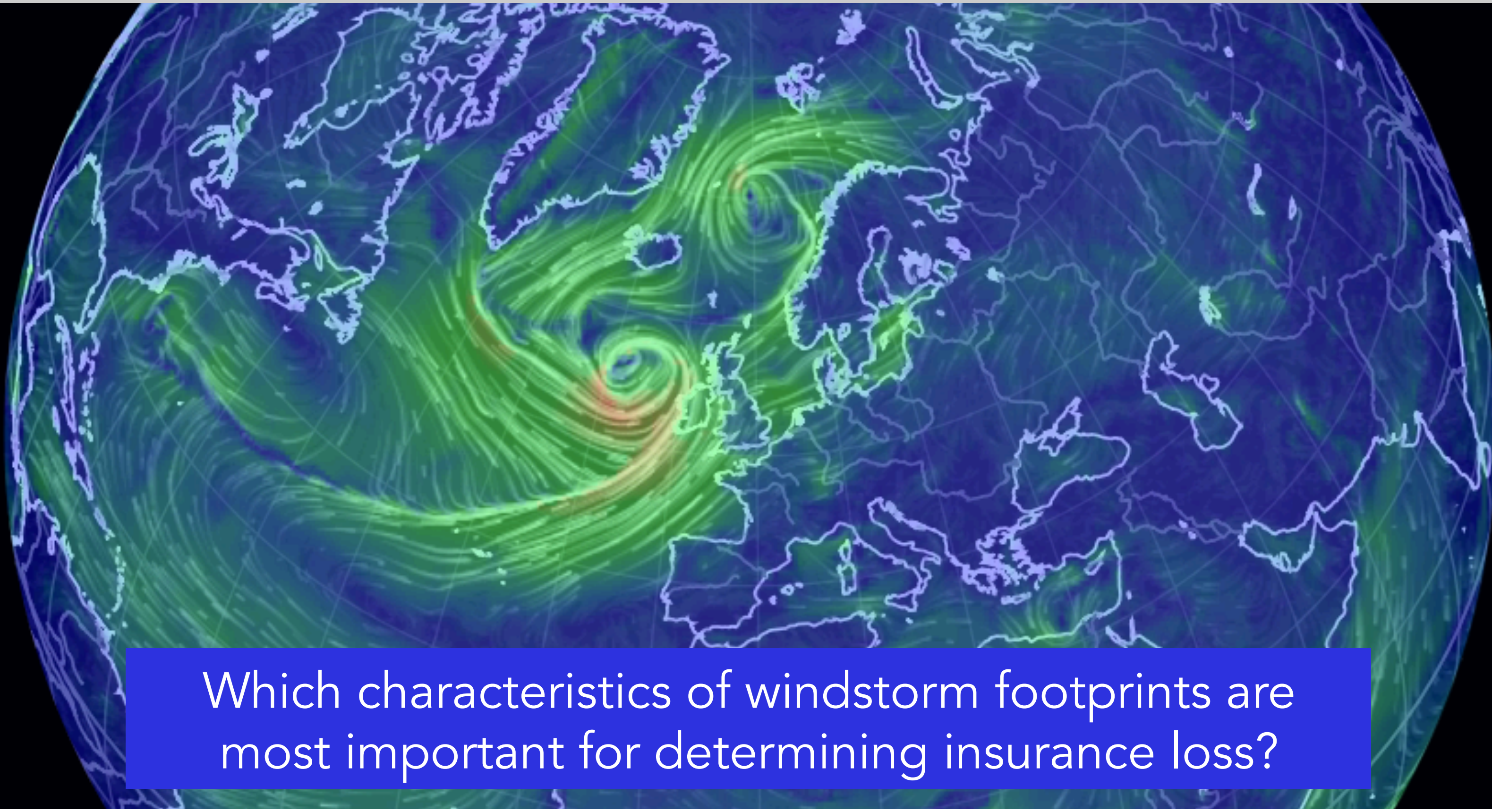
Modelling Dependence in Extremal Windstorm Footprints



Which characteristics of windstorm footprints are most important for determining insurance loss?

Laura Dawkins

Acknowledgements: David Stephenson, Ken Mylne, Ben Youngman, Chris Ferro

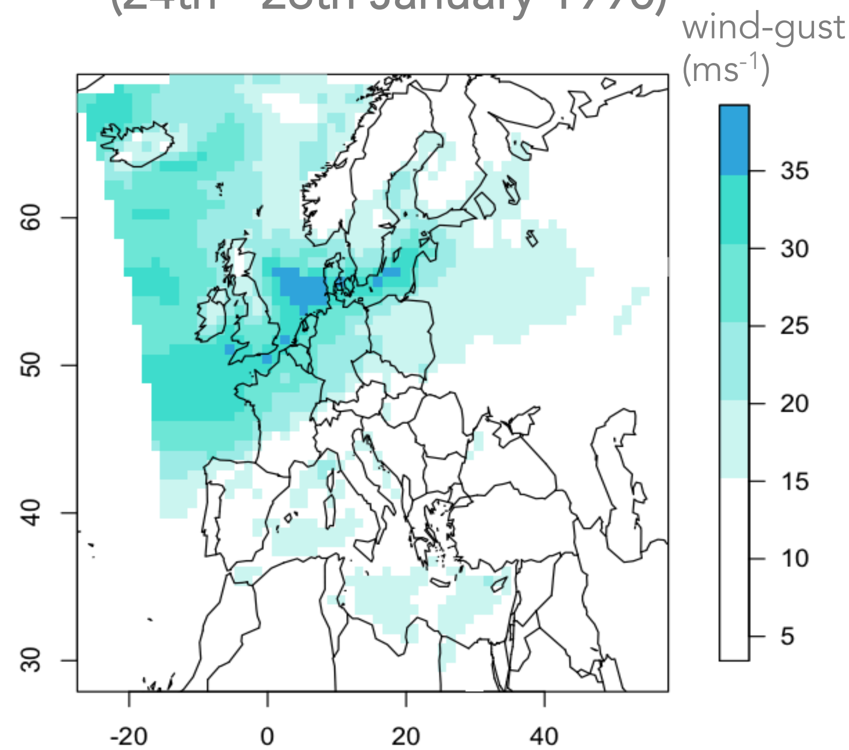
UNIVERSITY OF EXETER

lcd206@exeter.ac.uk

The Windstorm Footprint

The maximum 3-second wind-gust speed at each location in the 72 hour period covering the passage of the storm, centred on the time at which the maximum wind speed over land occurs

Footprint for historical windstorm Daria (24th - 26th January 1990)

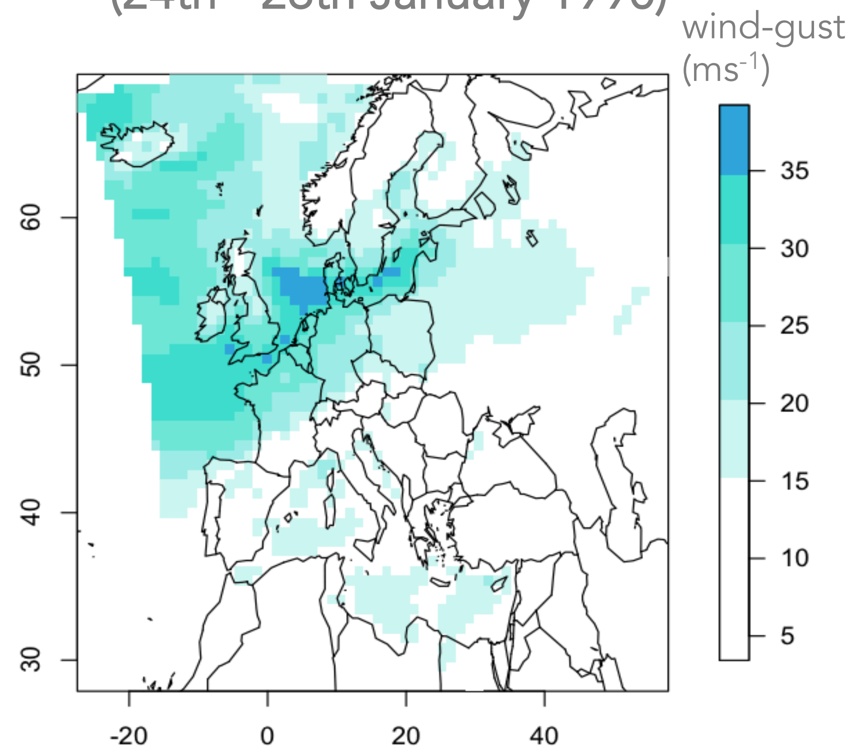


The Windstorm Footprint

The maximum 3-second wind-gust speed at each location in the 72 hour period covering the passage of the storm, centred on the time at which the maximum wind speed over land occurs

Created by dynamically downscaling ERA-Interim using the Met Office 25km resolution North Atlantic-European operational NWP model

Footprint for historical windstorm Daria (24th - 26th January 1990)

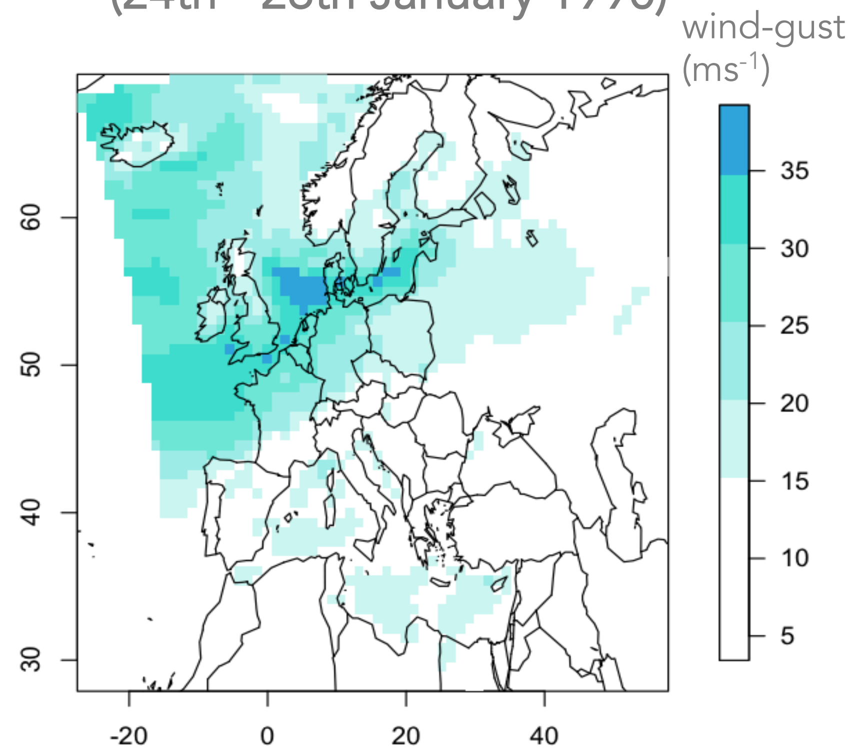


The Windstorm Footprint

The maximum 3-second wind-gust speed at each location in the 72 hour period covering the passage of the storm, centred on the time at which the maximum wind speed over land occurs

Created by dynamically downscaling ERA-Interim using the Met Office 25km resolution North Atlantic-European operational NWP model

Footprint for historical windstorm Daria (24th - 26th January 1990)

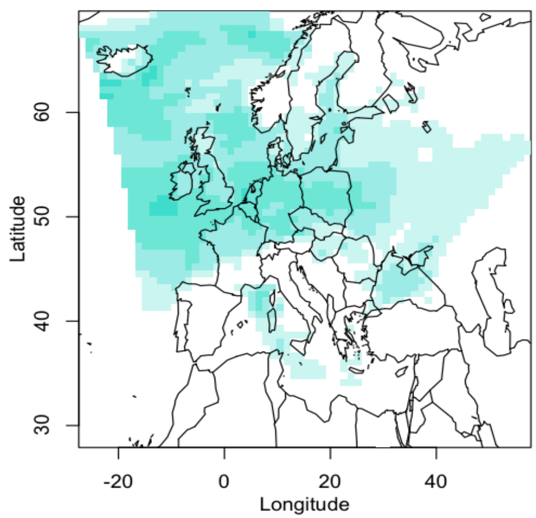


Data (kindly provided by Julia Roberts at the Met Office): footprints from the 5730 windstorms identified by Kevin Hodges' (Hodges, 1995) tracking algorithm, as occurring in the model domain in extended winters (October-March) 1979-2012. Statistical model is based on locations over land only.

Motivation

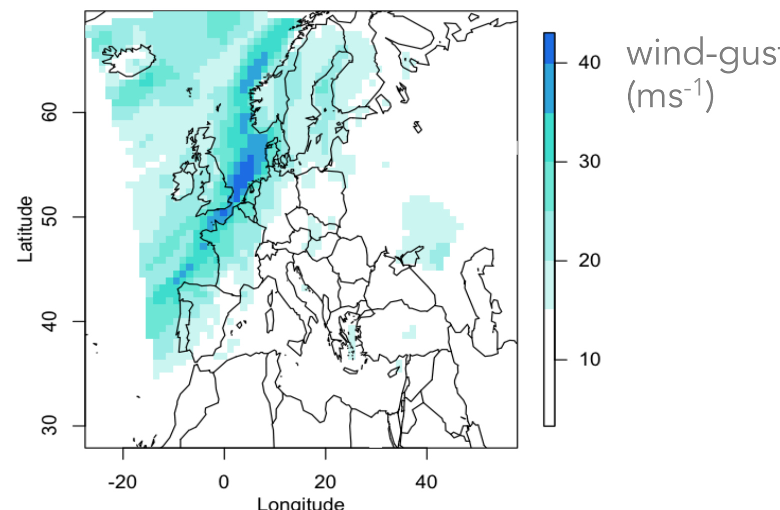
Windstorms causing extreme insurance loss have differing footprint characteristics, but which are most important for determining insurance loss?

Kyrill
17th-19th January 2007
Insurance Loss: $6.7bn



Low peak intensity
Large area
Large insurance loss

Great Storm '87
15th-17th October 1987
Insurance Loss: $6.3bn



High peak intensity
Small area
Large insurance loss

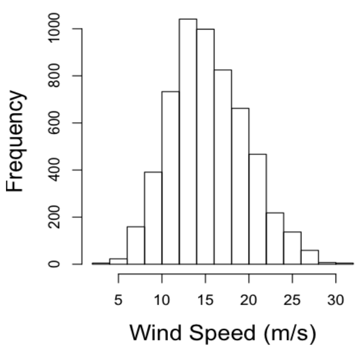
Develop a statistical model of the windstorm footprint, to explore the relationship between these characteristics and insurance loss using sensitivity analysis techniques

Bivariate Model

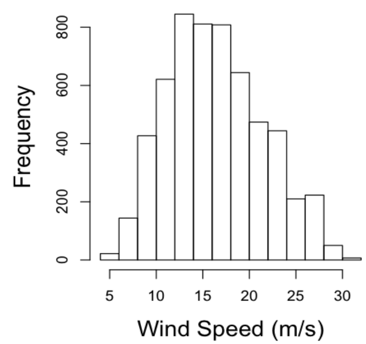
By Sklar's Theorem the joint distribution of footprint wind-gust speeds at two locations, denoted X_1 and X_2, can be though of in terms of…

…the marginal distributions of winds at each location …

London: F_1

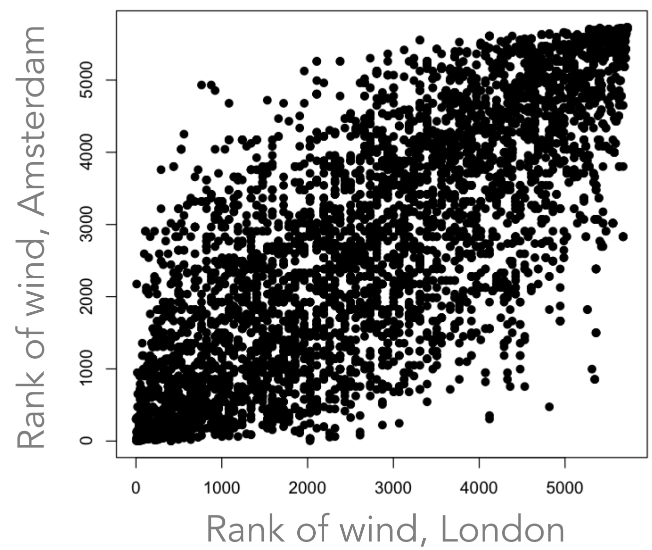


Amsterdam: F_2



… and their mutual dependence

Empirical Copula



$$\boxed{Pr(X_1 \leq x_1, X_2 \leq x_2) = F(x_1, x_2) = C(F_1(x_1), F_2(x_2))}$$

$$F_1(x_1) = Pr(X_1 \leq x_1)$$

Where C is often referred to as the copula function

Copulas

The copula contains all of the information about the joint distribution of X_1 and X_2 apart from the marginal distributions F_1 and F_2

Let $U = F_1(X_1) \quad V = F_2(X_2)$

Then, $C(u, v) = Pr(U \leq u, V \leq v)$

Copulas

The copula contains all of the information about the joint distribution of X_1 and X_2 apart from the marginal distributions F_1 and F_2

Let $U = F_1(X_1) \quad V = F_2(X_2)$

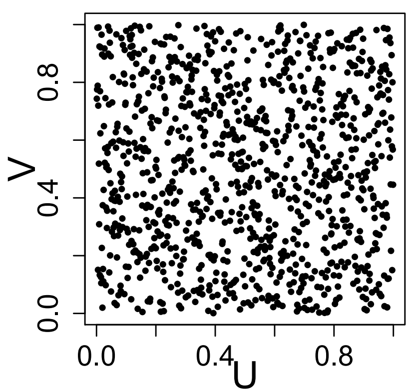
Then, $C(u, v) = Pr(U \leq u, V \leq v)$

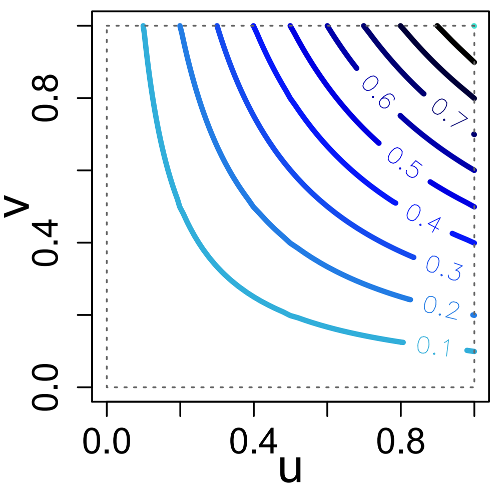
Independent

$$C(u, v) = uv$$

$$f(u, v) = Pr(U = u, V = v) = \frac{\partial^2 C}{\partial u \partial v} = 1$$

so (U,V) is
bivariate uniform





Copulas

The copula contains all of the information about the joint distribution of X_1 and X_2 apart from the marginal distributions F_1 and F_2

Let $$U = F_1(X_1) \quad V = F_2(X_2)$$

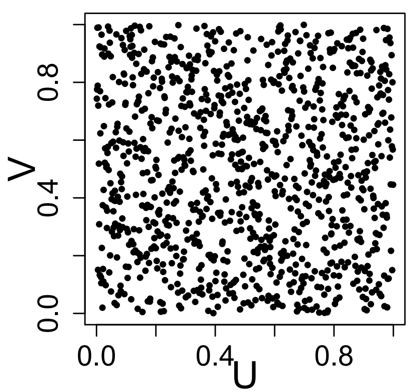
Then, $$C(u,v) = Pr(U \leq u, V \leq v)$$

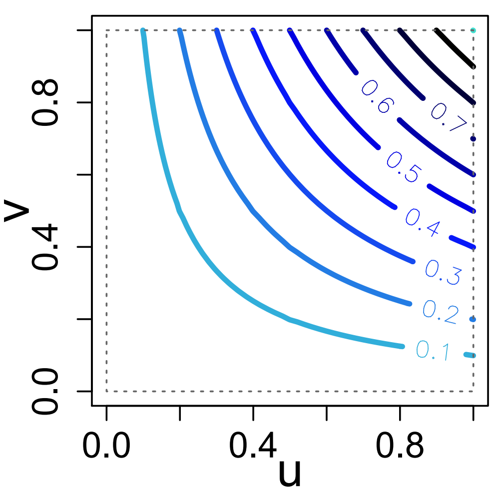
Independent

$$C(u,v) = uv$$

$$f(u,v) = Pr(U = u, V = v) = \frac{\partial^2 C}{\partial u \partial v} = 1$$

so (U,V) is bivariate uniform

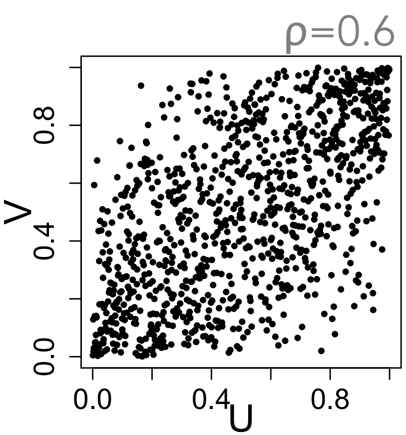


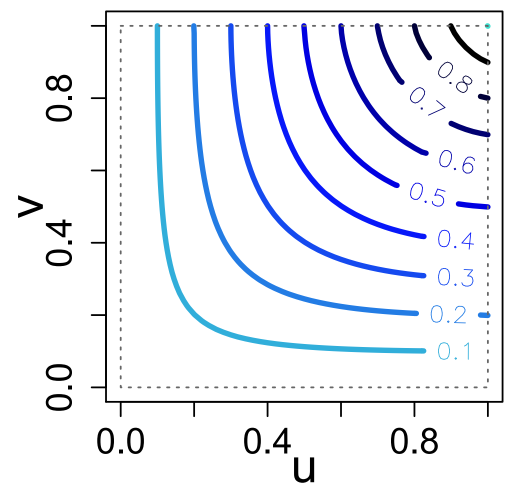


Gaussian

$$C(u,v) = \Phi_\Sigma(\Phi^{-1}(u), \Phi^{-1}(v))$$

$$(\Phi^{-1}(u), \Phi^{-1}(v))^T \sim MVN\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}\right)$$





Copulas

The copula contains all of the information about the joint distribution of X_1 and X_2 apart from the marginal distributions F_1 and F_2

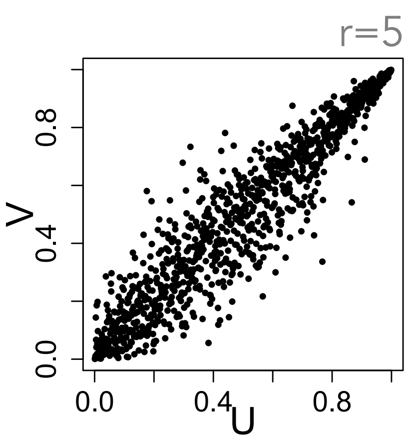
Let $U = F_1(X_1) \quad V = F_2(X_2)$

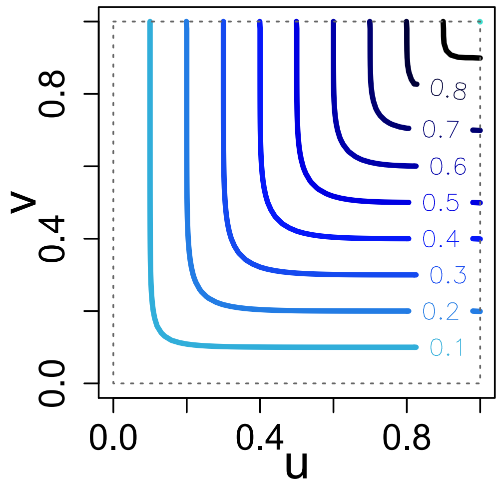
Then, $C(u,v) = Pr(U \leq u, V \leq v)$

Gumbel

$$C(u,v) = \exp(-((-\ln u)^r + (-\ln v)^r)^{1/r})$$

r=1, independent
r=∞, perfectly dependent





Copulas

The copula contains all of the information about the joint distribution of X_1 and X_2 apart from the marginal distributions F_1 and F_2

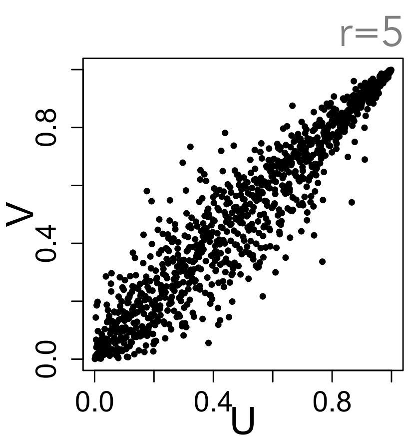
Let $U = F_1(X_1) \quad V = F_2(X_2)$

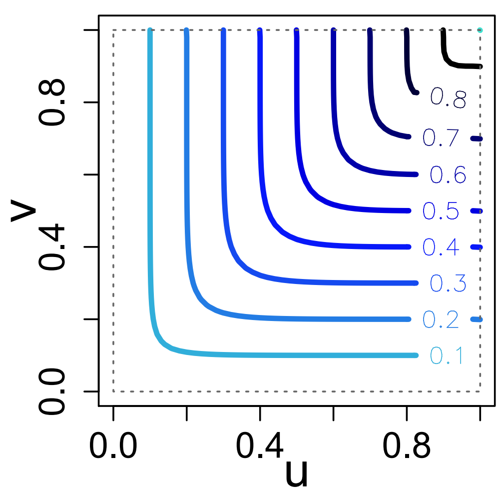
Then, $C(u,v) = Pr(U \leq u, V \leq v)$

Gumbel

$$C(u,v) = \exp(-((-\ln u)^r + (-\ln v)^r)^{1/r})$$

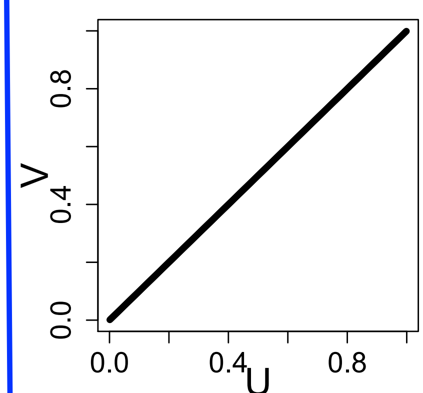
r=1, independent
r=∞, perfectly dependent

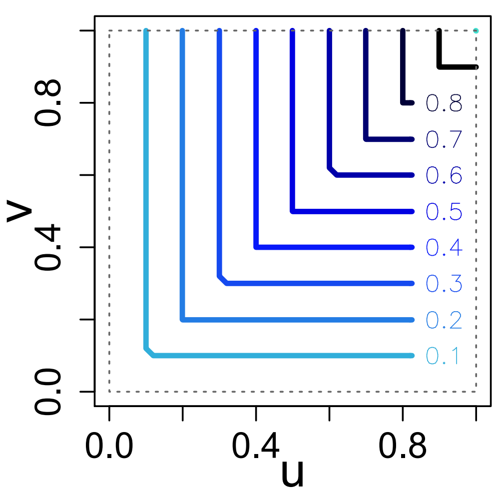




Perfectly Dependent

$$C(u,v) = min(u,v)$$



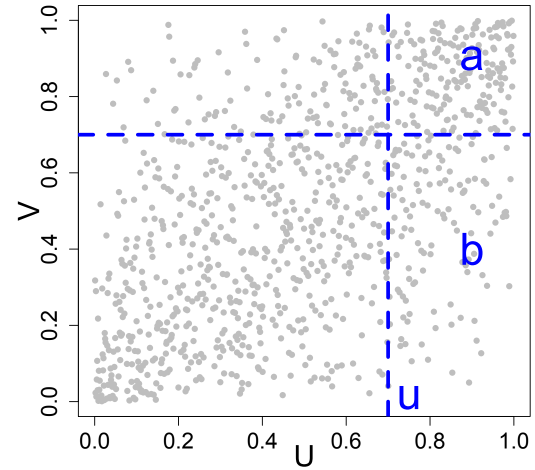


Which Copula?

- Different copulas assume different extremal dependence properties
- Explore these properties in the data to select the most appropriate copula

- Extremal dependence coefficient: $\chi(u) = Pr(V > u | U > u)$

$$
\begin{aligned}
\chi(u) &= Pr(V > u | U > u) \\
&= Pr(V > u, U > u)/Pr(U > u) \\
&= a/(a + b)
\end{aligned}
$$



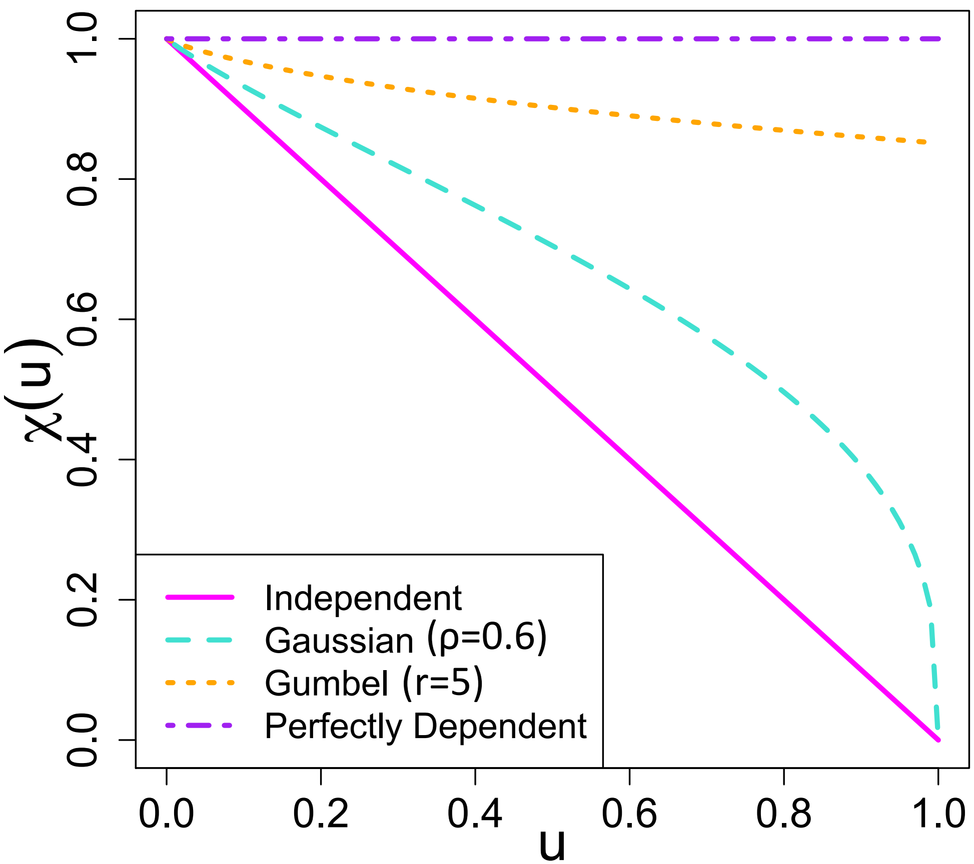
Which Copula?

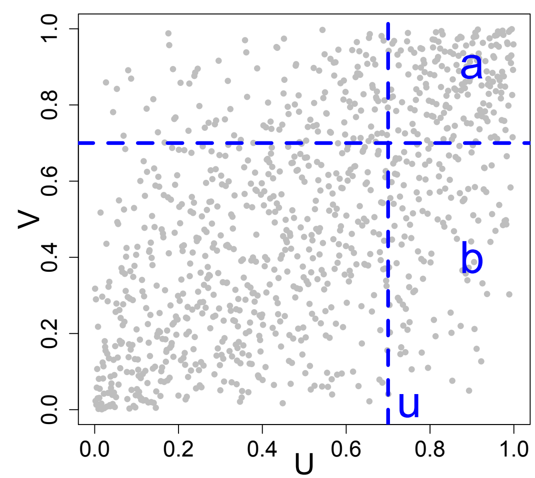
- Different copulas assume different extremal dependence properties
- Explore these properties in the data to select the most appropriate copula

- Extremal dependence coefficient: $\chi(u) = Pr(V > u | U > u)$

$$= Pr(V > u, U > u)/Pr(U > u)$$

$$= a/(a+b)$$





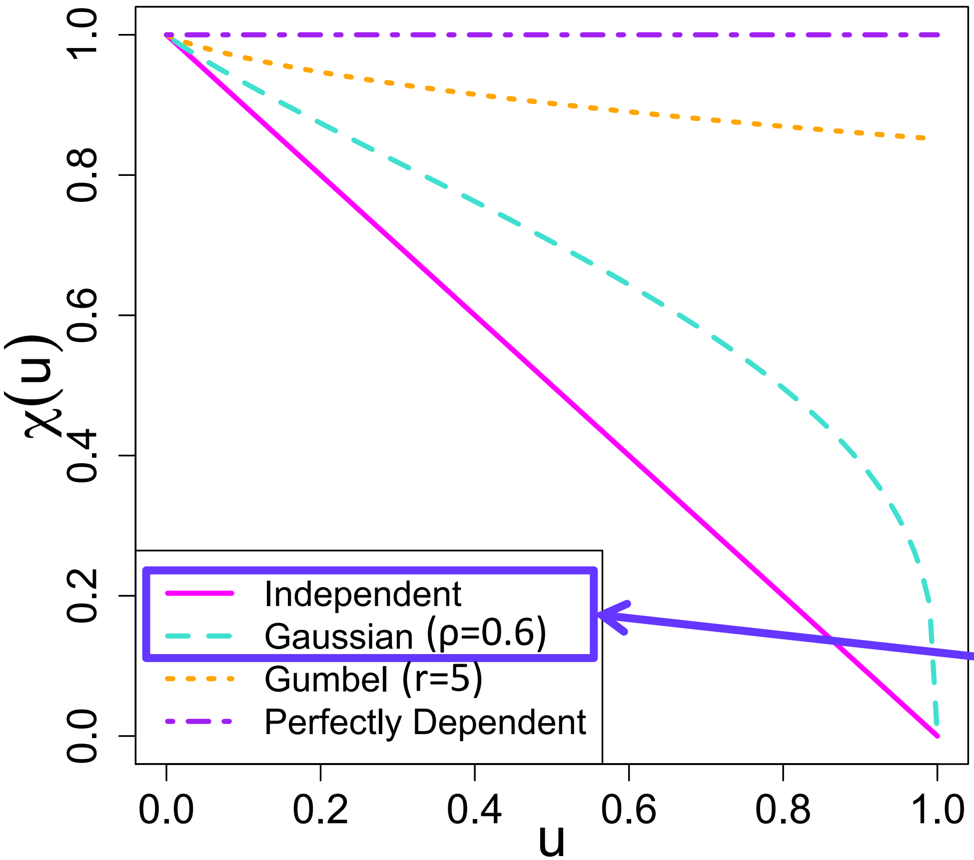
Which Copula?

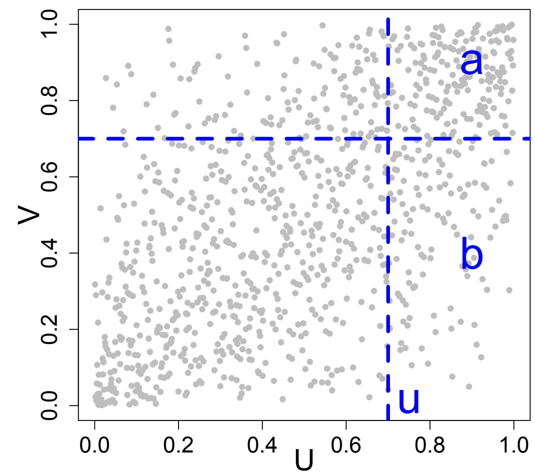
- Different copulas assume different extremal dependence properties
- Explore these properties in the data to select the most appropriate copula

- Extremal dependence coefficient: $\chi(u) = Pr(V > u | U > u)$

$$= Pr(V > u, U > u)/Pr(U > u)$$

$$= a/(a+b)$$





$\lim_{u \to 1} \chi(u) = 0$ - Asymptotic independence

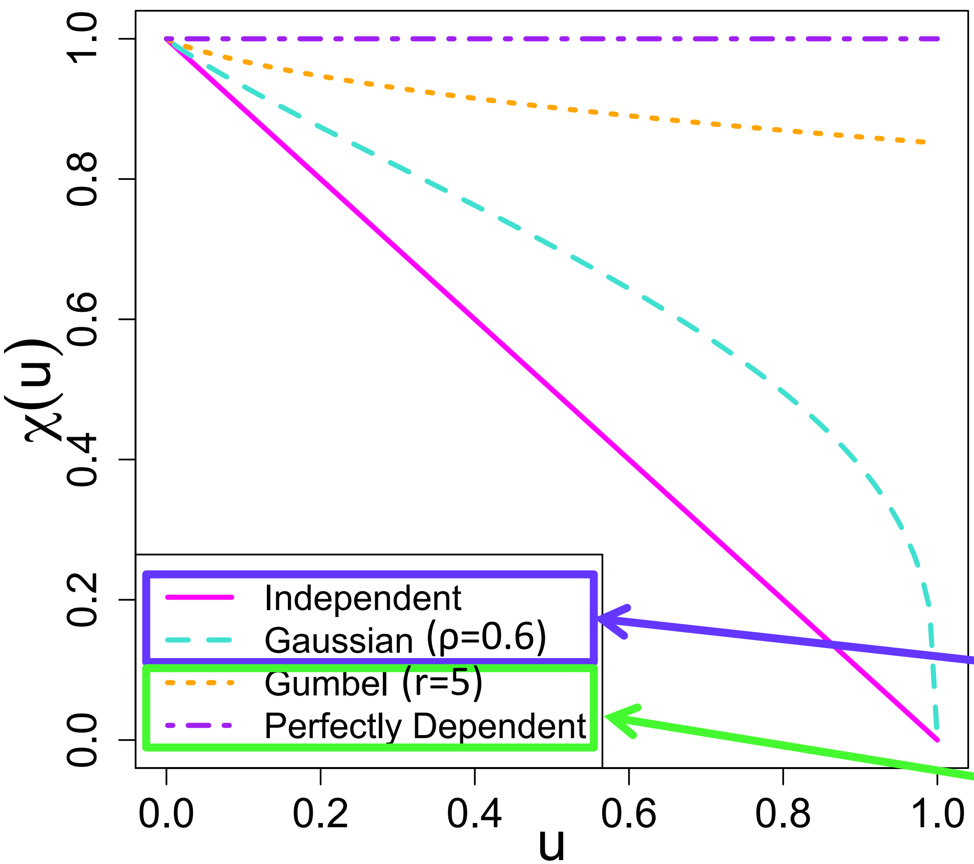
Which Copula?

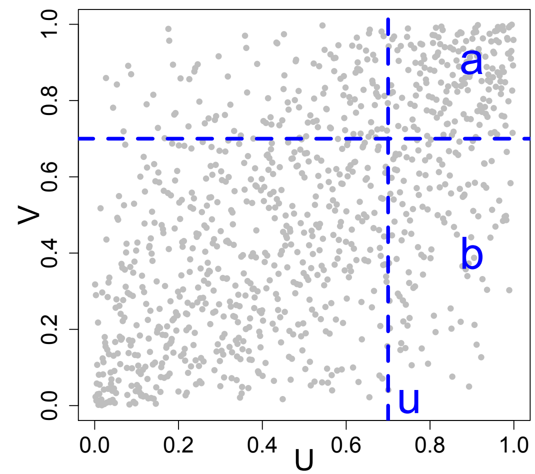
- Different copulas assume different extremal dependence properties
- Explore these properties in the data to select the most appropriate copula

- Extremal dependence coefficient: $\chi(u) = Pr(V > u | U > u)$

$$= Pr(V > u, U > u)/Pr(U > u)$$

$$= a/(a+b)$$





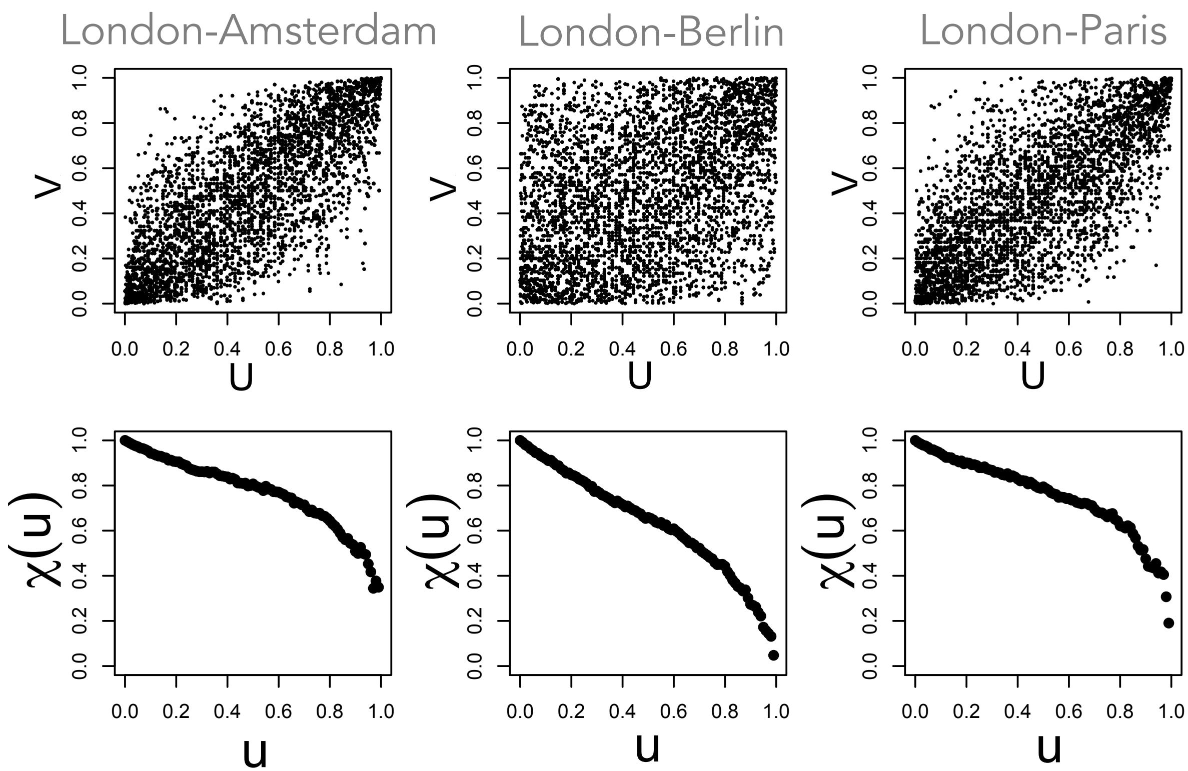
$\lim_{u \to 1} \chi(u) = 0$ - Asymptotic independence

$\lim_{u \to 1} \chi(u) > 0$ - Asymptotic dependence

Which Copula?

- How does the empirical χ(u) behave for footprint wind-gusts at pairs of locations?

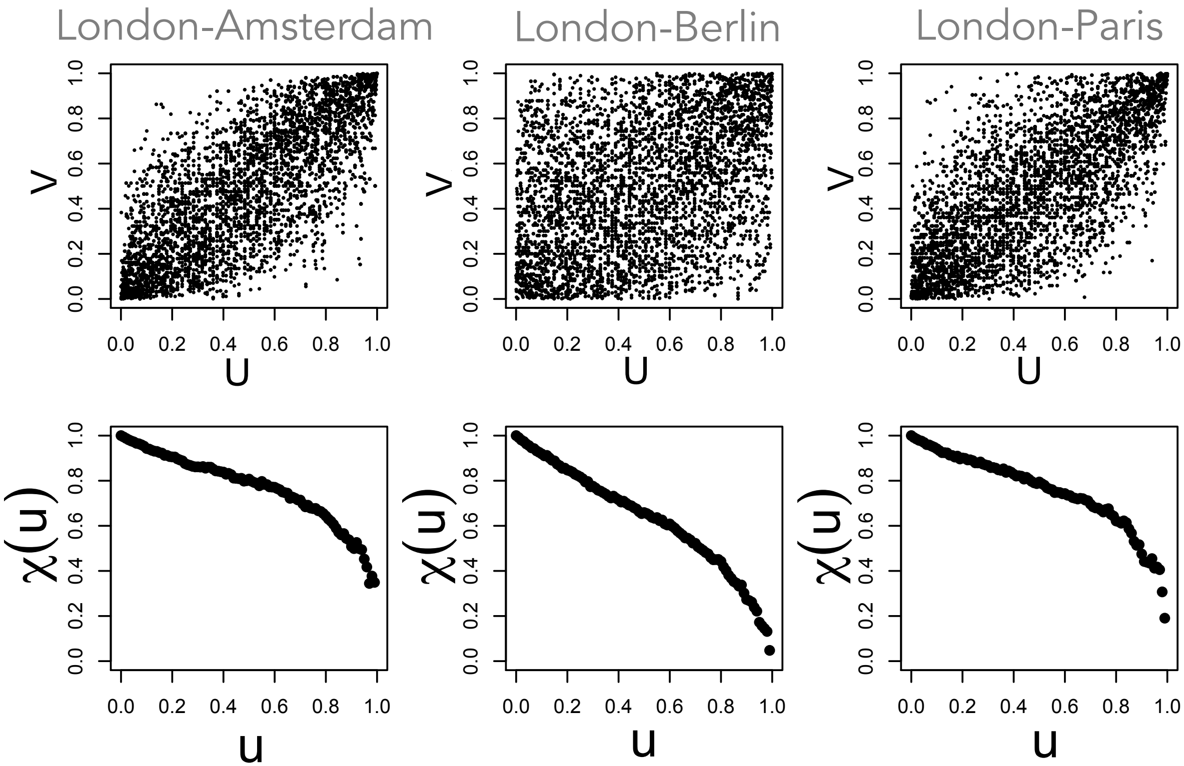
$$\boxed{\chi(u) = Pr(V > u | U > u)}$$



Which Copula?

- How does the empirical χ(u) behave for footprint wind-gusts at pairs of locations?

$$\boxed{\chi(u) = Pr(V > u | U > u)}$$



- The χ(u) curves resemble the Gaussian copula
- Less dependence between London-Berlin compared to Amsterdam and Paris
- Can't calculate the limit of χ(u), u→1 empirically due to the rarity of extreme events in the data set

Which Copula?

- Ferro (2007) proposes a probability model for the joint distribution of two variables for application in forecast verification of rare, extreme weather events (full reference on final slide)
- The model represents the diagonal of the copula as follows:

$$\boxed{\begin{aligned}\chi(u) &= Pr(V > u | U > u) \\ &= \kappa(1-u)^{\frac{1}{\eta}-1}\end{aligned}}$$

κ and η estimated from the data

- 0<η≤1 is known as the coefficient of tail dependence

Which Copula?

- Ferro (2007) proposes a probability model for the joint distribution of two variables for application in forecast verification of rare, extreme weather events (full reference on final slide)
- The model represents the diagonal of the copula as follows:

$$\boxed{\begin{aligned}\chi(u) &= Pr(V > u|U > u)\\ &= \kappa(1-u)^{\frac{1}{\eta}-1}\end{aligned}}$$

κ and η estimated from the data

- 0<η≤1 is known as the coefficient of tail dependence

η=1 $\lim_{u\to 1} \chi(u) > 0$ - Asymptotic dependence

η<1 $\lim_{u\to 1} \chi(u) = 0$ - Asymptotic independence

Which Copula?

- Ferro (2007) proposes a probability model for the joint distribution of two variables for application in forecast verification of rare, extreme weather events (full reference on final slide)
- The model represents the diagonal of the copula as follows:

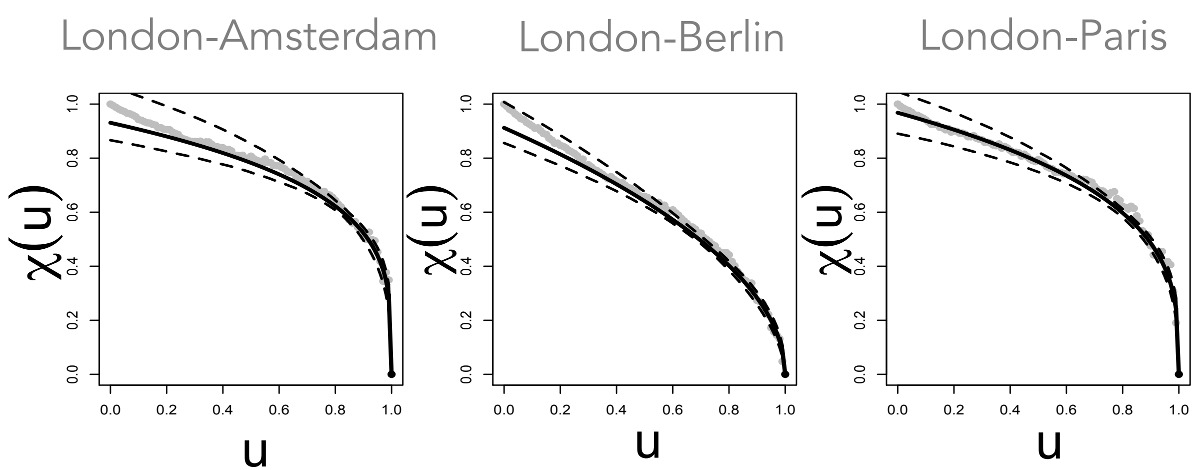
$$\boxed{\begin{aligned}\chi(u) &= Pr(V > u | U > u) \\ &= \kappa(1-u)^{\frac{1}{\eta}-1}\end{aligned}}$$

κ and η estimated from the data

- $0<\eta\leq 1$ is known as the coefficient of tail dependence

$\eta=1$ $\lim_{u\to 1} \chi(u) > 0$ - Asymptotic dependence

$\eta<1$ $\lim_{u\to 1} \chi(u) = 0$ - Asymptotic independence



Which Copula?

- Ferro (2007) proposes a probability model for the joint distribution of two variables for application in forecast verification of rare, extreme weather events (full reference on final slide)
- The model represents the diagonal of the copula as follows:

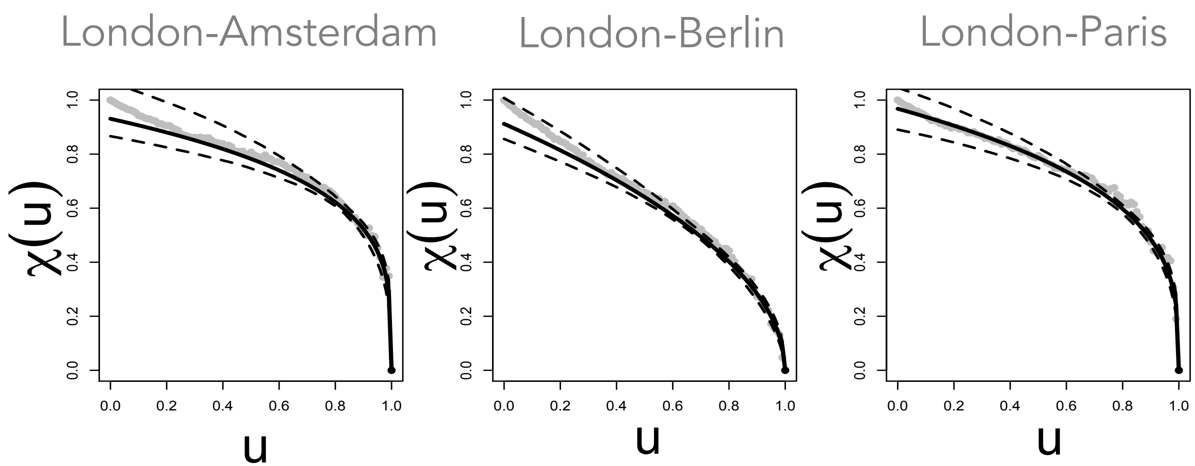
$$\boxed{\begin{aligned}\chi(u) &= Pr(V > u | U > u) \\ &= \kappa(1-u)^{\frac{1}{\eta}-1}\end{aligned}}$$

κ and η estimated from the data

- $0<\eta\leq 1$ is known as the coefficient of tail dependence

$\eta=1$ $\lim_{u\to 1} \chi(u) > 0$ - Asymptotic dependence

$\eta<1$ $\lim_{u\to 1} \chi(u) = 0$ - Asymptotic independence



- Estimate η for all pairs of locations: η < 1 (p value << 0.01)

Footprint wind-gusts are asymptotically independent – use the Gaussian Copula

Model Validation

My model

- Model margins using the Generalized Extreme Value (GEV) distribution
- Model the dependence using the Gaussian Copula

Bonazzi et al. (2012) (full reference on final slide)

- Model margins using the Generalized Pareto Distribution (GPD)
- Model the dependency using the Gumbel Copula
- Incorrectly assumes asymptotic dependence

Model Validation

My model

- Model margins using the Generalized Extreme Value (GEV) distribution
- Model the dependence using the Gaussian Copula

Bonazzi et al. (2012) (full reference on final slide)

- Model margins using the Generalized Pareto Distribution (GPD)
- Model the dependency using the Gumbel Copula
- Incorrectly assumes asymptotic dependence

Can the models realistically represent joint losses?

Conceptual loss function: $L(X_1, X_2) = H(X_1 > t) + H(X_2 > t)$

$H(a) = 1$ if a is true and 0 otherwise

Model Validation

Loss distribution for t=25ms^{-1}

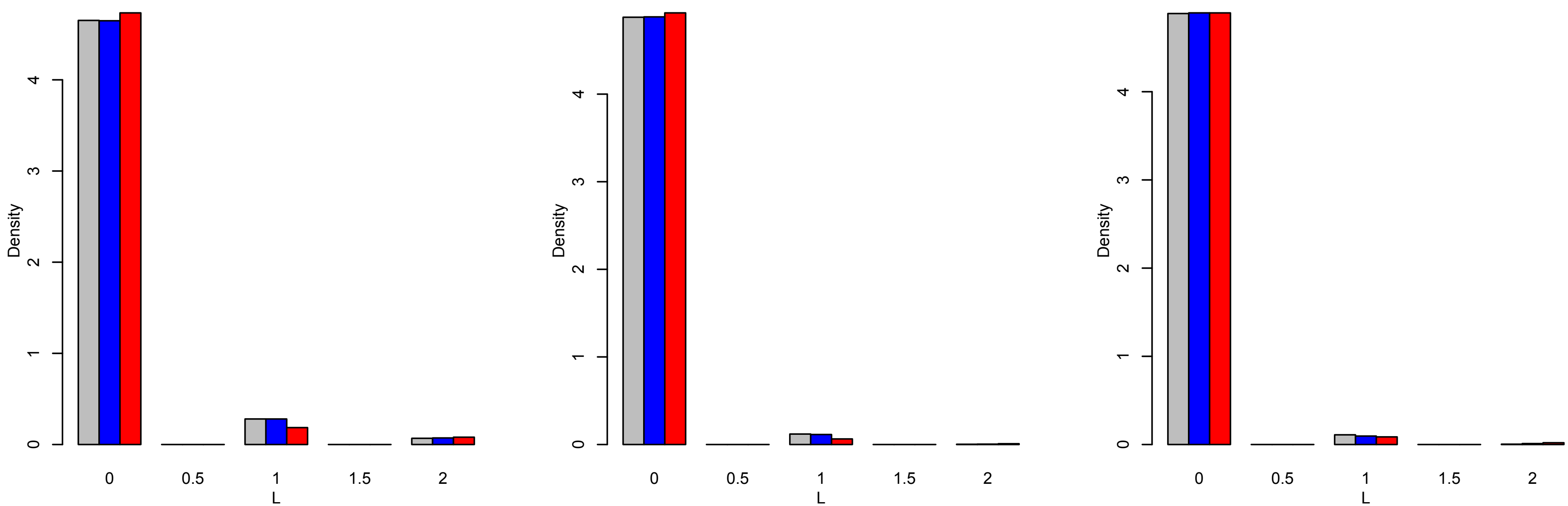
Shown to be a good approximation of insurance loss (Roberts et. al, 2014)

Can the models realistically represent joint losses?

$$L(X_1, X_2) = H(X_1 > t) + H(X_2 > t)$$

$H(a) = 1$ if a is true and 0 otherwise

London-Amsterdam London-Berlin London-Paris



Empirical My model Bonazzi

Model Validation

Loss distribution for t=25ms[-1]

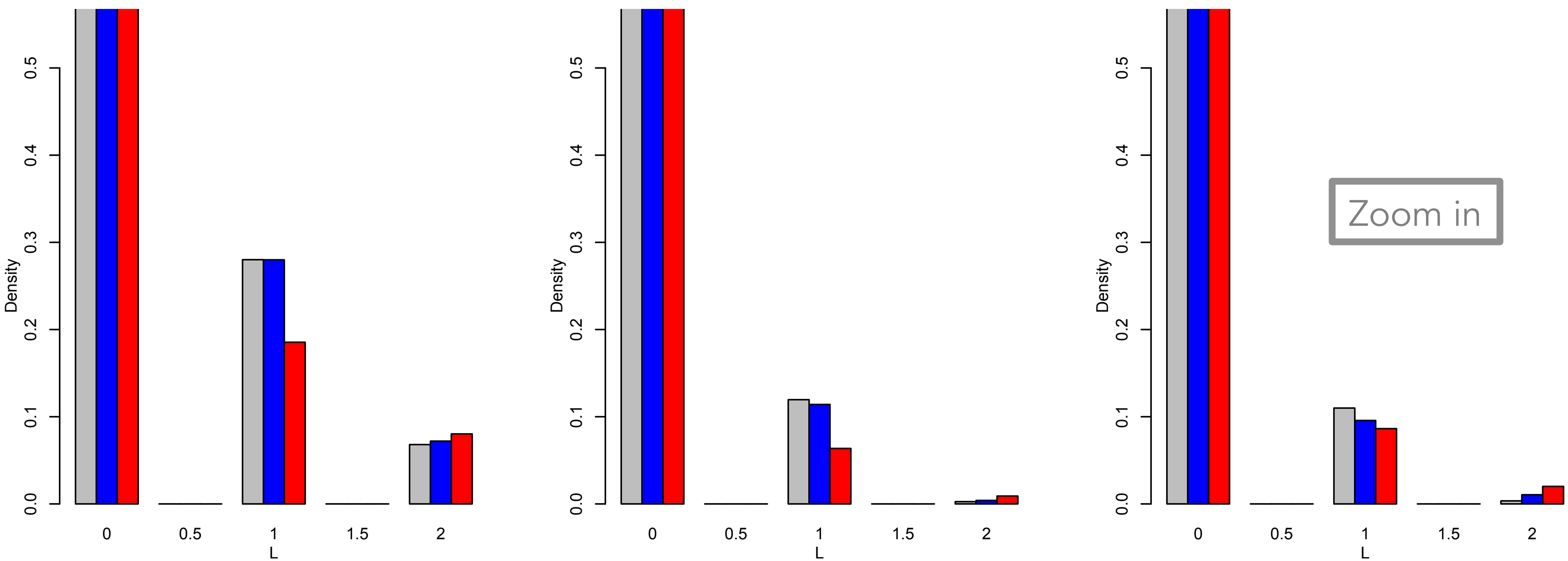
Shown to be a good approximation of insurance loss (Roberts et. al, 2014)

Can the models realistically represent joint losses?

$$L(X_1, X_2) = H(X_1 > t) + H(X_2 > t)$$

$H(a) = 1$ if a is true and 0 otherwise

London-Amsterdam London-Berlin London-Paris



Empirical My model Bonazzi

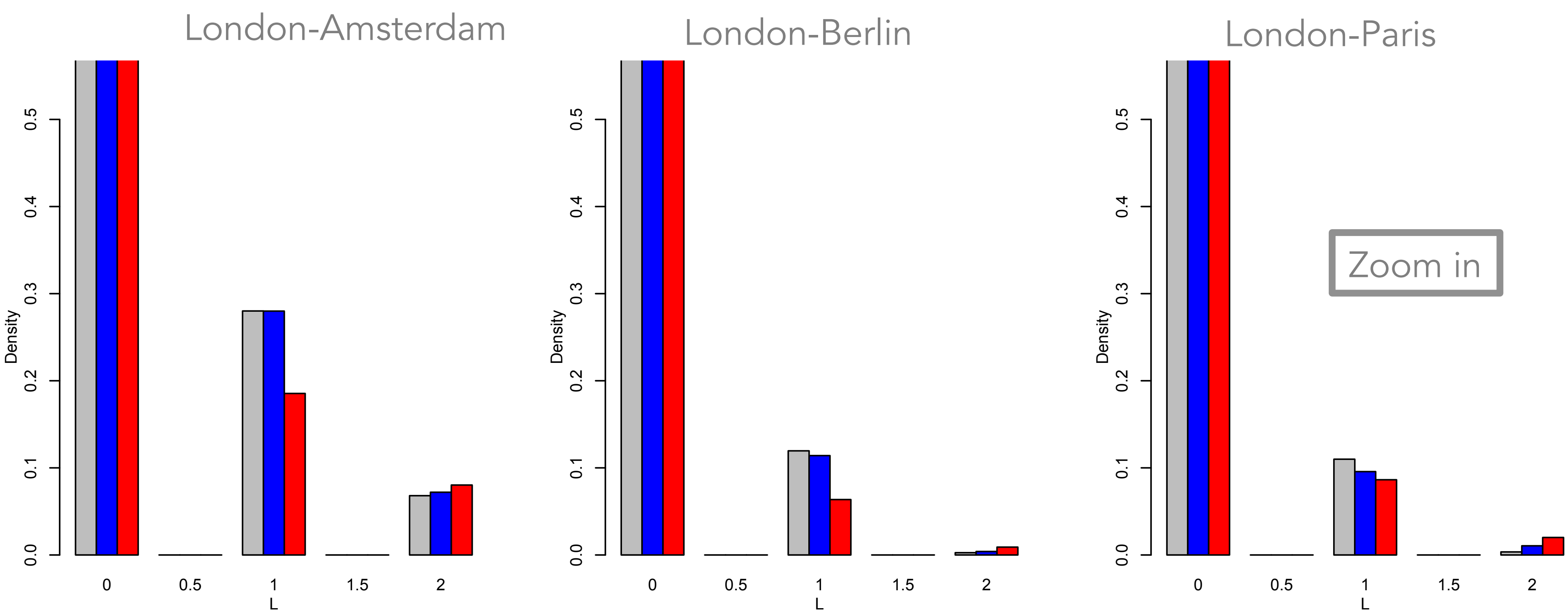
Model Validation

Loss distribution for t=25ms[-1]
Shown to be a good approximation of insurance loss (Roberts et. al, 2014)

Can the models realistically represent joint losses?

$$L(X_1, X_2) = H(X_1 > t) + H(X_2 > t)$$

$H(a) = 1$ if a is true and 0 otherwise



Empirical My model Bonazzi

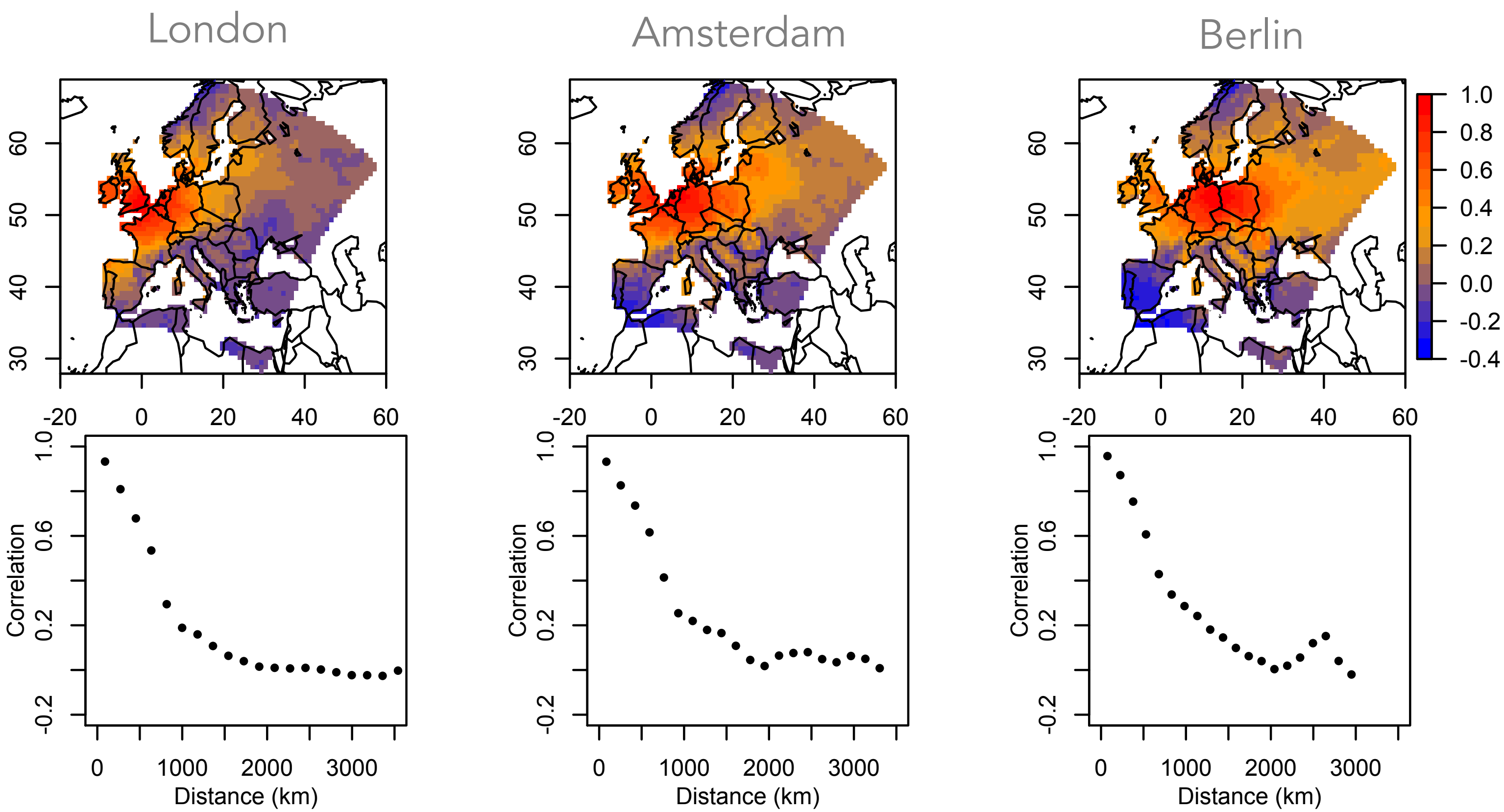
- Bonazzi's model, which assumes asymptotic dependence, over estimates the probability of joint extreme events
- My model does a better job of realistically representing joint losses

Spatial Model

- Using geostatistics to develop a spatial model for all locations at once
- The spatial correlation is modeled as a function of separation distance and direction
- Quicker to simulate from compared to a multivariate copula
- Fewer model parameters, simplifying the future sensitivity analysis experiment

Spatial Model

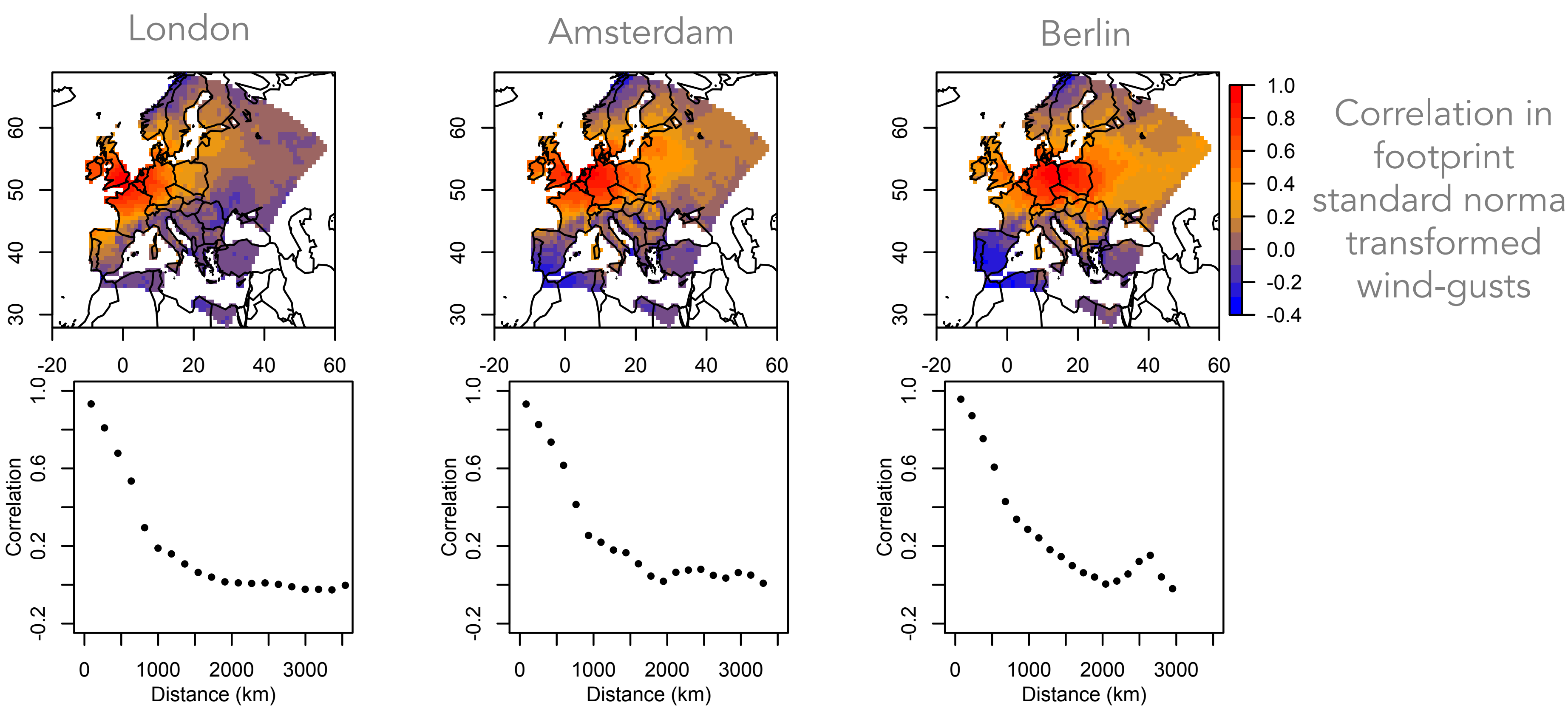
- Using geostatistics to develop a spatial model for all locations at once
- The spatial correlation is modeled as a function of separation distance and direction
- Quicker to simulate from compared to a multivariate copula
- Fewer model parameters, simplifying the future sensitivity analysis experiment



Correlation in footprint standard normal transformed wind-gusts

Spatial Model

- Using geostatistics to develop a spatial model for all locations at once
- The spatial correlation is modeled as a function of separation distance and direction
- Quicker to simulate from compared to a multivariate copula
- Fewer model parameters, simplifying the future sensitivity analysis experiment



- The correlation structure is similar for these locations suggesting that a geostatistical model may be able to capture the dependence structure of footprint wind-gusts.

Conclusions & Future Work

- Wind-gusts at pairs of locations are asymptotically independent

Conclusions & Future Work

- Wind-gusts at pairs of locations are asymptotically independent
- The Gaussian copula is able to realistically represent joint losses. The Gumbel copula over estimates the probability of joint extreme events because of the incorrect assumption of asymptotic dependence

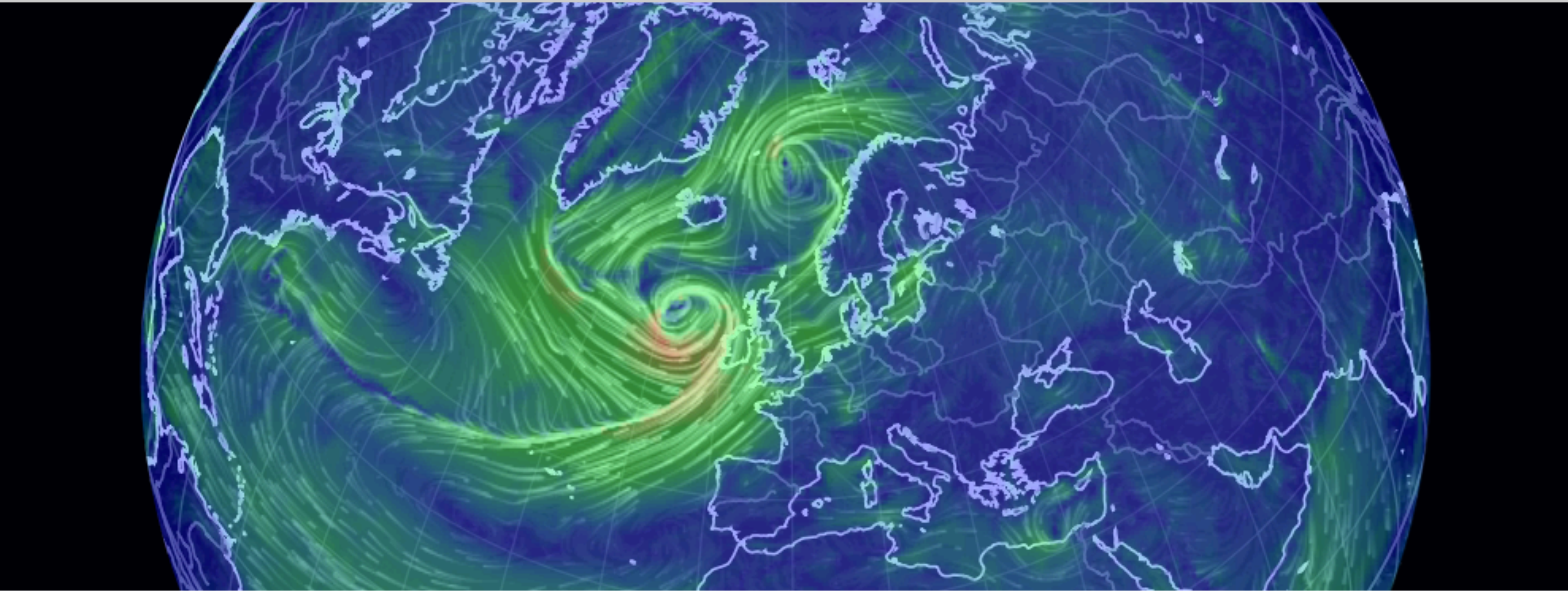
Conclusions & Future Work

- Wind-gusts at pairs of locations are asymptotically independent
- The Gaussian copula is able to realistically represent joint losses. The Gumbel copula over estimates the probability of joint extreme events because of the incorrect assumption of asymptotic dependence
- Now developing a geostatistical model for windstorm footprints which models spatial dependence as a function of separation distance and direction

Conclusions & Future Work

- Wind-gusts at pairs of locations are asymptotically independent
- The Gaussian copula is able to realistically represent joint losses. The Gumbel copula over estimates the probability of joint extreme events because of the incorrect assumption of asymptotic dependence
- Now developing a geostatistical model for windstorm footprints which models spatial dependence as a function of separation distance and direction
- Next: Finish fitting and validating the geostatistical model and carry out sensitivity analysis experiment

Thank you for listening
Any questions?



References

Bonazzi, A., Cusack, S., Mitas, C., and Jewson, S. (2012). The spatial structure of European wind storms as characterized by bivariate extreme-value copulas. Nat. Hazards Earth Syst. Sci., 12:1769–1782.

Ferro, C. A. T. (2007). A Probability Model for Verifying Deterministic Forecasts of Extreme Events. Weather Forecasting, 22:1089–1100.

Hodges, K. I. (1995). Feature tracking on the unit sphere. Monthly Weather Review, 123:3458–3465.

Roberts, J. F., Champion, A., Dawkins, L., Hodges, K. I., Shaffrey, L., Stephenson, D. B., Stringer, M., Thornton, H. and Youngman, B., (2014). The XWS open access catalogue of extreme windstorms in Europe from 1979–2012. Natural Hazards and Earth System Science.

UNIVERSITY OF EXETER

lcd206@exeter.ac.uk

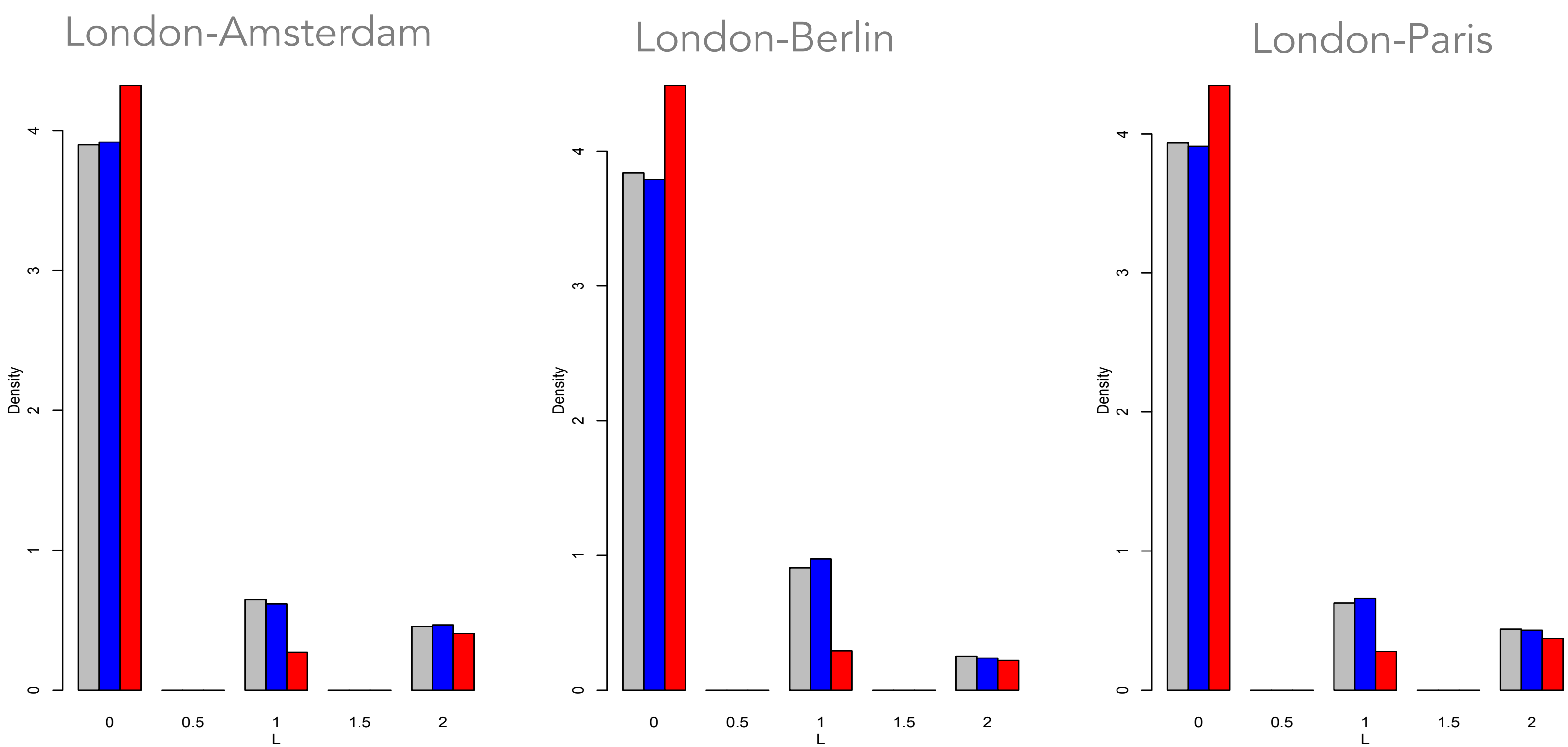
Model Validation

Loss distribution for t= local 98th percentile of climatology

Can the models realistically represent joint losses?

$$L(X_1, X_2) = H(X_1 > t) + H(X_2 > t)$$

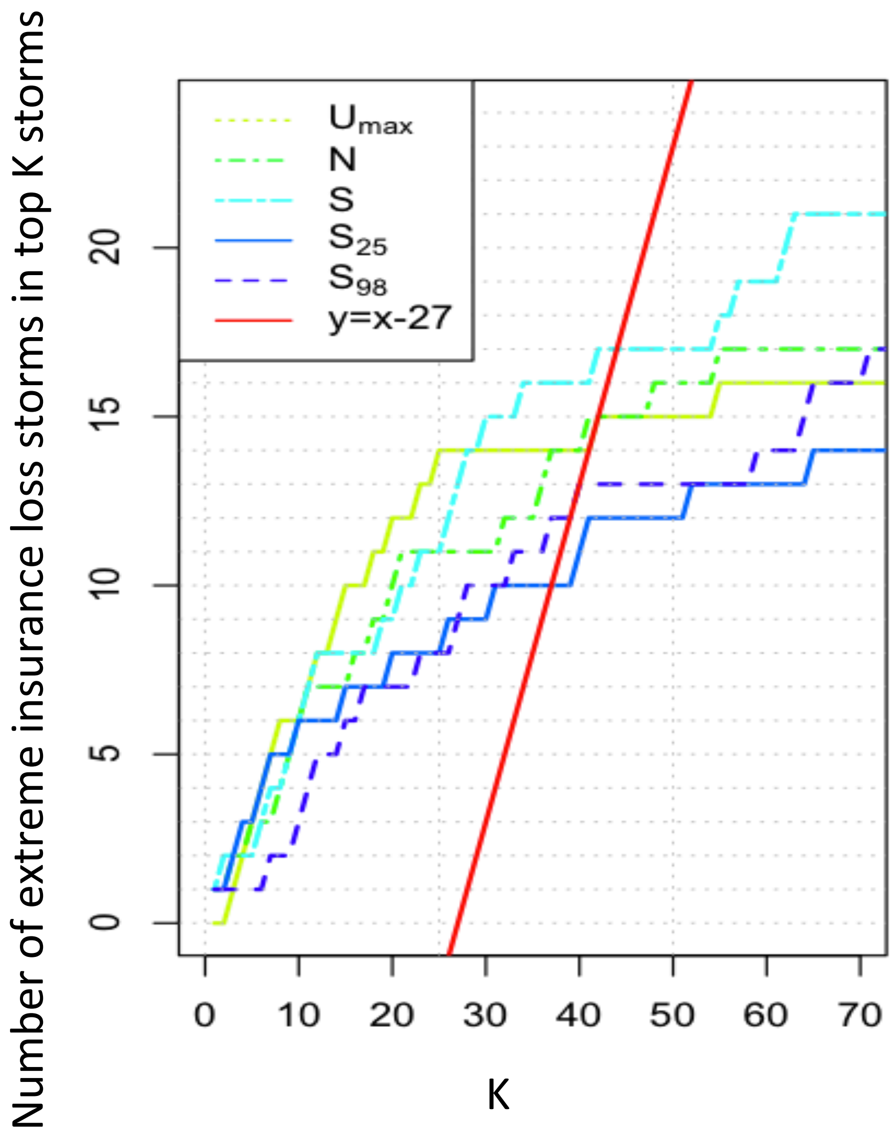
$H(a) = 1$ if a is true and 0 otherwise



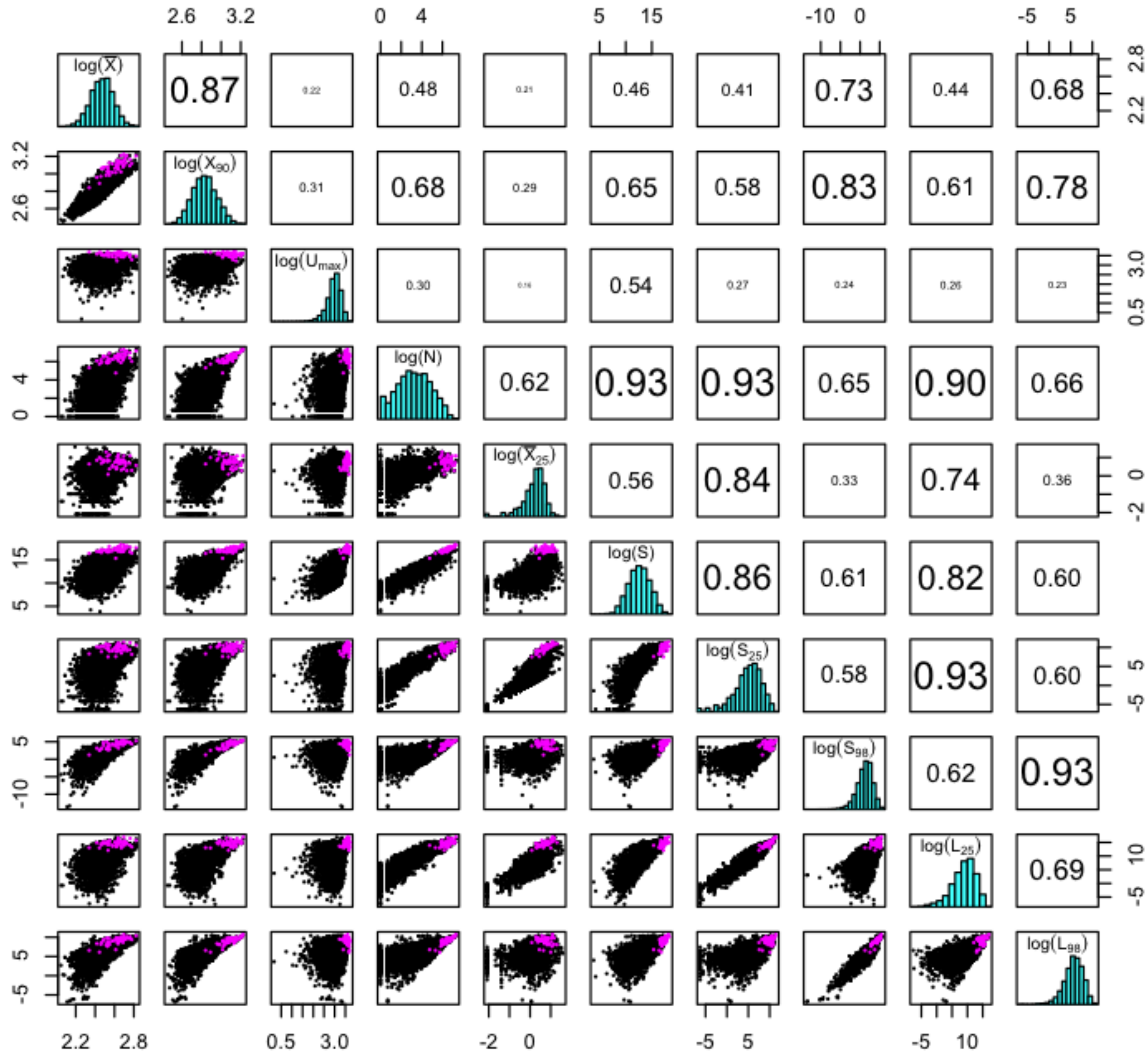
Empirical My model Bonazzi

SSIs

Name	Date of maximum 925hPa windspeed over land
Daria	25 Jan 1990
Lothar	26 Dec 1999
Kyrill	18 Jan 2007
Great Storm of '87	16 Oct 1987
Vivian	26 Feb 1990
Klaus	24 Jan 2009
Martin	27 Dec 1999
Xynthia	27 Feb 2010
Anatol	3 Dec 1999
Erwin	8 Jan 2005
Herta	3 Feb 1990
Emma	29 Feb 2008
Wiebke	28 Feb 1990
Gero	11 Jan 2005
Ulli	3 Jan 2012
Dagmar-Patrick	26 Dec 2011
Fanny	4 Jan 1998
Jeanette	27 Oct 2002
Lore	28 Jan 1994
Oratio	30 Oct 2000
Stephen	26 Dec 1998
Xylia	28 Oct 1998
Yuma	24 Dec 1997



SSIs



log($\overline{X}$)
log(X_{90})
log(U_{max})
log(N)
log($\overline{X}_{25}$)
log(S)
log(S_{25})
log(S_{98})
log(L_{25})
log(L_{98})
0.87
0.48
0.46
0.41
0.73
0.44
0.68
0.31
0.68
0.29
0.65
0.58
0.83
0.61
0.78
0.30
0.54
0.27
0.24
0.26
0.23
0.62
0.93
0.93
0.65
0.90
0.66
0.56
0.84
0.33
0.74
0.36
0.86
0.61
0.82
0.60
0.58
0.93
0.60
0.62
0.93
0.69

Spatial Model

- One such correlation function is the Matérn:

$$\rho(d) = \frac{1}{2^{\nu-1}\Gamma(\nu)} \left(\frac{2\sqrt{\nu t}}{\phi} \right)^{\nu} K_{\nu} \left(\frac{2\sqrt{\nu t}}{\phi} \right)$$

- Γ - gamma function
- K – modified Bessel function
- ϕ – spatial scale parameter
- ν - shape parameter – added flexibility of the Matérn model

- Estimate scale parameter, ϕ, for a fixed shape, ν
- Plot over empirical binned correlogram
- Best fit for $\nu = 0.5$
- Windstorm footprints are a rough spatial process with correlation dropping off quickly for small separation distance