Spatial Semiparametric Spectral Density Estimation for Multivariate Extremes

Mauricio Nascimento
Pennsylvania State University
Joint work with Ben Shaby

November 14, 2016

TABLE OF CONTENTS

MULTIVARIATE EXTREMES

- Wave heights and high still water level
- Concentrations of different air pollutants
- Wind speed and fire spread conditions

California Department of Forestry and Fire Protection(Cal Fire)

Figure: California Department of Forestry and Fire Protection[1]

[1] http://www.fire.ca.gov/lmu/header/2012_LMU_header_04.jpg

FIRE BURNS MORE THAN 172.000 ACRES ANNUALLY

Figure: Rocky Fire burns near Clearlake August, 2 2015[2]

[2] http://www.cnn.com/2015/07/18/us/california-freeway-fire/

DATA

- Wind speed
- Fosberg Fire Weather Index (FFWI)

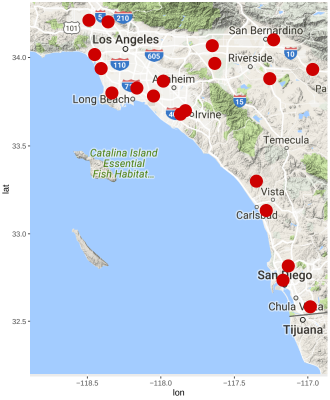


Figure: Location of 20 meteorological stations on California

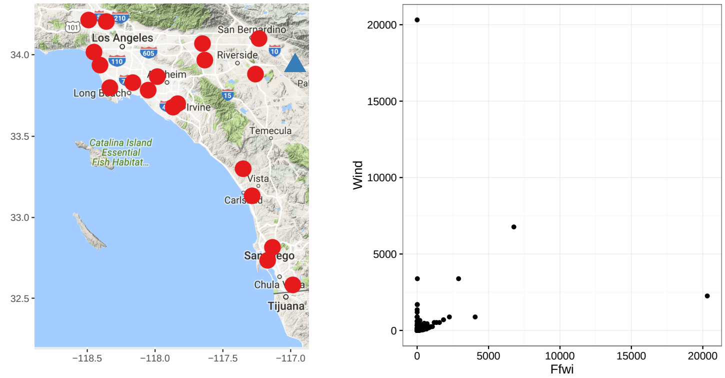


Figure: Scatterplot of Wind and FFWI on the Fréchet scale for Beaumont, CA

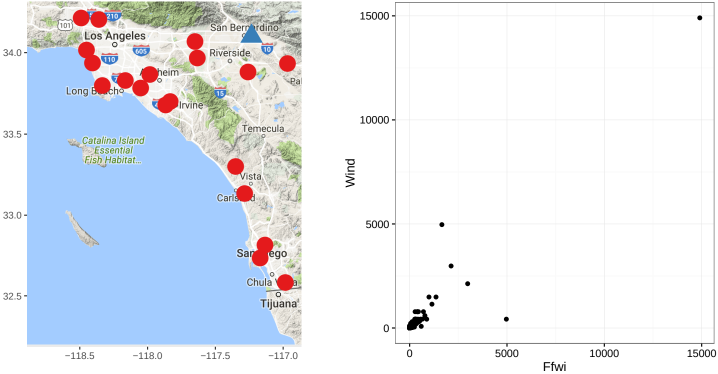
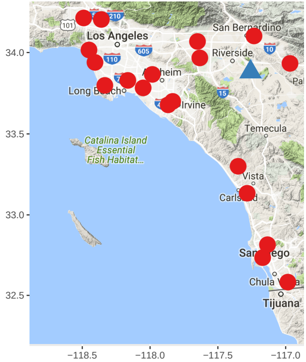


Figure: Scatterplot of Wind and FFWI on the Fréchet scale for San Bernardino, CA



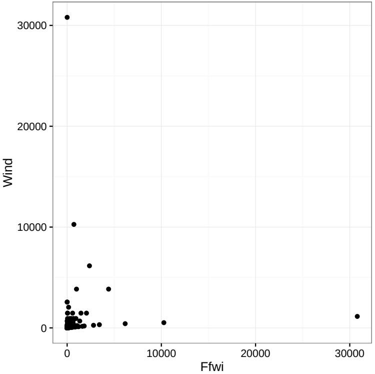


Figure: Scatterplot of Wind and FFWI on the Fréchet scale for Perris, CA

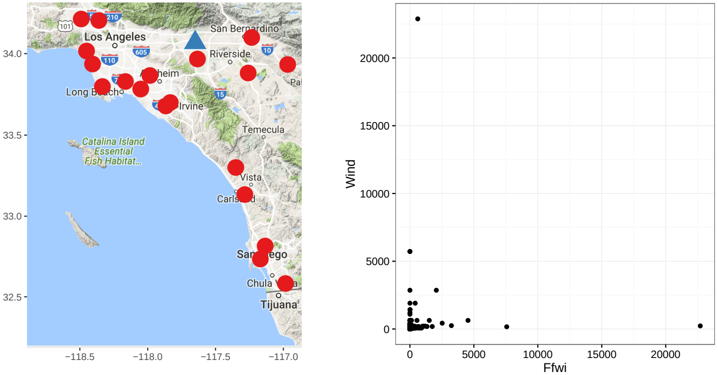


Figure: Scatterplot of Wind and FFWI on the Fréchet scale for Ontario, CA

MODEL FOR EXTREME DEPENDENCE

Assuming the individual margins are unit Frechet the distribution can be write it in polar coordinates $(R = \sum(\boldsymbol{x}), \boldsymbol{w} = \frac{\boldsymbol{x}}{R})$ as

$$P(R > r, \boldsymbol{w} \in A) \sim r^{-1} H(A)$$

where $H(A)$ is the angular measure in the unit simplex $\mathbb{S}^d$ with the following restriction

$$\int_{\mathbb{S}^d} w_i dH(\boldsymbol{w}) = \frac{1}{d} \; \forall i = 1, \ldots, d$$

NON-PARAMETRIC MODEL FOR DEPENDENCE USING A MIXTURE OF DIRICHLET[3]

$$h(\boldsymbol{w}) = \sum_{k=1}^{K} p_k Dirichlet(\boldsymbol{\mu_k}, \nu_k)$$

where

- p_k is the weight of the k_{th} component
- $\boldsymbol{\mu_k}$ is the mean vector of the k_{th} component
- ν_k is the concentration parameter of the k_{th} component

and the previous constraint becomes.

$$\int_{\mathbb{S}^d} w_i dH(\boldsymbol{w}) = \sum_{k=1}^{K} p_k \mu_{ik} = \frac{1}{d} \; \forall i = 1, ..., d$$

[3] Boldi, M.-O. and Davison, A. C. (2007), A mixture model for multivariate extremes. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 69: 217229.

RE-PARAMETRIZATION OF MODEL[4]

Original parameters

- $p_k \in [0,1], k = 1, ..., K$
- $\nu_k \in \mathbb{R}^+, k = 1, ..., K$
- $\mu_k \in \mathbb{S}^d, k = 1, ..., K$
- where $\sum_{k=1}^K p_k = 1$ and $\sum_{k=1}^K p_k \mu_{ik} = \frac{1}{d}$

New parameters

- $\epsilon_k \in [0,1], k = 1, ..., K-1$
- $\nu_k \in \mathbb{R}^+, k = 1, ..., K$
- $\boldsymbol{\mu}_k \in \mathbb{S}^d, k = 1, ..., K$
- where $(\boldsymbol{\epsilon}, \boldsymbol{\mu}_K) = T(\boldsymbol{\mu}, \boldsymbol{p})$

[4] Sabourin A, Naveau P. Bayesian Dirichlet mixture model for multivariate extremes: A re-parametrization. Computational Statistics and Data Analysis. 2014;71:542.

SPATIAL DEPENDENCE

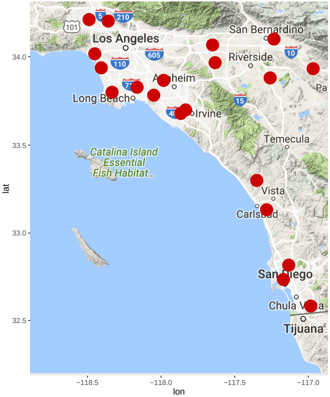


Figure: Location of 20 meteorological stations in California

Spatial dependence

$$h(\boldsymbol{w}_{ij}) = \sum_{k=1}^{K} p_{ki} Dirichlet(\boldsymbol{\mu_{ki}}, \nu_{ki}) \text{ for } i \in 1:I, j \in 1:J_i$$

$$(\boldsymbol{p_i}, \boldsymbol{\mu_{Ki}}) = T^{-1}(\boldsymbol{\epsilon_i}, \boldsymbol{\mu_i})$$

$$log(\nu_k) \sim GP(\boldsymbol{X\beta}_{\nu_k}, \Sigma_{\nu_k}) \text{ for } k \in 1, ..., K$$

$$logit(\epsilon_k) \sim GP(\boldsymbol{X\beta}_{\epsilon_k}, \Sigma_{\epsilon_k}) \text{ for } k \in 1, ..., K-1$$

$$\boldsymbol{\mu_k} \sim Dirichlet(\boldsymbol{\mu}_0, \nu_0) \text{ for } k \in 1, ..., K-1$$

$log(\nu)$

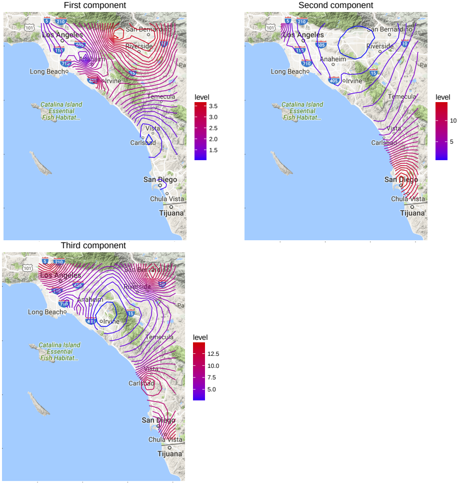


Figure: Spatial distribution of $log(\nu)$

$logit(\epsilon)$

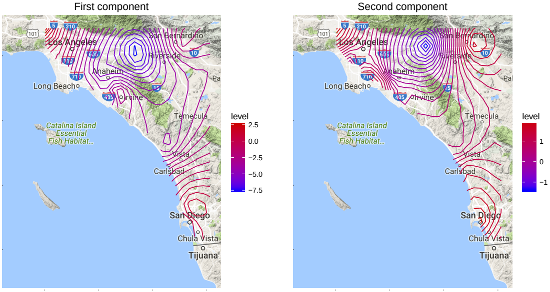


Figure: Spatial distribution of $logit(\epsilon)$

p

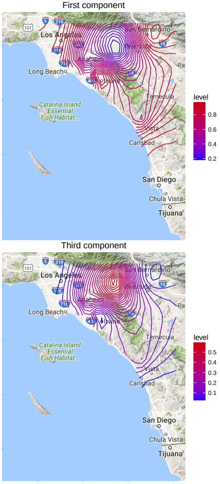


Figure: Spatial distribution of p

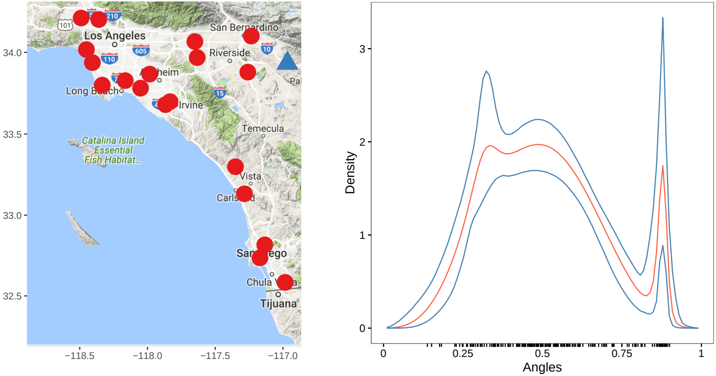


Figure: Posterior distribution of the angles for Beaumont, CA

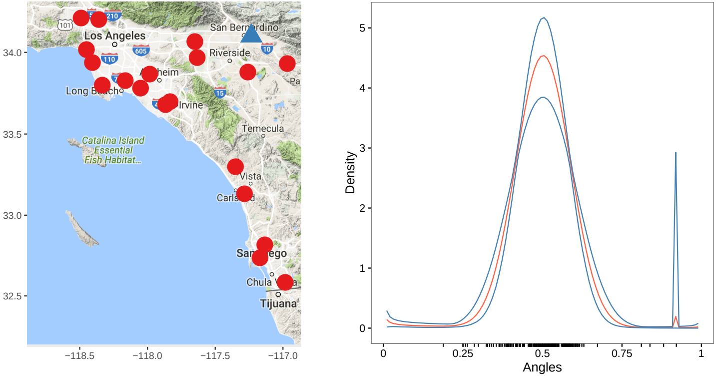


Figure: Posterior distribution of the angles for San Bernardino, CA

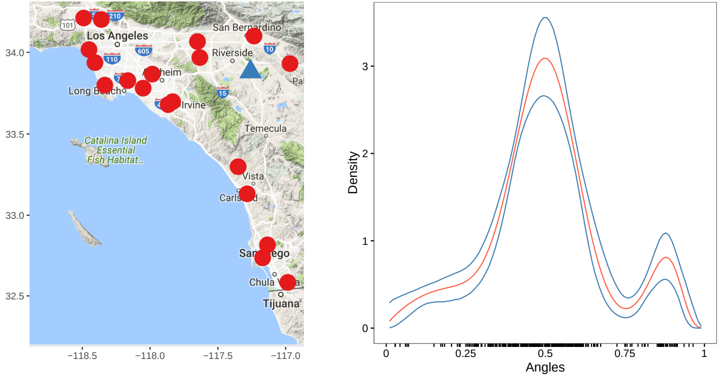


Figure: Posterior distribution of the angles for Perris, CA

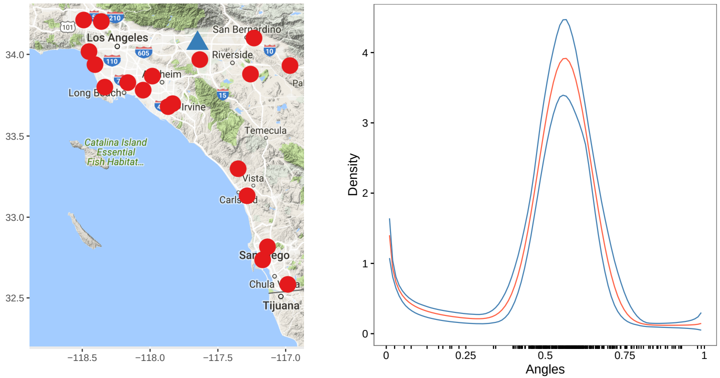


Figure: Posterior distribution of the angles for Ontario, CA

Current and Future work

- Estimating marginal distribution along side the dependence
  - Complicated because of censoring
- Estimating the number of components
  - Implemented reversible jump for non-spatial case

Thank You!