A Latent Mixture Model for Extreme Winds

Gregory Bopp

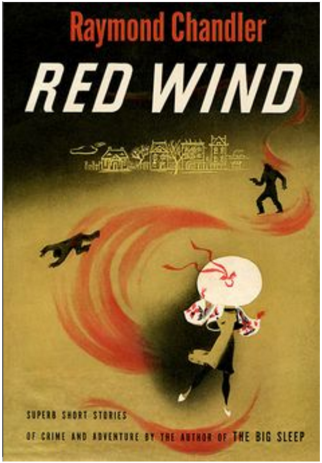
Pennsylvania State University

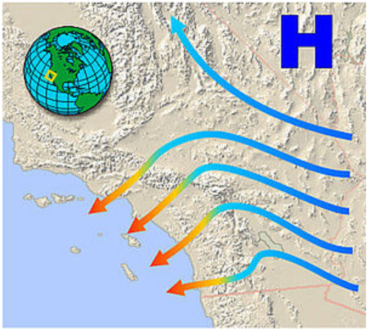
gxb951@psu.edu

October 23, 2016

Santa Ana Winds

"It was one of those hot dry Santa Anas that come down through the mountain passes and curl your hair and make your nerves jump and your skin itch." - *Raymond Chandler*





Santa Ana Winds

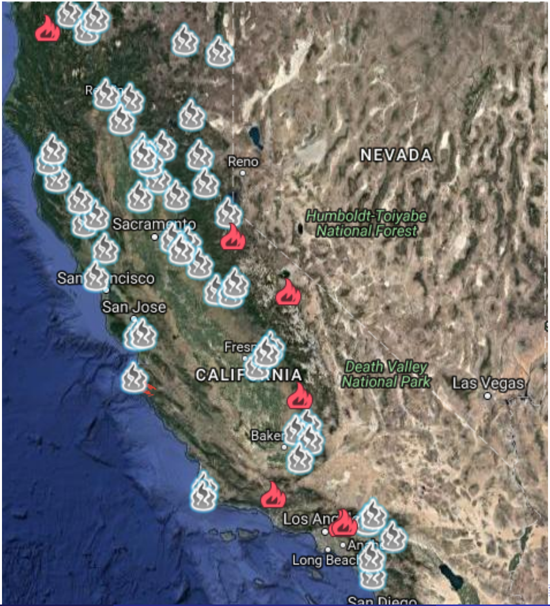
Fire	Year	Acres Burned	Cost
Cedar Fire	2003	273,246	$27 M
Esperanza Fire	2006	40,200	$9.9 M
Witch Fire	2007	197,990	$18 M

Data obtained from Cal Fire

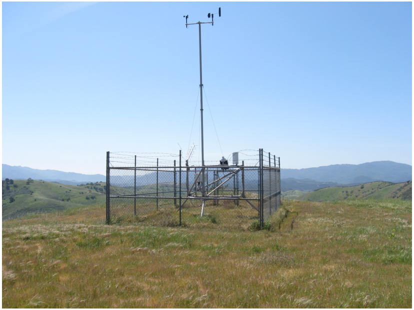
- SA fires often consume about half of the final burn area within the first day, and last about half as long as their non-SA counterparts in California[1]
- Santa Ana fires were responsible for $3.1 B in losses between 1990 and 2009[1].
- Santa Ana fires tend to encroach on coastal areas and pose greater risk to human lives and property loss than summer fires.

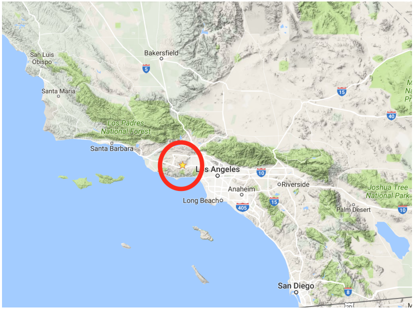
[1] Jin et al., (2015)

2016 Fires in California



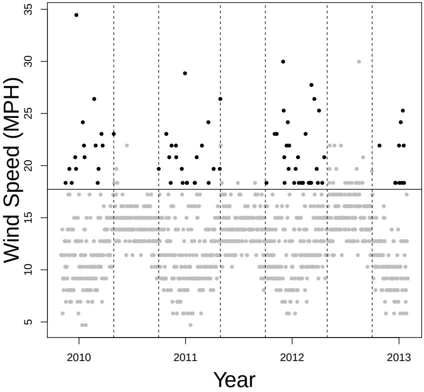
Cheeseboro, CA Weather Station





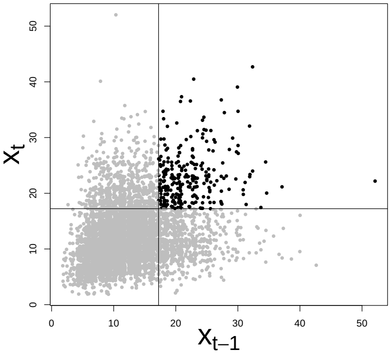
Wind Speed Time Series

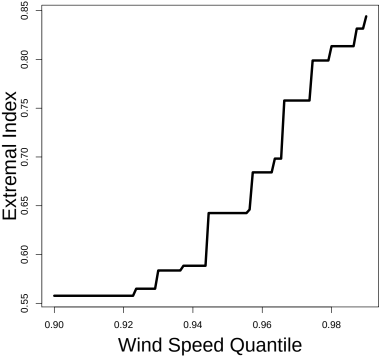
- Observation period: 1974 - 2014 (October - April)
- Weather station location: Cheeseboro, CA
- Data resolution: daily maxima



Wind EDA

Extremal Index: (limiting mean cluster size)$^{-1}$





How to Account for Temporal Dependence in Extremes?

The distinction between asymptotic dependence and independence has received considerable attention in multivariate extreme value statistics.

Desired model features

- Marginally extremes should follow and extremal distribution (e.g. GPD, GEV)
- Allow for either asymptotic dependence or independence
- Simplicity

GPD as a Gamma Mixture of Exponentials

Consider the following reparameterization of a generalized Pareto distribution (GPD) with shape ξ and scale σ. Let $\alpha = 1/\xi$ and $\beta = \sigma/\xi$

$$F(y; \alpha, \beta) = 1 - \left(1 + \frac{y}{\beta}\right)_{+}^{-\alpha}, \qquad y \geq 0$$

For $\alpha > 0$ a GPD(α, β) can be represented as a Gamma mixture of exponential distributions (Reiss and Thomas, 2007).

$$f(y; \alpha, \beta) = \int_0^\infty h(y; \lambda) g(\lambda; \alpha, \beta) d\lambda$$

for

$$h(y; \lambda) = \lambda e^{-y\lambda} \qquad \text{and} \qquad g(\lambda; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{\beta\lambda}$$

Motivation for Hierarchical Model

GPD marginal density for excesses:

$$f(y;\alpha,\beta) = \int_0^\infty [\lambda e^{-y\lambda}][\frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{\beta\lambda}]d\lambda$$

Hierarchical Model

If

$$Y|\Lambda = \lambda \sim \text{Exp}(\lambda)$$
$$\Lambda|\alpha,\beta \sim \text{Gamma}(\alpha,\beta)$$

then for $\xi = 1/\alpha$ and $\sigma = \beta/\alpha$

$$Y \sim GPD(\sigma,\xi)$$

Bayesian Hierarchical Model

- **Latent Process** $\{\Lambda_t, t \geq 1\}$
  - First order Markov process Gamma margins
  - Different choices will yield different dependence characteristics in the response
- **Conditional Observation Model** $[Y_t | \Lambda_t]$
  - Conditionally independent, exponentially distributed excesses conditional on latent stage

Inference Procedure

- Variable-at-a-time Metropolis Hastings algorithm to integrate out latent process, $\{\Lambda_t, t \geq 1\}$
- Can induce either asymptotic dependence or independence (Bortot and Gaetan, 2014)

Observation Model

Notation

- $\{X_t, t \geq 1\}$: stationary time series
- $\{Y_t, t \geq 1\}$: excesses series
- u: high threshold

$$Y_t = (X_t - u)1_{\{X_t > u\}}$$

Conditional excess model

$$P(X_t > u | \Lambda_t) = \exp(-\kappa \Lambda_t) \qquad \text{(probability of exceedance)}$$

$$Y_t | (\Lambda_t, X_t > u) \sim \mathrm{Exp}(\Lambda_t) \qquad \text{(conditional excess distribution)}$$

Latent Stage: Warren (1992) Process

Markov process with Gamma margins

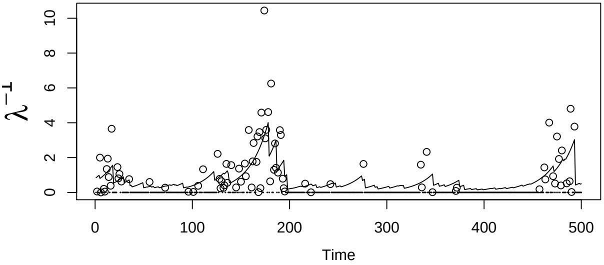
$$\Lambda_0 \sim \text{Gamma}(\alpha, \beta)$$
$$\Pi_t | \Lambda_{t-1} \sim \text{Poisson}\left(\Lambda_{t-1} \frac{\beta\rho}{1-\rho}\right)$$
$$\Lambda_t | \Pi_t \sim \text{Gamma}(\Pi_t + \alpha, \frac{\beta}{1-\rho})$$

for $\alpha, \beta > 0$, $0 \leq \rho < 1$

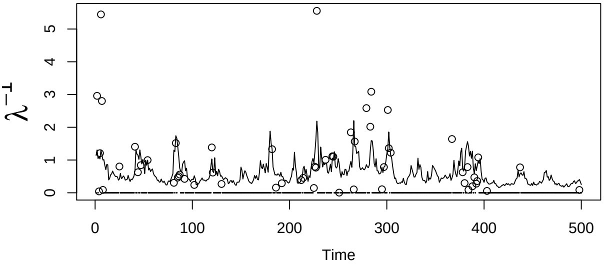
- Λ_t has Gamma(α, β) margins
- Results in asymptotic independence among exceedances
- Dependence modeled through ρ parameter

$$\text{corr}(\Lambda_t, \Lambda_{t+k}) = \rho^{|k|}$$

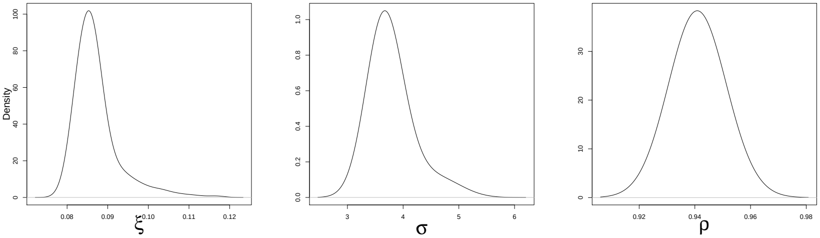
Gaver−Lewis



Warren



Posterior Summaries



Priors

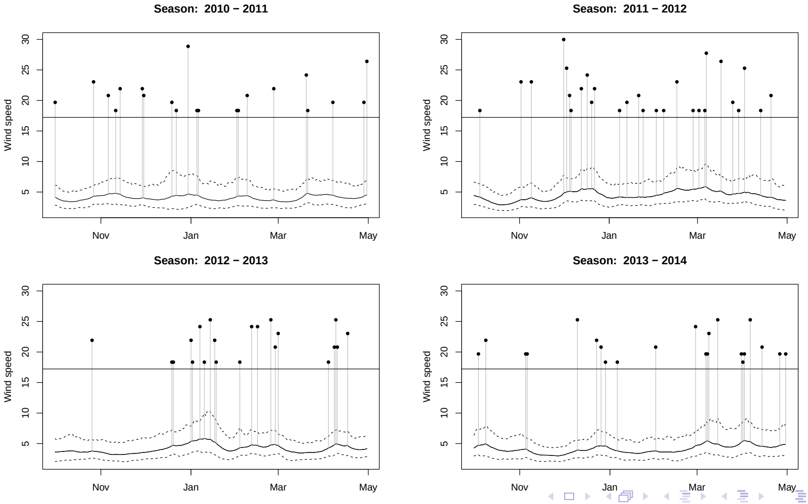
- $\alpha \sim$ Trunc. Normal($\mu = 0$, $\sigma = 1/2$, lb$= 0$, ub $= \infty$)
- $\beta \sim$ Trunc. Cauchy(lb $= 0$, ub $= \infty$)
- $\kappa \sim$ Trunc. Cauchy(lb $= 0$, ub $= \infty$)
- $\rho \sim$ Unif(0,1)

Posterior Means

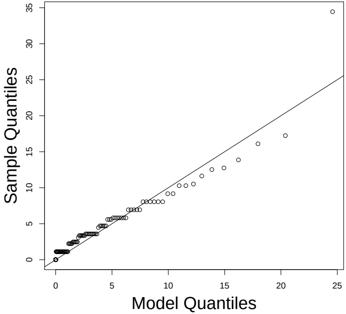
Parameter	Est. (95% Cred. Int.)
$\hat{\xi}$	0.087 (0.081, 0.104)
$\hat{\sigma}$	3.792 (3.455, 4.829)
$\hat{\rho}$	0.941 (0.927, 0.959)
$\hat{\kappa}$	9.695 (8.844, 12.415)

Estimated Latent Process

- Lines: posterior mean and 95% credible intervals for $1/\lambda_t$
- Points: y_t (exceedances)



Model Fit



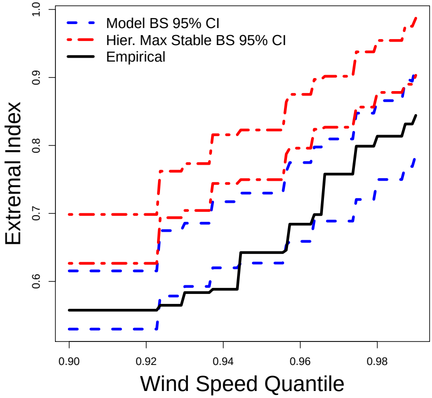
Model Fit

Model	DIC	LPML
Gamma Mixture	2289.7	-3109.4
Max-Stable[1]	2378.6	-3205.2

[1] Reich et al. (2014)

Extremal Dependence

Extremal Index: : $\theta = (\text{limiting mean cluster size})^{-1}$



Summary

Bayesian model for extremal dependence that allows for asymptotic dependence or independence while preserving *GPD* margins

Comments

- The $\xi > 0$ constraint can be relaxed with marginal transformations

Drawbacks

- Dependence is only controlled through a single parameter, ρ
- Cannot smoothly transition between independence and dependence
- Lack of separation between latent process parameters and marginal excess distribution parameters

Future Work

- Other Gamma processes with more flexible dependence structures
- Extend to spatial setting

Thank you!

Questions?

References

- Jin Y, Goulden M, Faivre N, Veraverbeke S, Sun F, Hall A, et al. (2015). Identification of two distinct fire regimes in Southern California: implications for economic impact and future change. Environmental Research Letters. 10(9).
- Paola Bortot and Carlo Gaetan. (2013). A latent process model for temporal extremes. *Scandinavian Journal of Statistics*. 41(3):606621.
- Reich BJ, Shaby BA, Cooley D. (2014). A Hierarchical Model for Serially-Dependent Extremes: A Study of Heat Waves in the Western US. Journal of Agricultural, Biological, and Environmental Statistics. 19(1):119-35.
- Reiss, R. and Thomas, M. (2007). *Statistical Analysis of Extreme Values: with Applications to Insurance, Finance, Hydrology and Other Fields*. 3rd; Birkhauser;