Statistical simulation of extreme European windstorms and other natural hazard events

Ben Youngman

University of Exeter

Storms workshop

Bern

2 September 2015

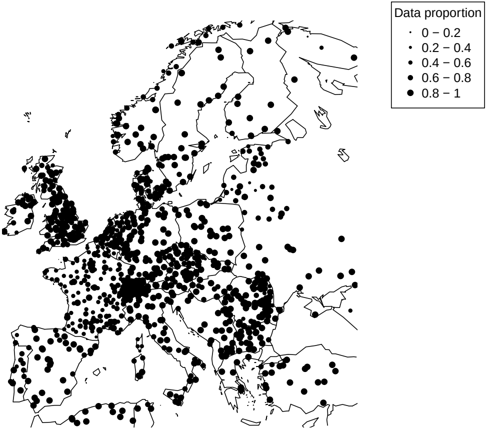
Stochastic European windstorm event set

Overview

- Develop statistical model for simulating windstorm footprints
- Model based primarily on wind gust measurements
- Quick simulations of high-resolution footprints
- Statistically sound basis
- Reliable model checks
- Sanity check for catastrophe models

The data

- National climatic data centre (NCDC) global summary of the day (GSOD) measurements
- Daily maximum wind gusts: DJF 1990–2014



The marginal model

- $Y(s,t)$ is wind gust measurement for location s at time t
- Extreme threshold $u = 22\text{m/s}$
- Assume generalised Pareto model

$$Y(s,t) - u \,|\, Y(s,t) > u \sim GPD(\psi(s,t),\, \xi(s,t))$$

with $pr(Y(s,t) \leq y \,|\, Y(s,t) > u) =$

$$F_u(y;\, \psi(s), \xi(s)) = 1 - \left(1 + \xi(s,t)\frac{y-u}{\psi(s,t)}\right)^{-1/\xi(s,t)}$$

The marginal model

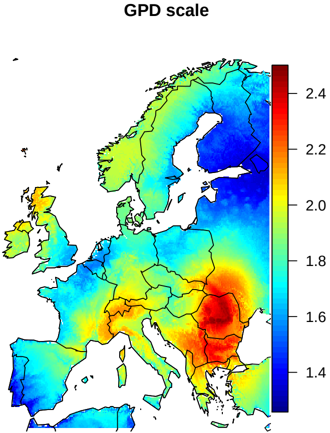
- We're simulating footprints for a region R
- Must know $\psi(s,t)$ and $\xi(s,t)$ for all $s \in R$
- Temporally stationary, at the moment, so $\psi(s,t) = \psi(s)$ and $\xi(s,t) = \xi(s)$
- Spline forms used to capture spatial variation

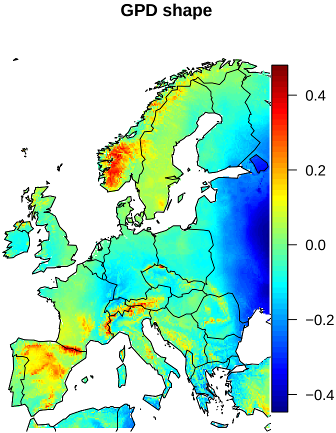
  $$\xi(s) = g_1(\text{elevation}) + g_2(\text{mean wind speed}) + g_3(\text{longitude, latitude})$$

  where
  - $g_1(\ )$ and $g_2(\ )$ are cubic regression splines
  - $g_3(\ ,\ )$ is a thin-plate spline
- Similarly for $\log(\psi(s))$

Marginal model

- Plots of estimates of $\psi(s)$ and $\xi(s)$





Spatial model

- Normal marginal distributions can be achieved:
- First

$$F_u(Y(s,t);\psi(s),\xi(s)) \sim Uniform([0,1])$$

- Then

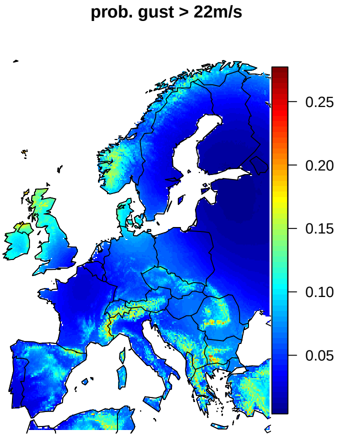
$$Z(s,t) := \Phi^{-1}(F_u(Y(s,t);\psi(s),\xi(s))) \sim N(0,1)$$

where $\Phi^{-1}(\,)$ is the inverted standard normal distribution

- We can then turn to geostatistical models for $Z(s,t)$
- These let dependence exist between $Z(s,t)$ and $Z(s',t)$ for different locations s and s'
- More details in Model-based Geostatistics (2007) by Diggle & Ribeiro

Exceedance rate model

- Finally need to estimate exceedance rate $pr(Y(s,t) > u)$
- Logistic regression is used with various covariates with flexible GAM forms, with logit link function and mean function $\mu(s)$



Footprint simulation algorithm

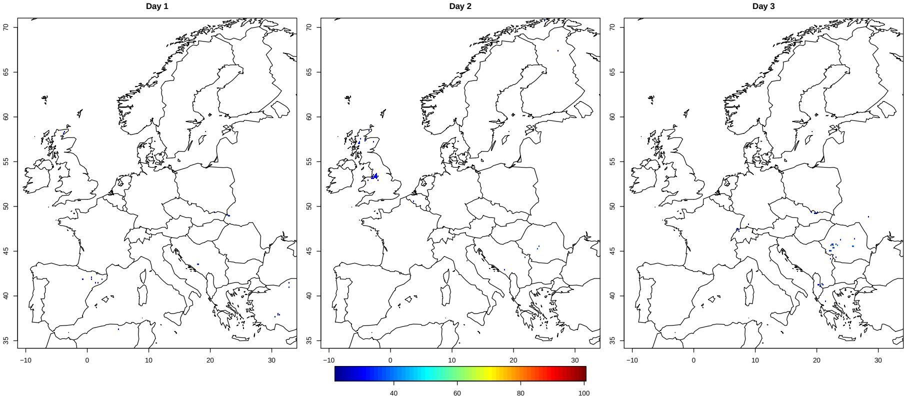
1. Choose times $t = 1, \ldots, T$ and locations $s_1, \ldots, s_D$ (perhaps a grid) for which simulated gust speeds are sought
2. Set $t = 1$
3. Simulate $Z(s_1, t), \ldots, Z(s_D, t)$ from suitable geostatistical model
4. For $d = 1, \ldots, D$, if $Z(s_d, t) > logit^{-1}(\mu(s))$ set $Y(s_d, t) > u$
5. Convert margins to original scale, so that if $Y(s_d, t) > u$

$$Y(s_d, t) = F_u^{-1}(\Phi(Z(s_d)); \psi(s_d), \xi(s_d))$$

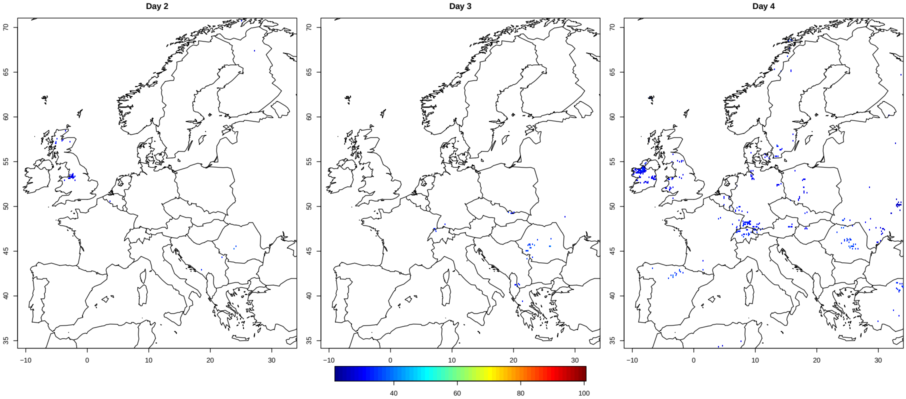
6. Update $t = t + 1$ and return to Step 3

Simulated events

- Some simulated events, chosen arbitrarily

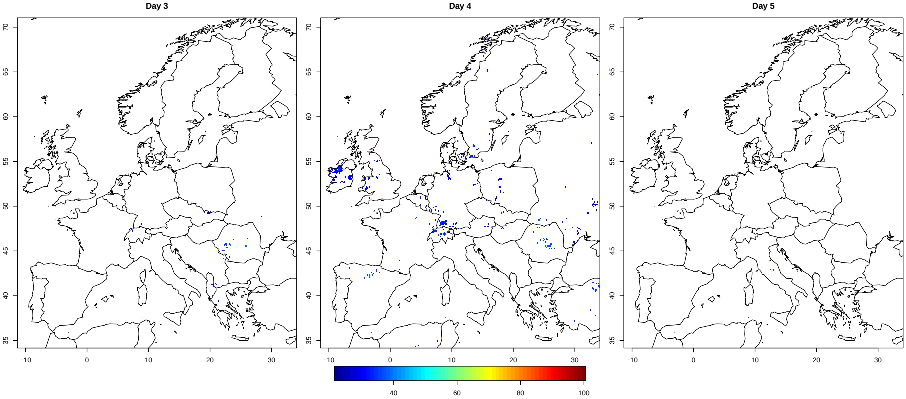


- Some simulated events, chosen arbitrarily



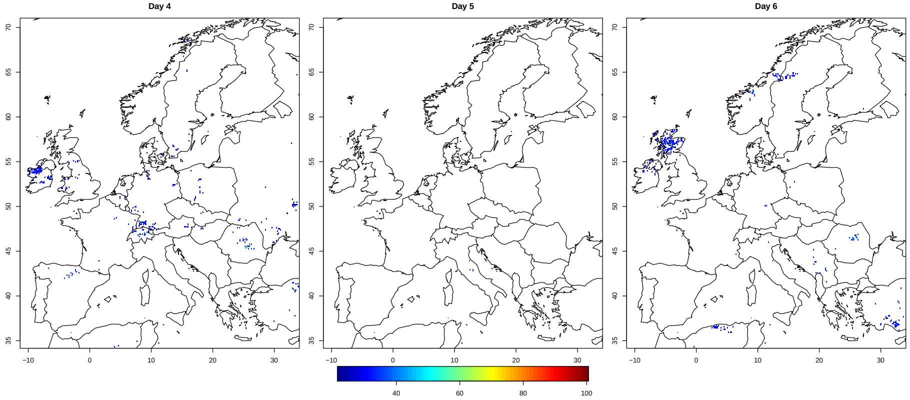
Simulated events

- Some simulated events, chosen arbitrarily



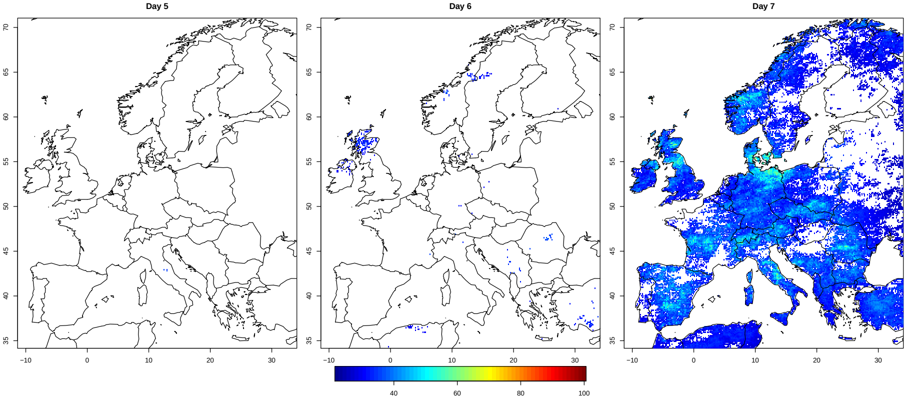
Simulated events

- Some simulated events, chosen arbitrarily



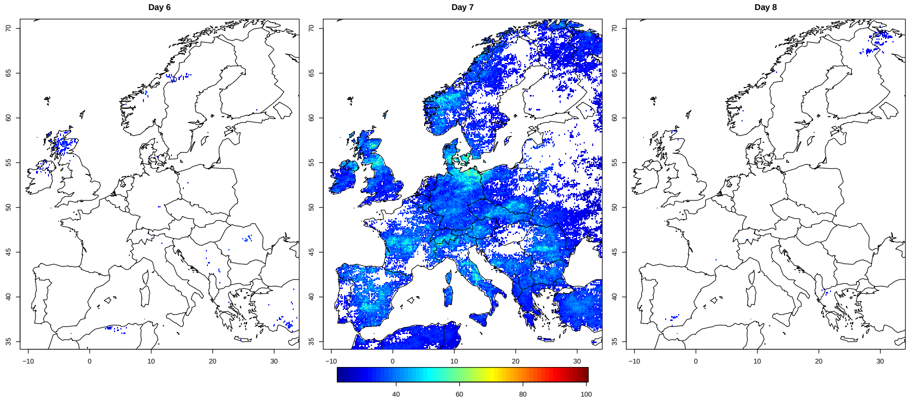
Simulated events

- Some simulated events, chosen arbitrarily



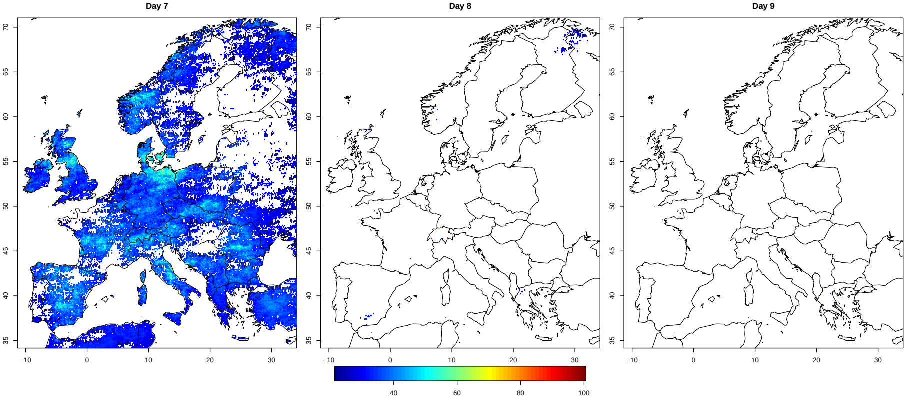
Simulated events

- Some simulated events, chosen arbitrarily



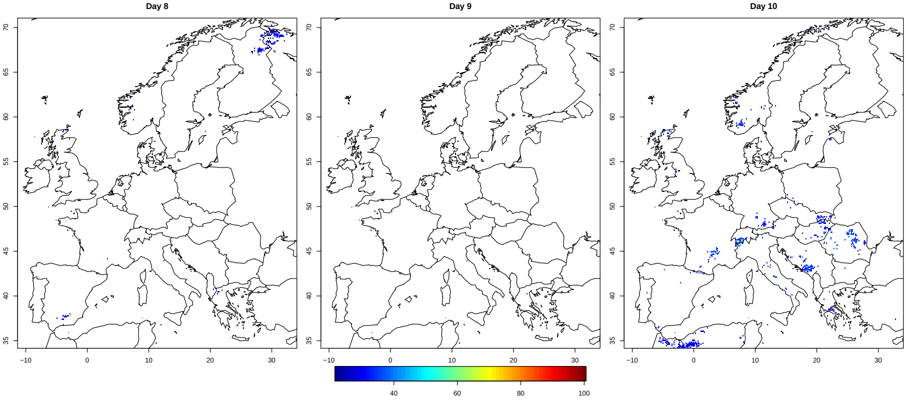
Simulated events

- Some simulated events, chosen arbitrarily



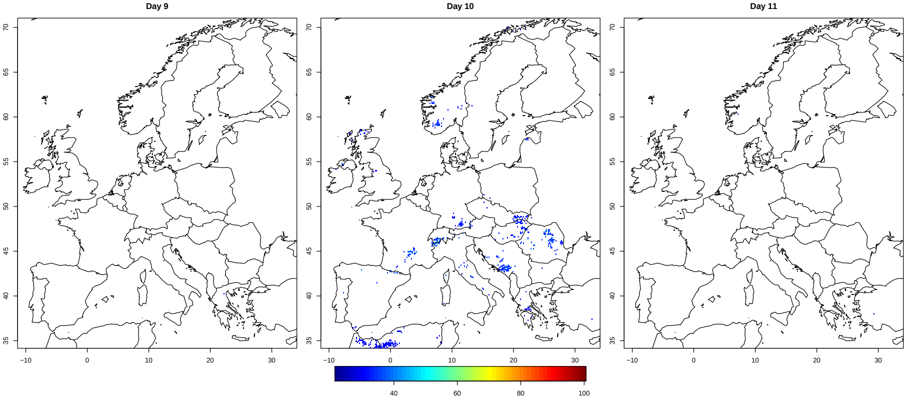
Simulated events

- Some simulated events, chosen arbitrarily



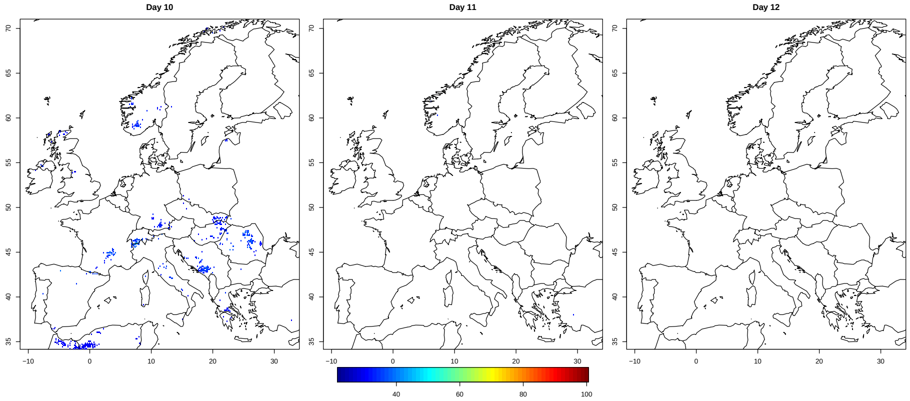
Simulated events

- Some simulated events, chosen arbitrarily



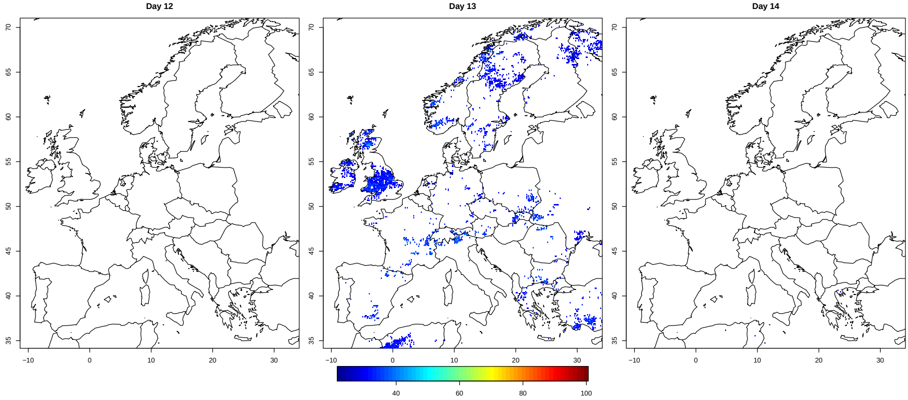
Simulated events

- Some simulated events, chosen arbitrarily



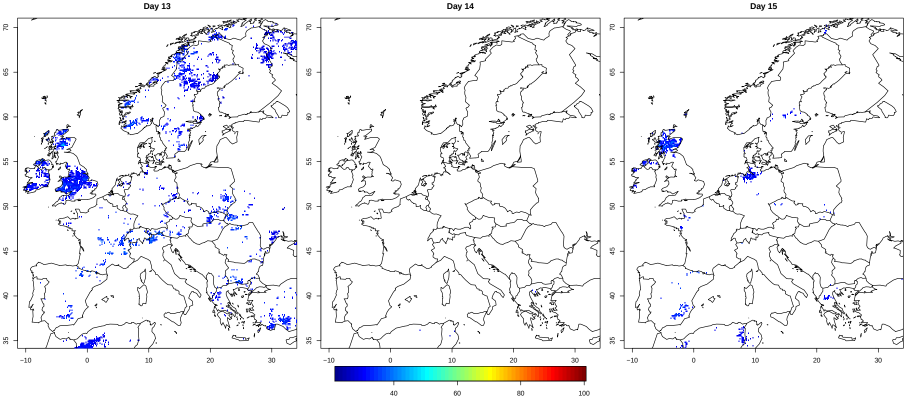
Simulated events

- Some simulated events, chosen arbitrarily



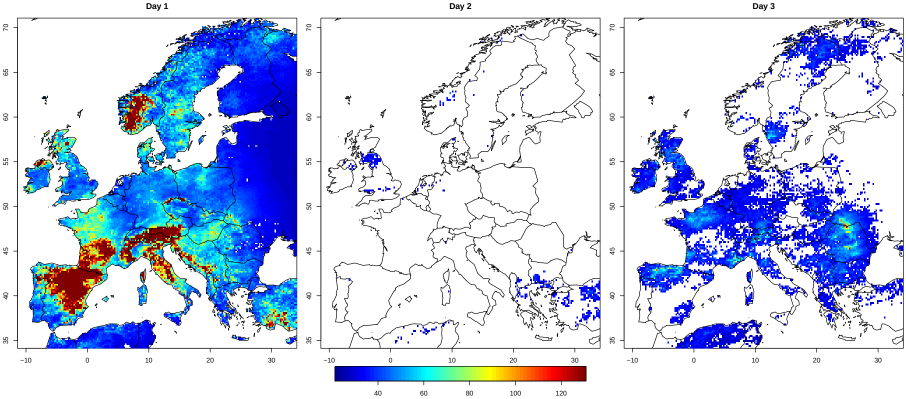
Simulated events

- Some simulated events, chosen arbitrarily



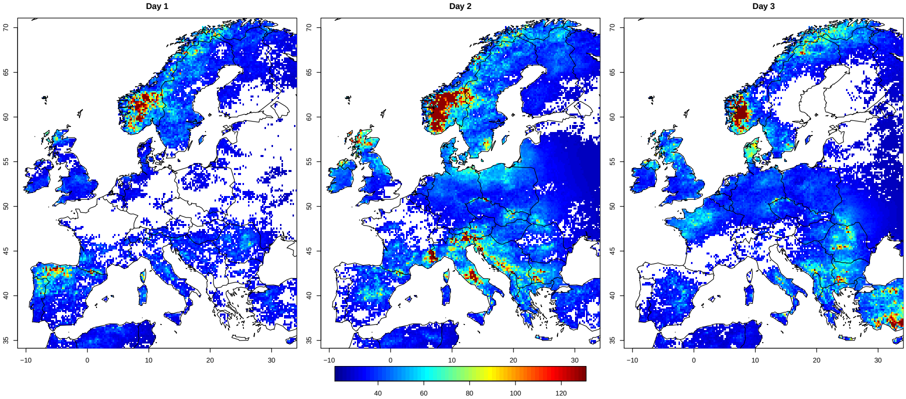
'Extreme' simulated events

- Largest storm severity index, over three days



'Extreme' simulated events

- Largest exceedance area, relative to 22m/s threshold



- Simulation of entire years, not just DJF winters
- Temporally nonstationary generalised Pareto distribution parameters
  - trends over times
  - dependence on climate indices
- Max-stable models
- Quality control on station data
- Simulation of ‘actual’ gust speeds, as opposed to measurements