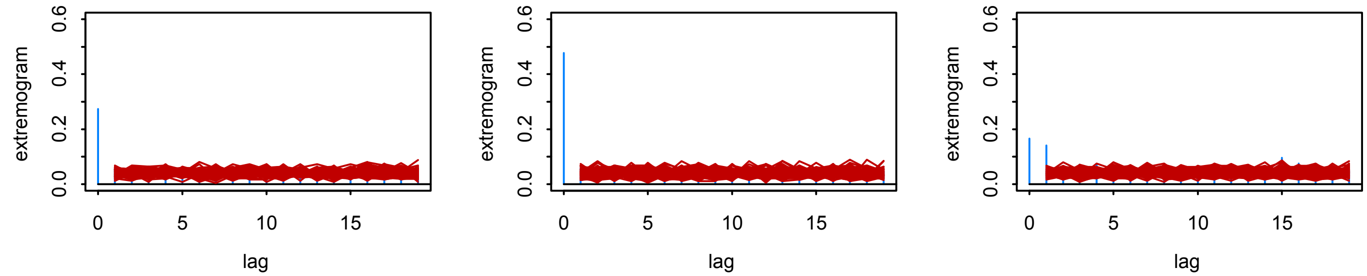
$$X_t = (6.28e-7 + .057X_{t-1}^2 + .939\sigma_{t-1}^2)^{1/2} Z_t,$$

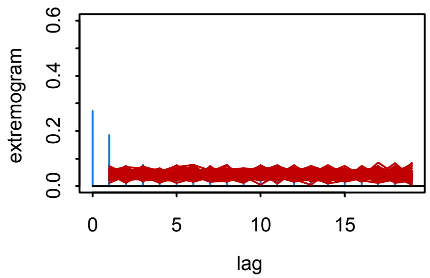
$\{Z_t\} \sim IID\ t(6.14),$



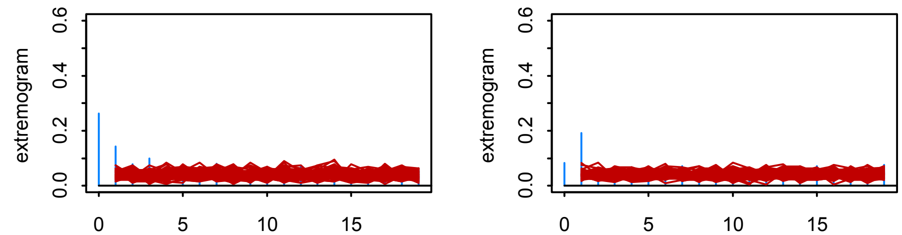

FTSE
S&P
DAX
NIK
extremogram
lag
0.0 0.2 0.4 0.6
0 5 10 15

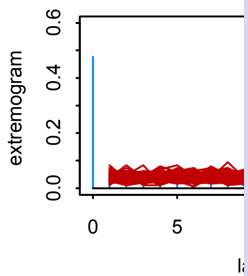
FTSE





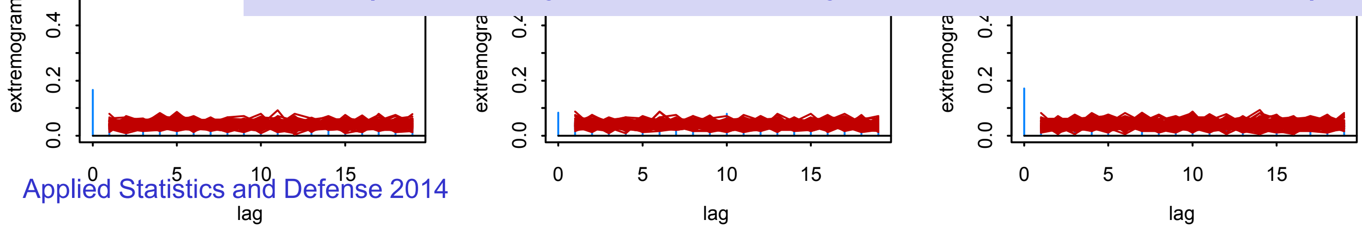
S&P





Second row (conditional on
lag 1. Given a significant ***le***
left tail event in FTSE, DAX,
- FTSE, DAX close at 1
so the ripple effect of S&
(possibly current day for FTSE and DAX).

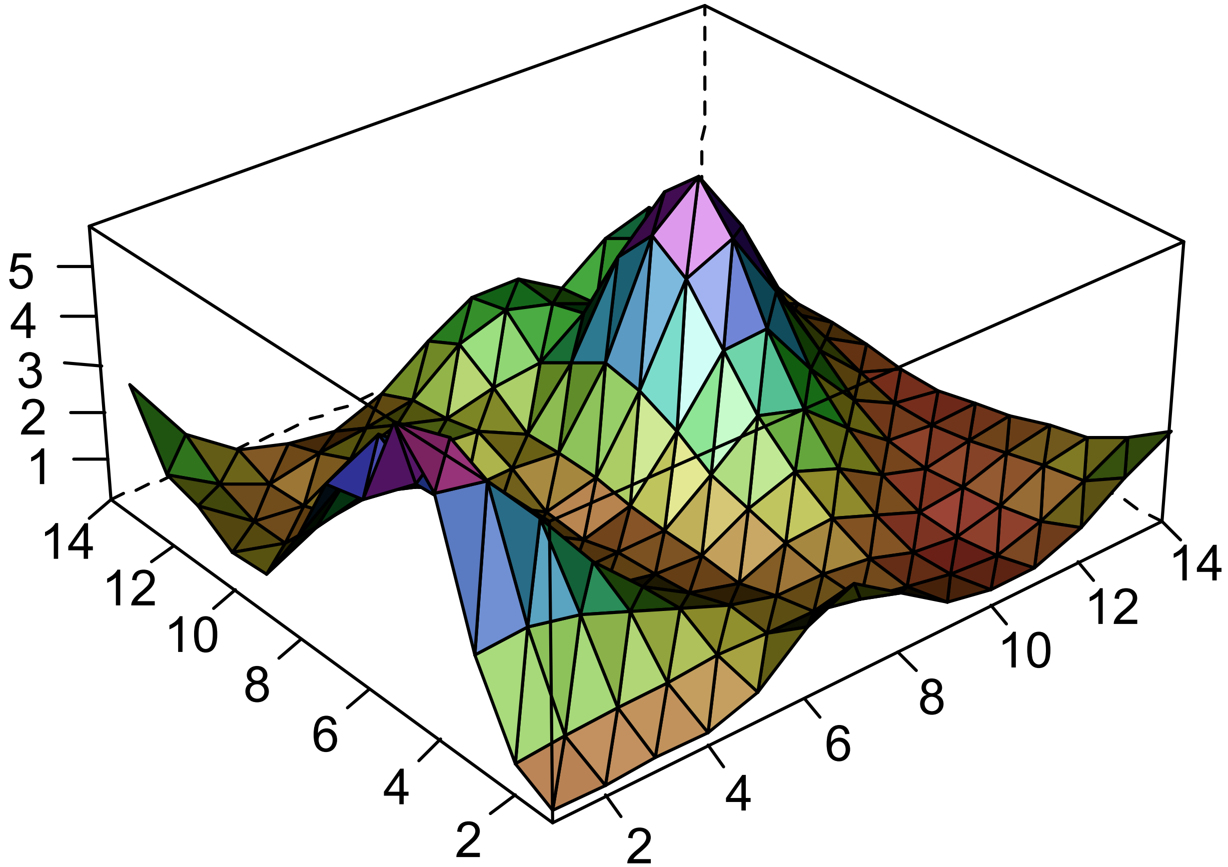
No symmetry at lag 1 (compare second column and second row).
-Extreme event in FTSE and DAX will have an impact the same day on S&P (not so much for Nikkei).



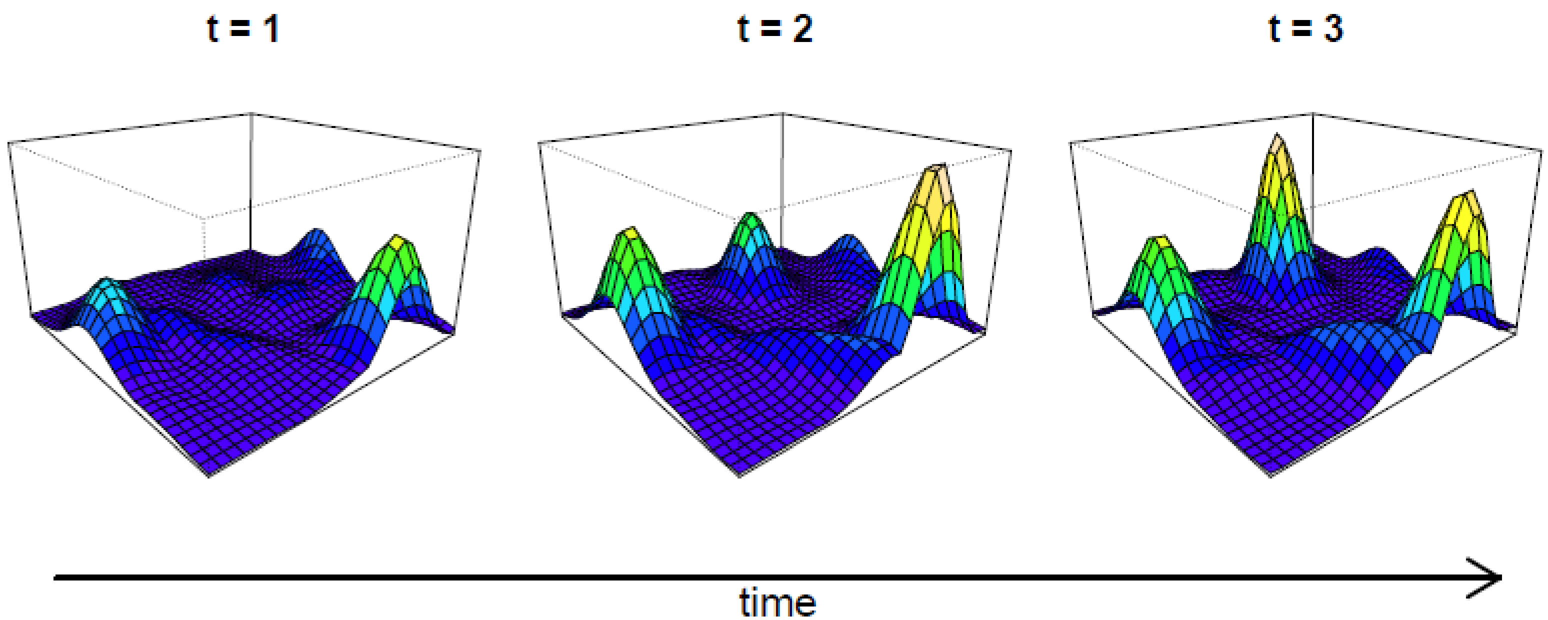
NIK

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 $s \in \mathbb{Z}^2$).



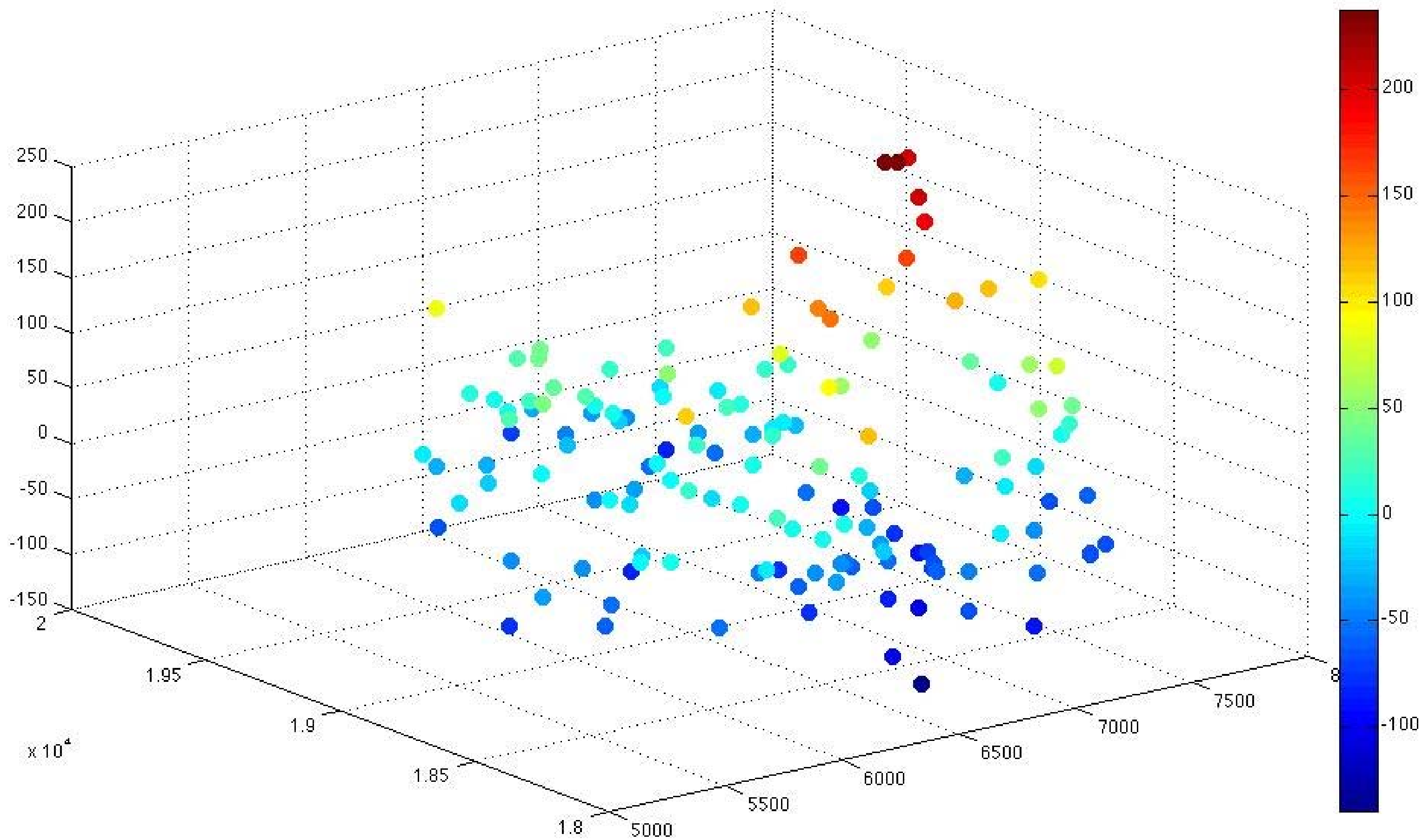
Extremal Dependence in Space and Time



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

Illustration with French Precipitation Data

Data from Naveau et al. (2009). Precipitation in Bourgogne of France; 51 year maxima of daily precipitation. Data has been adjusted for seasonality and orographic effects.

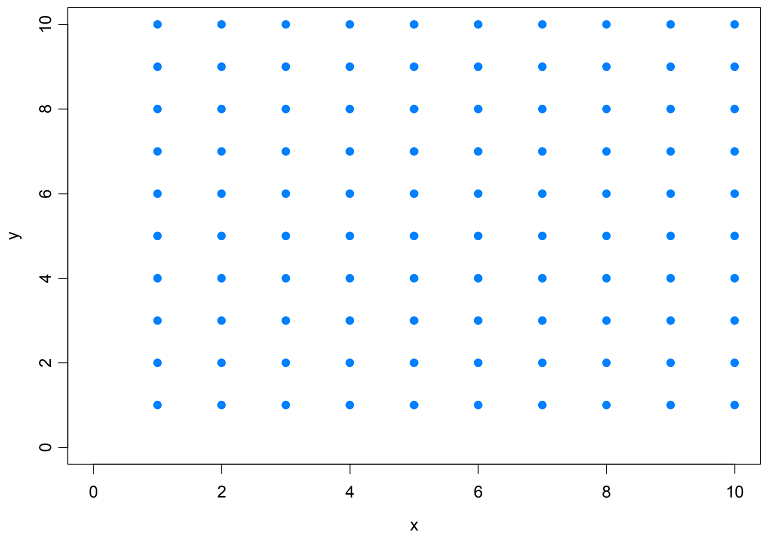


Lattice vs cont space

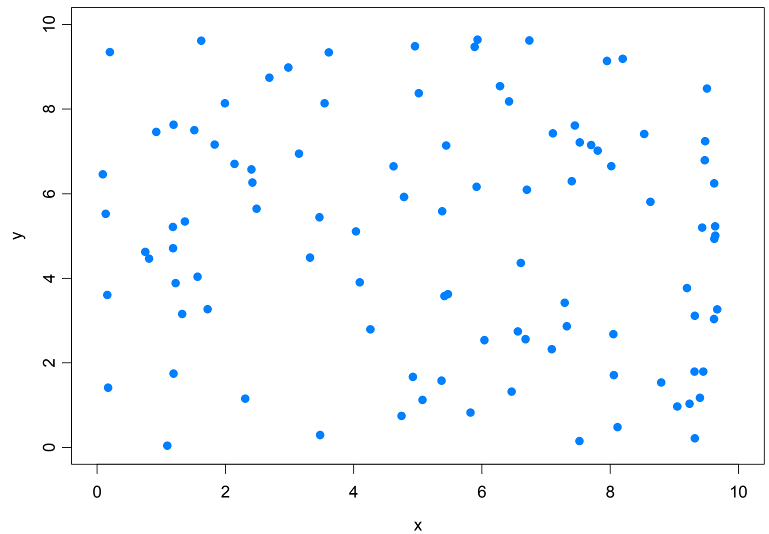
Setup: Let $X(s)$ be a RV stationary (isotropic?) spatial process defined on $s \in \mathbb{R}^2$ (or on a regular lattice $s \in \mathbb{Z}^2$). Consider the former—latter is more straightforward.

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

regular grid

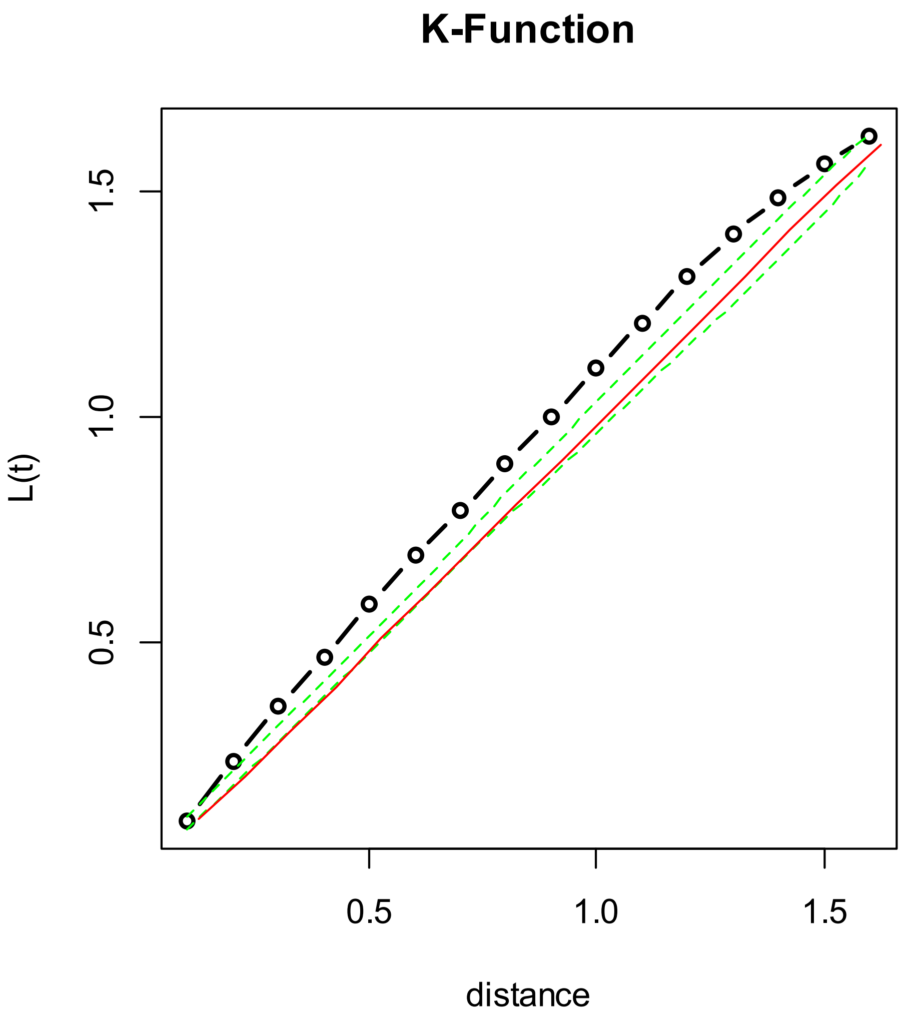


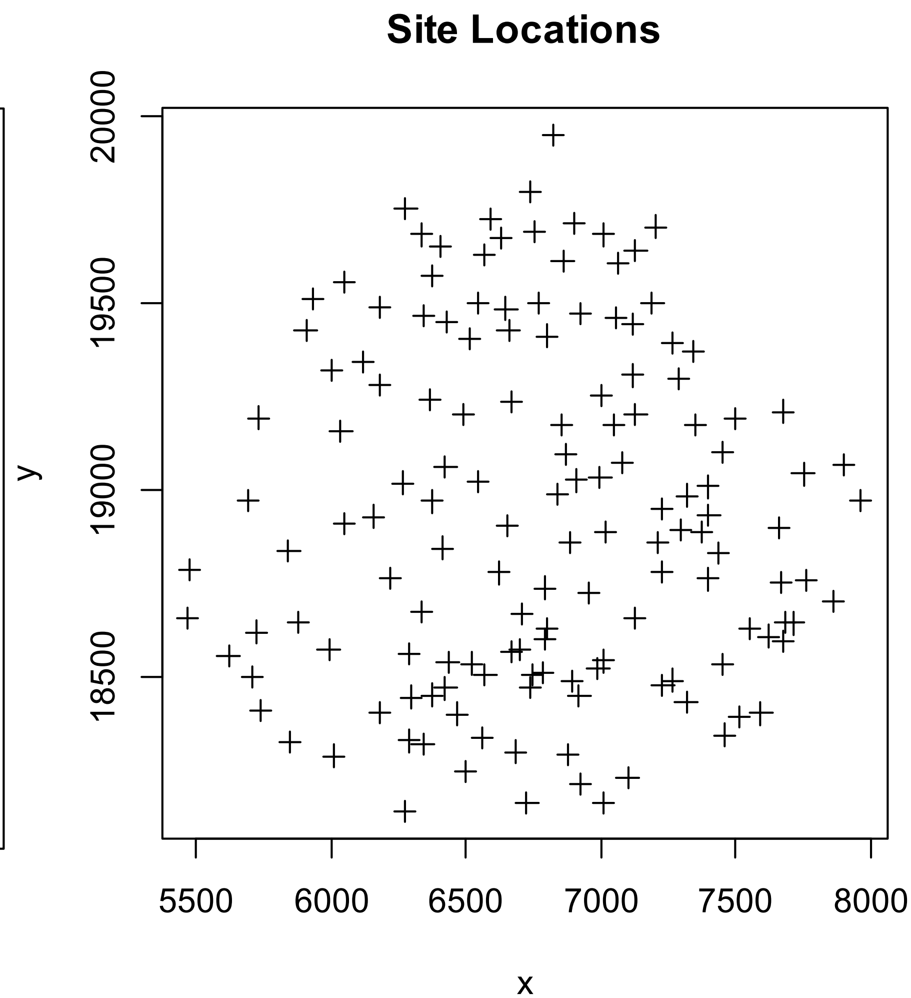
random pattern



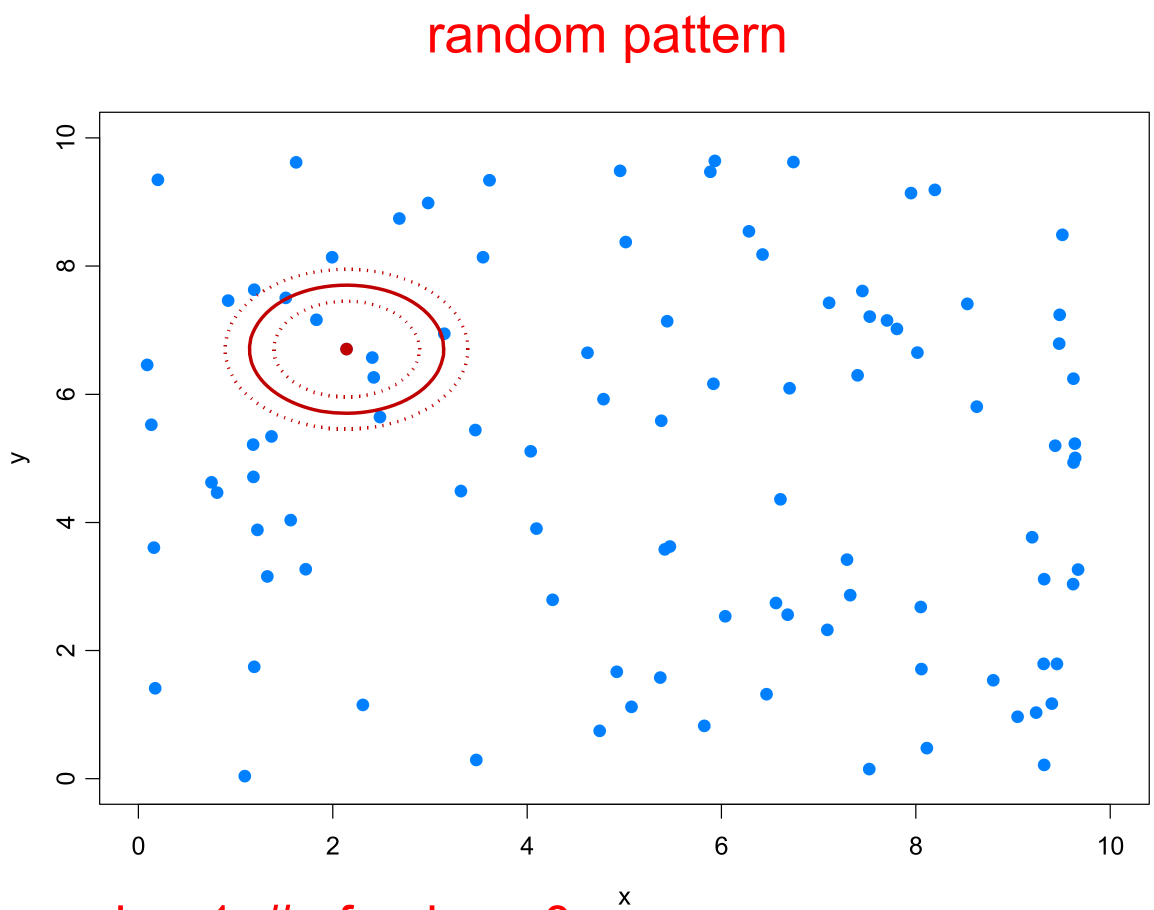
France Precipitation Data

regular or random?





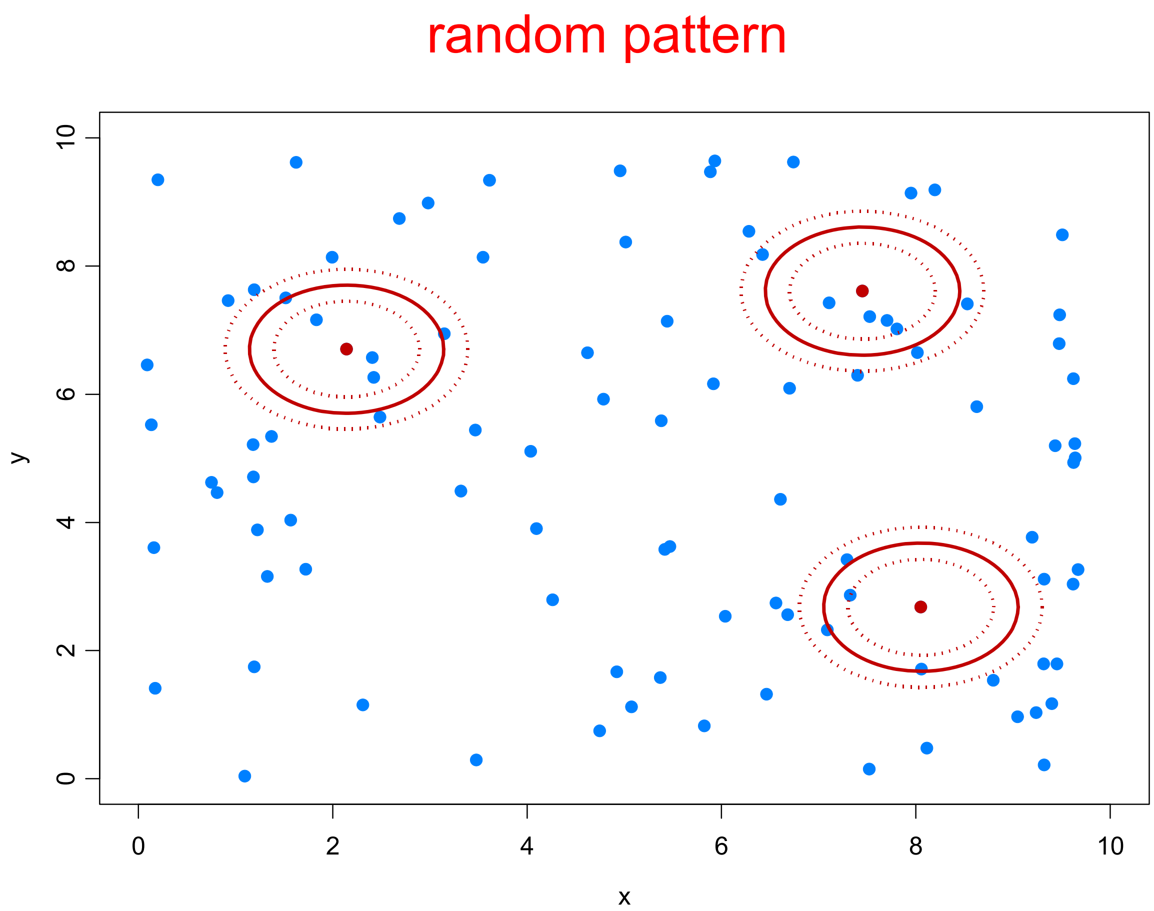
Random pattern



h = 1; # of pairs = 0

h = 1 $\pm$ .25

Random pattern



Note:

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

Estimating the extremogram--random pattern

Setup: Suppose we have observations, $X(s_1), \dots, X(s_N)$ at locations $s_1, \dots, s_{N_n}$ 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

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

Kernel estimate of ρ:

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

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

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.

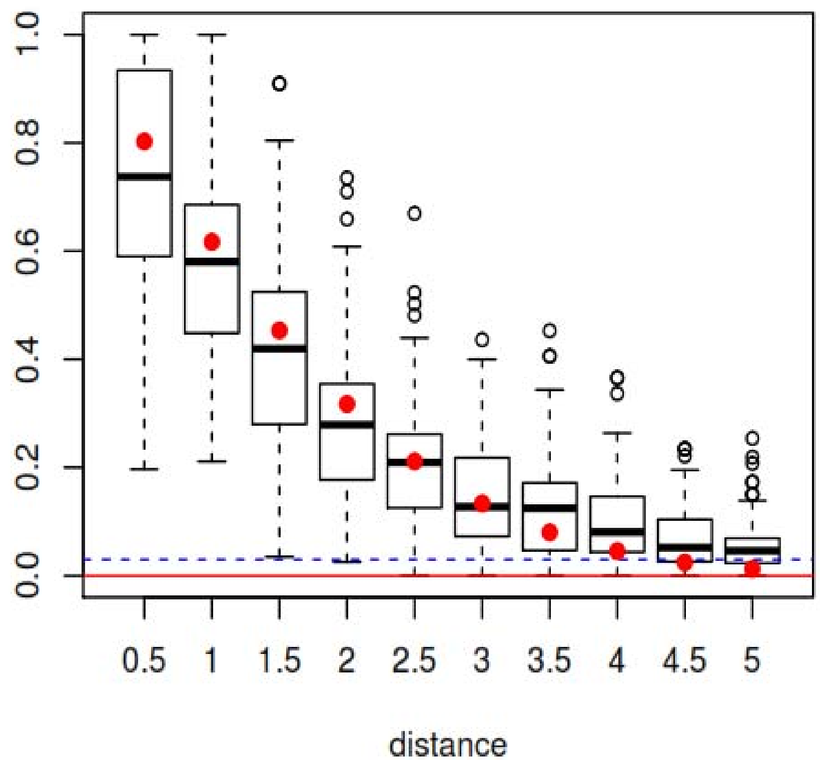
Limit theory:

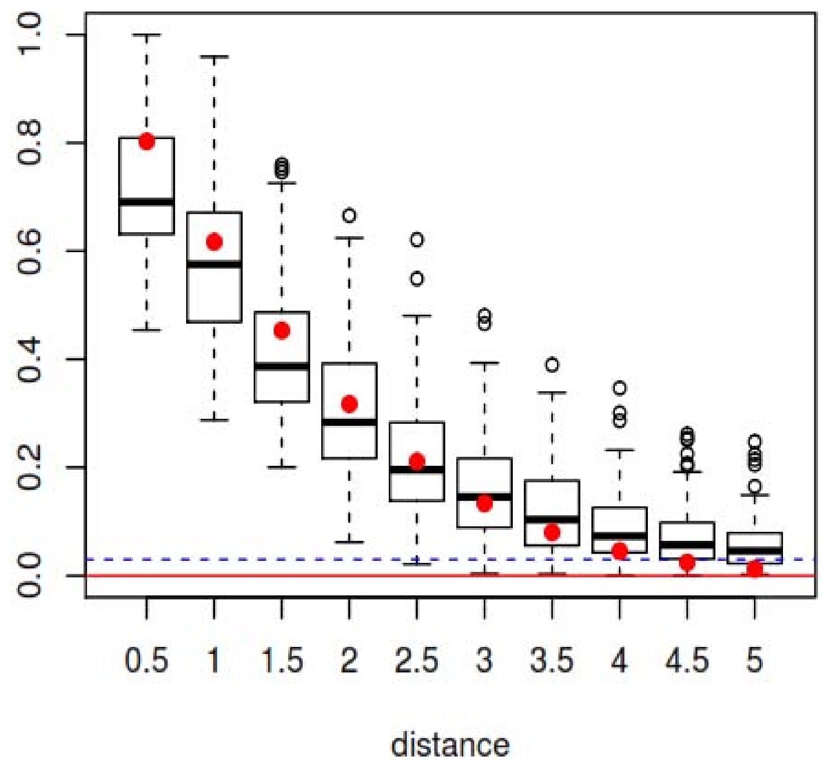
$$\left(\frac{|S_n|\lambda_n^2}{m_n}\right)^{\frac{1}{2}} \left(\hat{\rho}_{A,B}(h) - \rho_{A,B,m}(h)\right) \to N(0, \Sigma),$$

Simulation of spatial extremogram

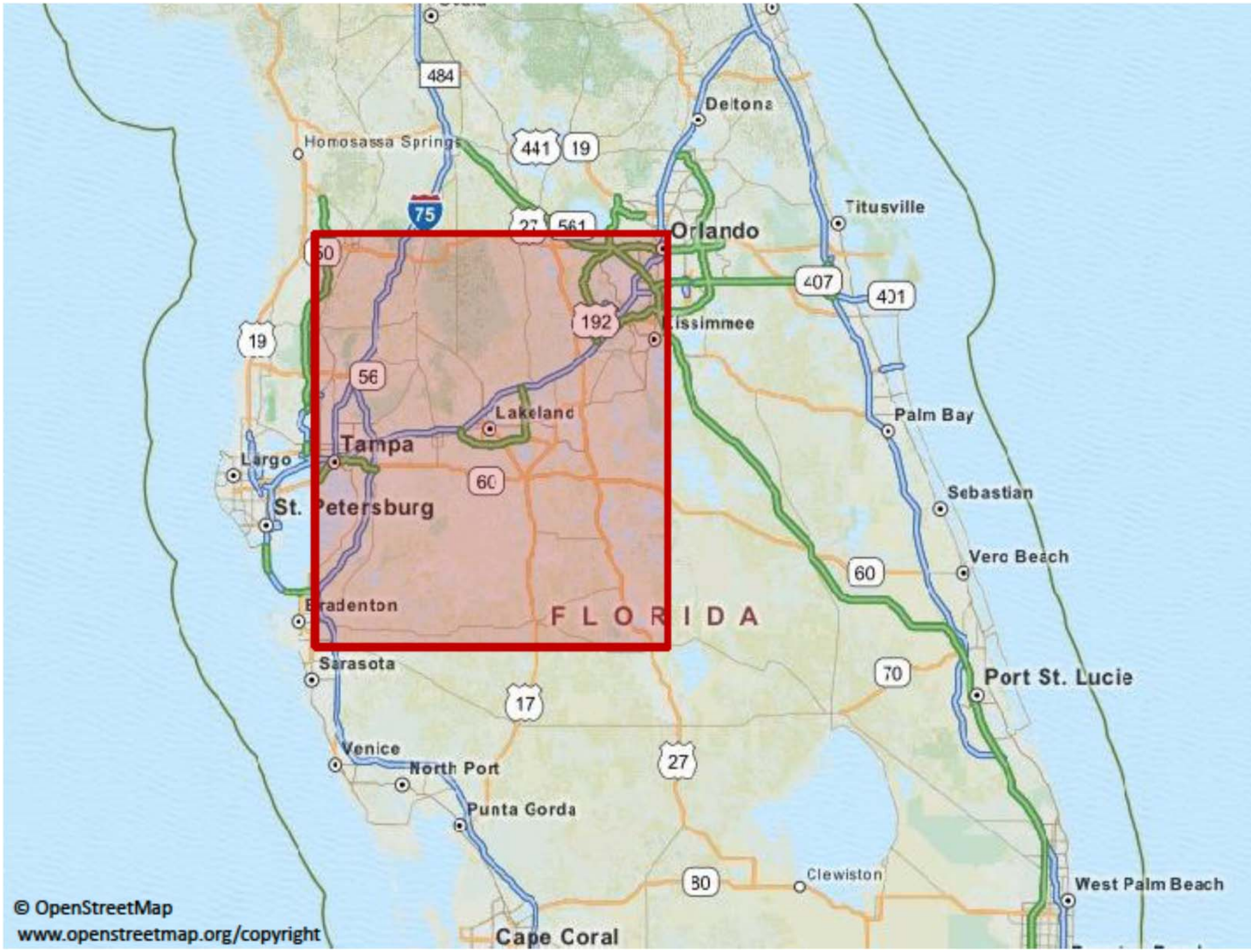
Box-plots based on 100 replications of BR on nonlattice

$\lambda_n = 1/\log n$ (left), $\lambda_n = 5/\log n$ (right)





Data Example: extreme rainfall in Florida



Data Example: extreme rainfall in Florida

Radar data:

Rainfall in inches measured in 15-minutes intervals at points of a spatial 2x2km grid.

Region:

120x120km, results in 60x60=3600 measurement points in space.

Take only wet season (June-September).

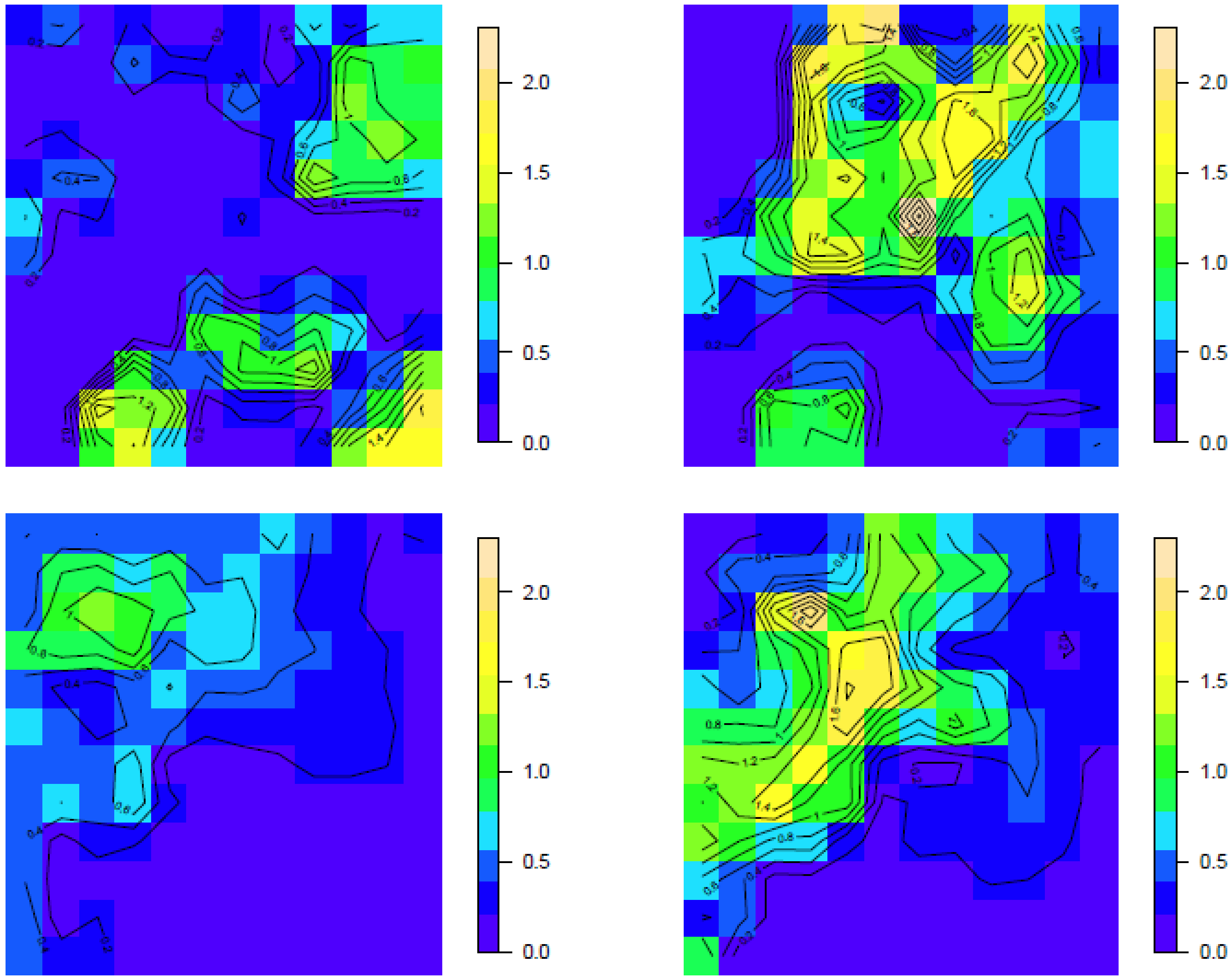
Block maxima in space: Subdivide in 10x10km squares, take maxima of rainfall over 25 locations in each square. This results in 12x12=144 spatial maxima.

Temporal domain: Analyze daily maxima and hourly accumulated rainfall observations.

Fit our extremal space-time model to daily/hourly maxima.

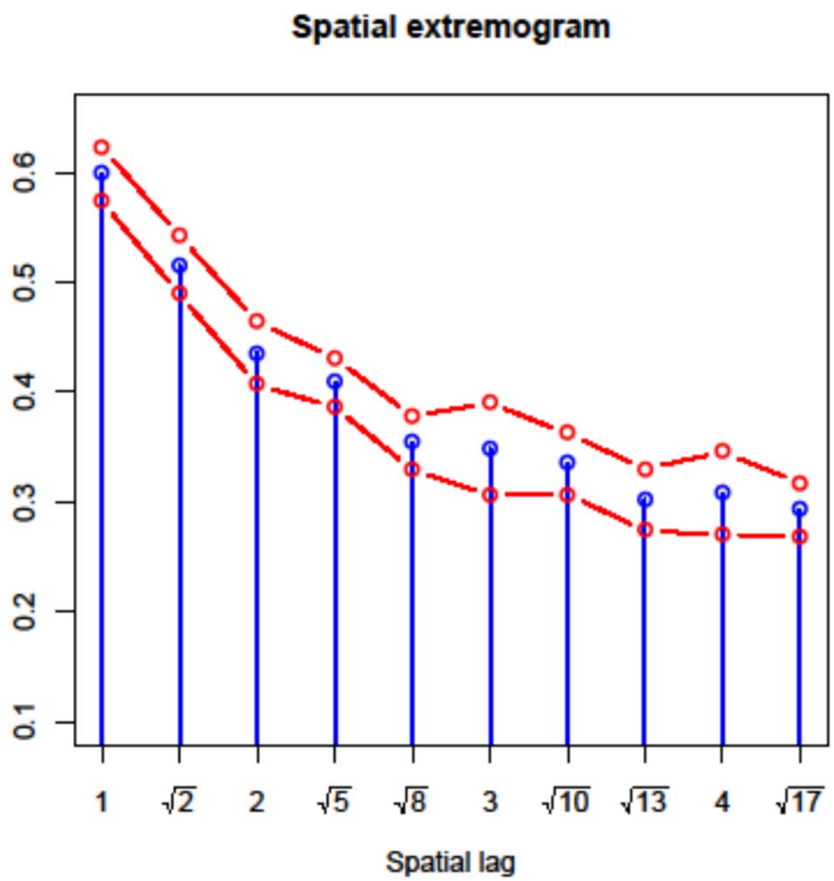
Data Example: extreme rainfall in Florida

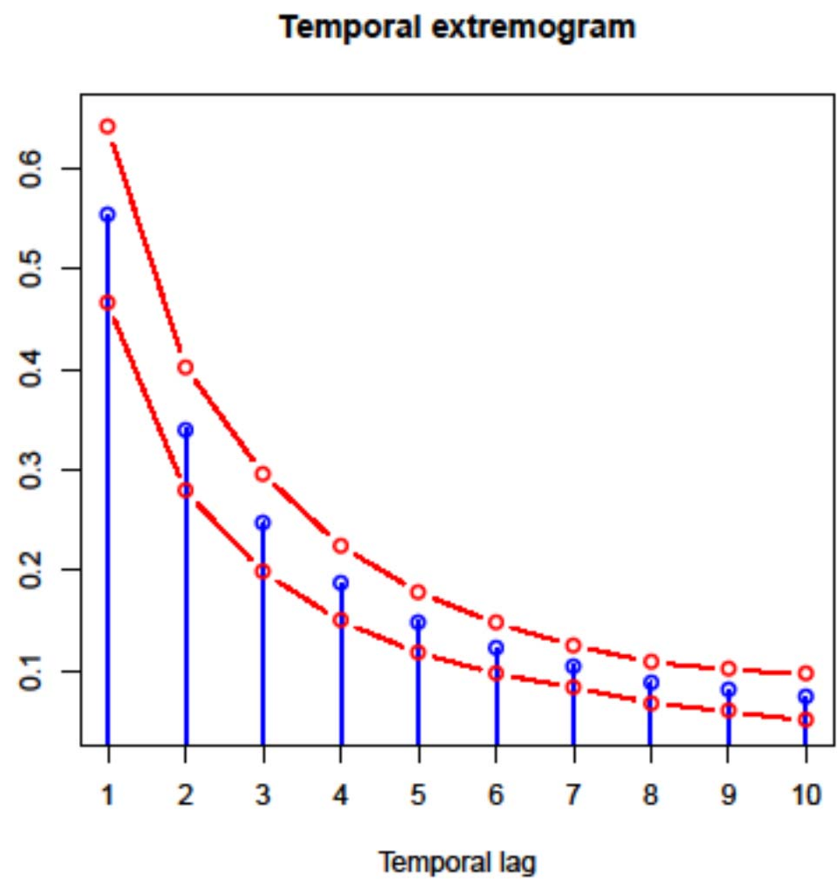
Hourly accumulated rainfall fields for four time points.



Data Example: extreme rainfall in Florida

Empirical extremogram in space (left) and time (right)

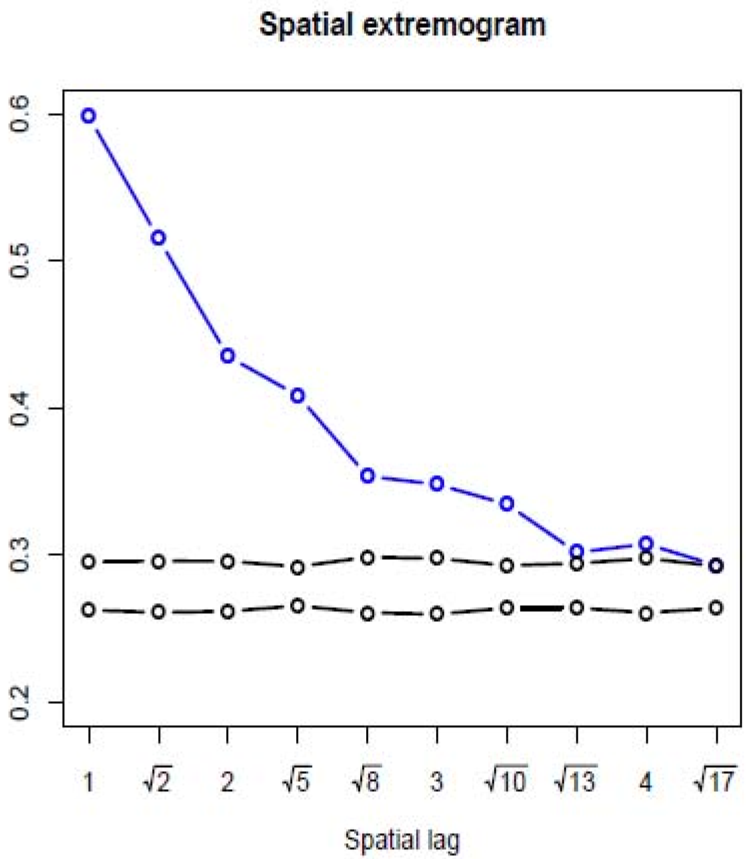


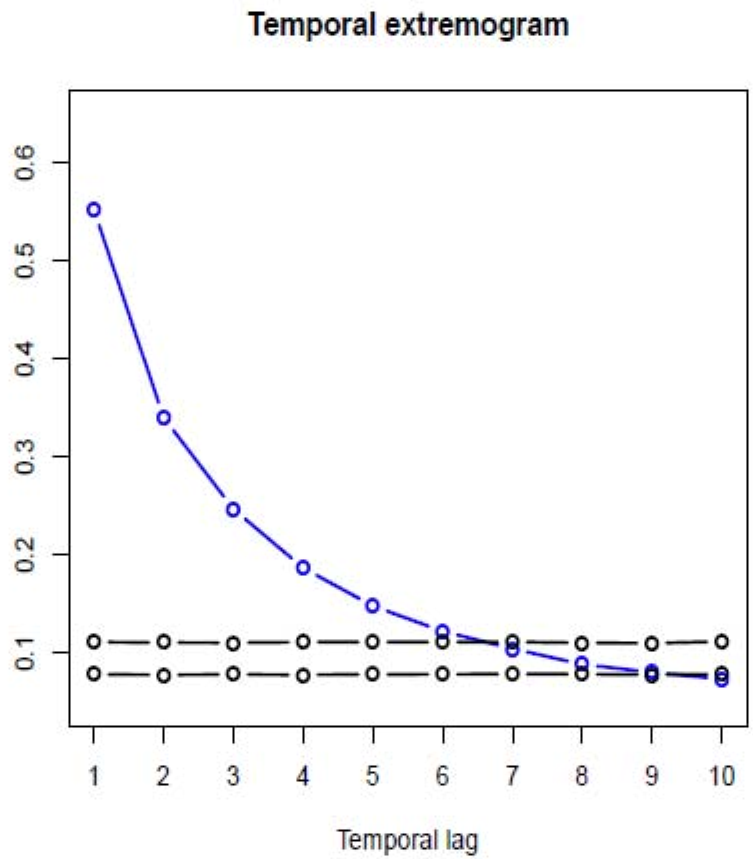


Data Example: extreme rainfall in Florida

Empirical extremogram in space (left) and time (right):

spatial indep for lags > 4; temporal indep for lags > 6.





Data Example: extreme rainfall in Florida

Computing conditional return maps.

Estimate $z_c(s,t)$ such that

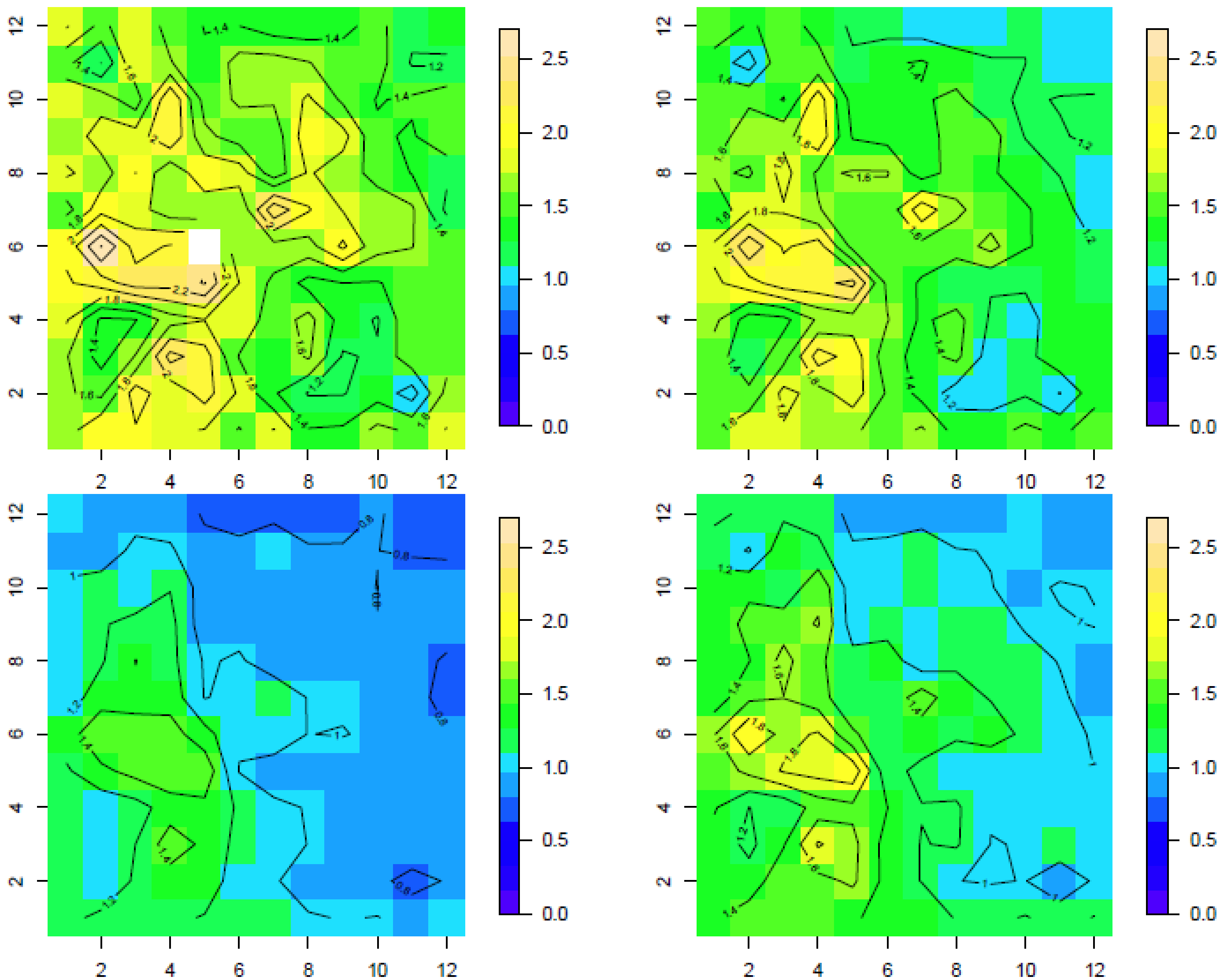
$$P(Z(s,t) > z_c(s,t) \mid Z(s^*,t^*) > z^*) = p_c,$$

where z^* satisfies $P(Z(s^*,t^*) > z^*) = p^*$ is pre-assigned.

A straightforward calculation shows that $z_c(s,t)$ must solve,

$$p_c = 1 - \frac{1}{p^*} \exp\left\{-\frac{1}{z_c(s,t)}\right\} + \frac{1}{p^*} F_{(BR)}(z_c(s,t), 1 - p^*)\,,$$

100-hour return maps ($p_c = .01$): $s^* = (5,6)$, time lags = 0,2,4,6 hours (left to right on top and then right to left on bottom), quantiles in inches.



Wrap-up

- *Extremogram* is another potential tool for estimating extremal dependence that may be helpful for discriminating between models on the basis of extreme value behavior.
- Permutation procedures are a *quick* and *clean* way to test for significant values in the extremogram and other statistics.
- *Bootstrapping* may prove useful for constructing CI's for the extremogram and also for assessing extremal dependence.
- The *Extremogram* can provide insight on extremal dependence between components in a multivariate time series.
- Extensions to *spatial* and *space-time* domains are possible.
- Theory for the *extremogram* has been developed for spatial data that are observed at *unequal spaced locations.*