# Spatio temporal analysis of extreme wind velocities for infrastructure desing. Case study Colombia

---

A Thesis
Presented to
The Division of Instituto for Geoinformatics - IFGI
University of Münster

---

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in Geospatial Technologies

---

Alexys Herleym Rodríguez Avellaneda

Jan 2020

Approved for the Division
()

Dr. Edzer Pebesma | Dr. Juan C. Reyes

# Acknowledgements

I want to thank a few people.

# Preface

This is an example of a thesis setup to use the reed thesis document class (for LaTeX) and the R bookdown package, in general.

# Table of Contents

# List of Tables

# List of Figures

# Abstract

The preface pretty much says it all.

Second paragraph of abstract starts here.

# Dedication

You can have a dedication here if you wish.

# Introduction

Placeholder

# Chapter 1

# Data

Placeholder

## 1.1 IDEAM

## 1.2 ISD

## 1.3 ERA5

## 1.4

## 1.5 Data Download and Organization

## 1.6 Data Standarzation

# Chapter 2

# Theoretical Framework

## 2.1 Probability Concepts

Poisson process is an stochastic method that relies in the concepts of probability distributions. The main functions related to probability for extreme value analysis will be described below.

### 2.1.1 Probability Density Function - *pdf*

Pdf defines the probability that a continuos variable falls between two points, this is, in *pdf* the proability is related to the area below the curve (integral) between two points, as for continuos probability distributions the probability at a single point is zero. The term density is directly related to the probability of a portion of the curve, if the density function has high values the probability will be greater in comparison with the same portion of curve for low values.

$$\int_a^b f(x)dx = Pr[a \leq X \leq b]$$

Equation (2.1) is the Gumbel *pdf*.

$$\mathrm{f(x)} = \frac{1}{\beta} \exp\left\{-\frac{\mathrm{x}-\mu}{\beta}\right\} \exp\left\{-\exp\left\{-\left(\frac{\mathrm{x}-\mu}{\beta}\right)\right\}\right\}, \quad -\infty < \mathrm{x} < \infty \tag{2.1}$$

where $\exp\{.\} \mapsto \mathrm{e}^{\{.\}}$, $\beta$ is the scale parameter, and $\mu$ is the location parameter. Location ($\mu$) has the effect to shift the *pdf* to left or right along 'x' axis, thus, if location value is changed the effect is a movement of *pdf* to the left (small value for location), or to the right (big value for location). Scale has the effect to stretch ($\beta > 1$) of compress ($0 < \beta < 1$) the *pdf*, if scale parameter is close to zero the pdf approaches a spike.

Figure 2.1 shows *pdf* with location ($\mu$) = 100 and scale ($\beta$) = 40, using equation (2.1).

```
location = 100
scale = 40
.x <- seq(0, 300, length.out=1000)
pdfG <- function(x) {
  1/location *exp(-(x-location)/scale)*exp(-exp(-(x-location)/scale))
  }
.y = pdfG(.x)
plot(.x, .y, col="green", lty=4,
     xlab="Velocities Km/h", ylab="Density Function - Gumbel Distribution",
     main=paste("Gumbel - Density Function Gumbel Distribution\n", "Location=",
     round(location,2), " Scale=", round(scale,2)), type="l",
     cex.axis = 0.5, cex.lab= 0.6, cex.main=0.7, cex.sub=0.6)
```

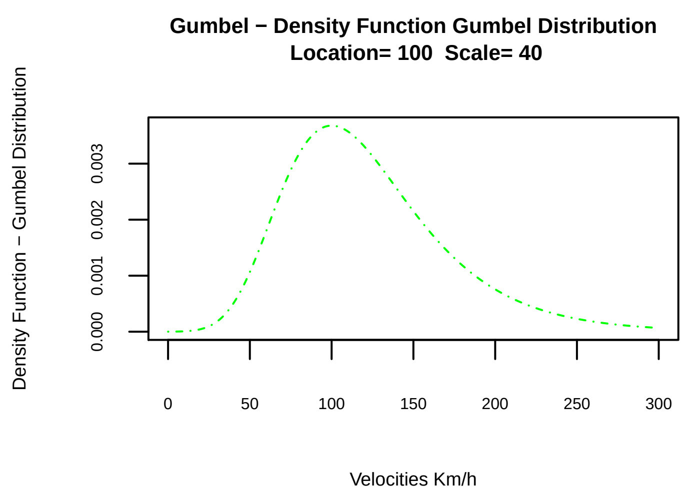


Figure 2.1: Gumbel pdf

Figure 2.2 shows *pdf* with location $(\mu) = 100$ and scale $(\beta) = 40$, using function `dgumbel` of the package `RcmdrMisc`

```
location = 100
scale = 40
.x <- seq(0, 300, length.out=1000)
dfG = dgumbel(.x, location=location, scale=scale)
plot(.x, dfG, col="red", lty=4,
     xlab="Velocities Km/h", ylab="Density Function - Gumbel Distribution",
     main=paste("Gumbel - Density Function Gumbel Distribution\n", "Location=",
     round(location,2), " Scale=", round(scale,2)), type="l",
     cex.axis = 0.5, cex.lab= 0.6, cex.main=0.7, cex.sub=0.6)
```

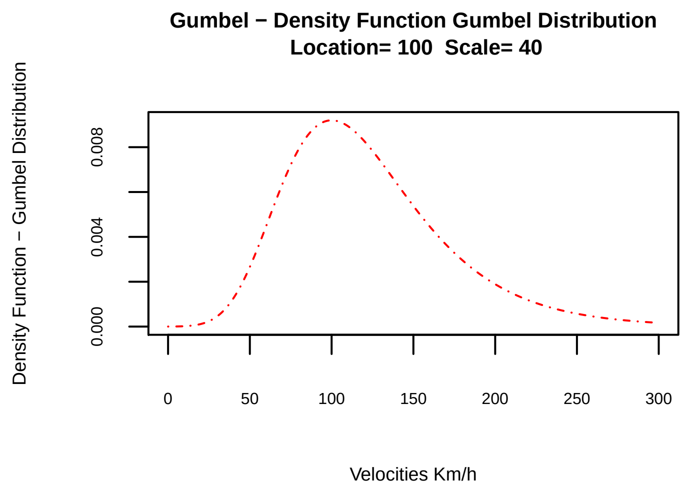


Figure 2.2: Gumbel pdf - dgumbel function

### 2.1.2 Cumulative Distribution Funtcion - *cdf*

*Cdf* is the probability of taking a value less than or equal to x. That is

$$F(x) = Pr[X < x] = \alpha$$

For a continuous variable, *cdf* can be expressed as the integral of its *pdf*.

$$F(x) = \int_{-\infty}^{x} f(x)dx$$

Equation (2.2) is the Gumbel *cdf*.

$$\mathrm{F(x)} = \exp\left\{-\exp\left[-\left(\frac{\mathrm{x}-\mu}{\beta}\right)\right]\right\}, \quad -\infty < \mathrm{x} < \infty \tag{2.2}$$

Figure 2.3 shows Gumbel *cdf* with location ($\mu$) = 100 and scale ($\beta$) = 40, using equation (2.2). As previously done with *pdf*, similar result can be achieved using function `pgumbel` of package `RcmdrMisc`.

```
location = 100
scale = 40
.x <- seq(0, 300, length.out=1000)
cdfG <- function(x) {
  exp(-exp(-(x-location)/scale))
  }
.y = cdfG(.x)
plot(.x, .y, col="green", lty=4,
     xlab="Velocities Km/h", ylab="Probability",
     main=paste("Gumbel - Cumulative Distribution Function\n", "Location=",
     round(location,2), " Scale=", round(scale,2)), type="l",
     cex.axis = 0.5, cex.lab= 0.6, cex.main=0.7, cex.sub=0.6)
```

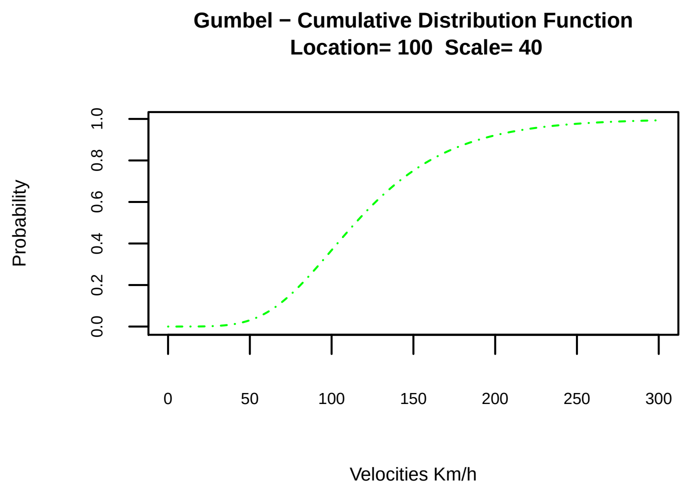


Figure 2.3: Gumbel cdf

### 2.1.3 Percent Point Function - *ppf*

*Ppf* is the inverse of *cdf*, also called the *quantile* function. This is, from a specific probability get the corresponding value x of the variable.

$$x = G(\alpha) = G(F(x))$$

Equation (2.3) is the Gumbel *ppf*.

$$\mathrm{G}(\alpha) = \mu - \beta \ln(-\ln(\alpha)) \quad 0 < \alpha < 1 \tag{2.3}$$

Figure 2.4 shows Gumbel *ppf*, using equation (2.3). Similar result can be achieved using function `qgumbel` of package `RcmdrMisc`.

```
location = 100
scale = 40
.x <- seq(0, 1, length.out=1000)
ppfG <- function(x) {
  location - (scale*log(-log(x)))
  }
.y = ppfG(.x)
plot(.x, .y, col="green", lty=4,
     ylab="Velocities Km/h", xlab="Probability",
     main=paste("Gumbel - Percent Point Function\n", "Location=",
     round(location,2), " Scale=", round(scale,2)), type="l",
     cex.axis = 0.5, cex.lab= 0.6, cex.main=0.7, cex.sub=0.6)
```

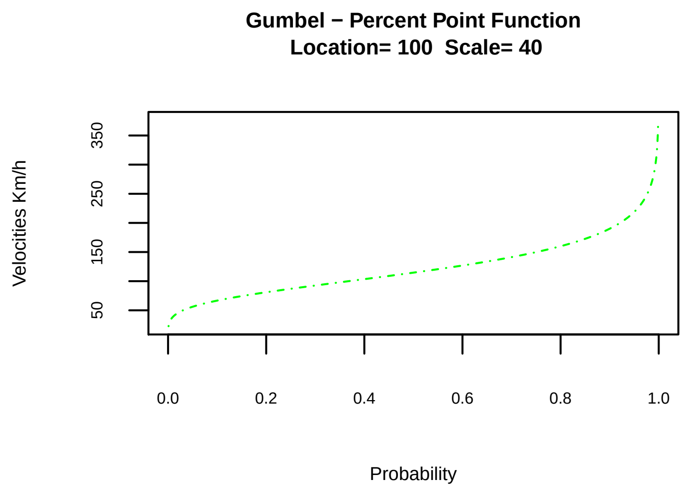


Figure 2.4: Gumbel cdf

### 2.1.4 Hazard Function - *hf*

Using $S(x) = 1 - F(x)$ as survival function -*sf*, the probability that a variable takes a value greather than x $S(x) = Pr[X > x] = 1 - F(x)$, the *hf* is the ratio between *pdf* and *sf*.

$$h(x) = \frac{f(x)}{S(x)} = \frac{f(x)}{1 - F(x)}$$

Equation (2.4) is the Gumbel *ppf*.

$$\mathrm{h(x)} = \frac{1}{\beta} \frac{\exp(-(\mathrm{x} - \mu)/\beta)}{\exp(\exp(-(\mathrm{x} - \mu)/\beta)) - 1} \tag{2.4}$$

Figure 2.5 shows Gumbel *hf*, using equation (2.4).

```
location = 100
scale = 40
.x <- seq(0, 3000, length.out=1000)
hfG <- function(x) {
  (1/scale)*(exp(-(x-location)/scale))/(exp(exp(-(x-location)/scale))-1)
  }
.y = hfG(.x)
plot(.x, .y, col="green", lty=4,
     xlab="Velocities Km/h", ylab="Hazard",
     main=paste("Gumbel - Hazard Function\n", "Location=",
     round(location,2), " Scale=", round(scale,2)), type="l",
     cex.axis = 0.5, cex.lab= 0.6, cex.main=0.7, cex.sub=0.6)
```

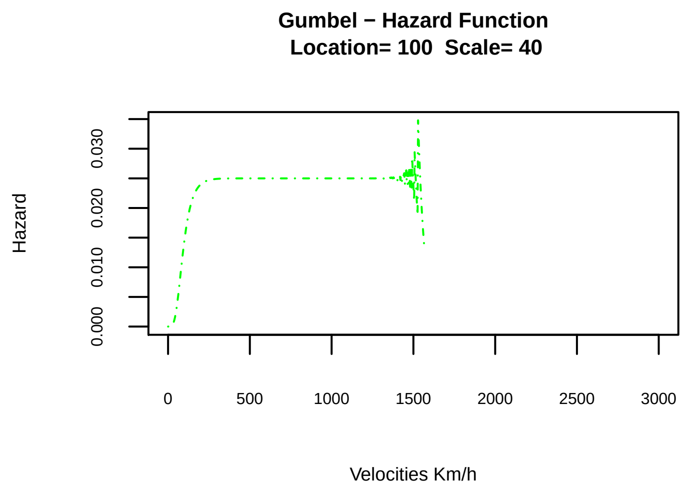


Figure 2.5: Gumbel cdf

```
#library(reliaR)
#plot(.x, hgumbel(.x, mu=location, sigma=scale))
#plot(.x, hra.gumbel(.x, mu=location, sigma=scale))
```

### 2.1.5 Cumulative Hazard Function

$$\sum_{j=1}^{n}(\delta\theta_j)^2 \leq \frac{\beta_i^2}{\delta_i^2+\rho_i^2}\left[2\rho_i^2+\frac{\delta_i^2\beta_i^2}{\delta_i^2+\rho_i^2}\right] \equiv \omega_i^2$$

$$\sum_{\mathrm{j}=1}^{\mathrm{n}}(\delta\theta_{\mathrm{j}})^2 \leq \frac{\beta_{\mathrm{i}}^2}{\delta_{\mathrm{i}}^2+\rho_{\mathrm{i}}^2}\left[2\rho_{\mathrm{i}}^2+\frac{\delta_{\mathrm{i}}^2\beta_{\mathrm{i}}^2}{\delta_{\mathrm{i}}^2+\rho_{\mathrm{i}}^2}\right] \equiv \omega_{\mathrm{i}}^2 \tag{2.5}$$

We can reference this combustion of glucose reaction via Equation (2.1).

$$\sum_{j=1}^{n}(\delta\theta_j)^2 \leq \frac{\beta_i^2}{\delta_i^2+\rho_i^2}\left[2\rho_i^2+\frac{\delta_i^2\beta_i^2}{\delta_i^2+\rho_i^2}\right] \equiv \omega_i^2$$

TeX is the best way to typeset mathematics. Donald Knuth designed TeX when he got frustrated at how long it was taking the typesetters to finish his book, which contained a lot of mathematics. One nice feature of *R Markdown* is its ability to read LaTeX code directly.

If you are doing a thesis that will involve lots of math, you will want to read the following section which has been commented out. If you're not going to use math, skip over or delete this next commented section.

$$\sum_{j=1}^{n}(\delta\theta_j)^2 \leq \frac{\beta_i^2}{\delta_i^2+\rho_i^2}\left[2\rho_i^2+\frac{\delta_i^2\beta_i^2}{\delta_i^2+\rho_i^2}\right] \equiv \omega_i^2$$

From Informational Dynamics, we have the following (Dave Braden):

Another equation:

$$\det \begin{vmatrix} c_0 & c_1 & c_2 & \dots & c_n \\ c_1 & c_2 & c_3 & \dots & c_{n+1} \\ c_2 & c_3 & c_4 & \dots & c_{n+2} \\ \vdots & \vdots & \vdots & & \vdots \\ c_n & c_{n+1} & c_{n+2} & \dots & c_{2n} \end{vmatrix} > 0$$

Lapidus and Pindar, Numerical Solution of Partial Differential Equations in Science and Engineering. Page 54

$$\int_t \left\{ \sum_{j=1}^{3} T_j \left( \frac{d\phi_j}{dt} + k\phi_j \right) - kT_e \right\} w_i(t)\ dt = 0, \qquad i = 1, 2, 3.$$

L&P Galerkin method weighting functions. Page 55

$$\sum_{j=1}^{3} T_j \int_0^1 \left\{ \frac{d\phi_j}{dt} + k\phi_j \right\} \phi_i\ dt = \int_0^1 k\, T_e \phi_i dt, \qquad i = 1, 2, 3$$

Another L&P (p145)

$$\int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} f\Big(\xi, \eta, \zeta\Big) = \sum_{k=1}^{n}\sum_{j=1}^{n}\sum_{i=1}^{n} w_i w_j w_k f\Big(\xi, \eta, \zeta\Big).$$

Another L&P (p126)

$$\int_{A_e} (\,\cdot\,) dx dy = \int_{-1}^{1}\int_{-1}^{1} (\,\cdot\,) \det[J] d\xi d\eta.$$

## 2.2 Annual Excedance Probability - Pa

Chemical formulas will look best if they are not italicized. Get around math mode's automatic italicizing in LaTeX by using the argument `$\mathrm{formula here}$`, with your formula inside the curly brackets. (Notice the use of the backticks here which enclose text that acts as code.)

So, $\mathrm{Fe_2^{2+}Cr_2O_4}$ is written `$\mathrm{Fe_2^{2+}Cr_2O_4}$`.
Exponent or Superscript: $\mathrm{O^-}$
Subscript: $\mathrm{CH_4}$

To stack numbers or letters as in $\mathrm{Fe_2^{2+}}$, the subscript is defined first, and then the superscript is defined.
Bullet: $\mathrm{CuCl} \bullet \mathrm{7H_2O}$
Delta: $\Delta$
Reaction Arrows: $\longrightarrow$ or $\xrightarrow{solution}$
Resonance Arrows: $\leftrightarrow$
Reversible Reaction Arrows: $\rightleftharpoons$

### 2.2.1 Typesetting reactions

You may wish to put your reaction in an equation environment, which means that LaTeX will place the reaction where it fits and will number the equations for you.

$$C_6H_{12}O_6 + 6O_2 \longrightarrow 6CO_2 + 6H_2O \tag{2.6}$$

We can reference this combustion of glucose reaction via Equation (2.6).

### 2.2.2 Other examples of reactions

$NH_4Cl_{(s)} \rightleftharpoons NH_{3(g)} + HCl_{(g)}$
$MeCH_2Br + Mg \xrightarrow[below]{above} MeCH_2 \bullet Mg \bullet Br$

## 2.3 Return Period

Many of the symbols you will need can be found on the math page `http://web.reed.edu/cis/help/latex/math.html` and the Comprehensive LaTeX Symbol Guide (`http://mirror.utexas.edu/ctan/info/symbols/comprehensive/symbols-letter.pdf`).

## 2.4 Compound Excedance Probability - Pn

You will probably find the resources at `http://www.lecb.ncifcrf.gov/~toms/latex.html` helpful, particularly the links to bsts for various journals. You may also be interested in TeXShade for nucleotide typesetting (`http://homepages.uni-tuebingen.de/beitz/txe.html`). Be sure to read the proceeding chapter on graphics and tables.

## 2.5 Extreme Value Analysis Overview

### 2.5.1 Main Methods

#### Epochal methods

#### Peak Over Threshold

#### GPD

#### Poisson Process

### 2.5.2 Commond Distributions for Extreme Values

### 2.5.3 Methods for parameters estimation

### 2.5.4 Return Period

### 2.5.5 Wind Speed at Return Period

## 2.6 Peaks Over Threshold - Poisson Process

$$\lambda\left(\mathrm{y},\mathrm{t}\right)\left\{\lambda_s(y), fortinthunderstormperiod\lambda_n s(y), fortinnon-thunderstormperiod\right. \tag{2.7}$$

### 2.6.1 Data Requirements

### 2.6.2 Maximun Likelihood

After $n$ such encounters the posterior density for $\theta$ is

$$f(x) = \frac{1}{\beta}\exp\left\{-\frac{x-\mu}{\beta}\right\}\exp\left\{-\exp\left\{-\left(\frac{x-\mu}{\beta}\right)\right\}\right\}, \quad -\infty < x < \infty$$

$$L(\eta) = \left(\prod_{i=1}^{I}\lambda\left(y_i,t_i\right)\right)\exp\left\{-\int_D \lambda\left(y,t\right)dydt\right\}$$

$$\pi(\theta|X_1 < y_1,\ldots,X_n < y_n) \propto \pi(\theta)\prod_{i=1}^{n}\int_{-\infty}^{y_i}\exp\left(-\frac{(x-\theta)^2}{2\sigma^2}\right)\, dx$$

$$\mathrm{L}(\eta) = \left(\prod_{\mathrm{i}=1}^{\mathrm{I}}\lambda\left(\mathrm{y_i},\mathrm{t_i}\right)\right)\exp\left\{-\int_{\mathrm{D}}\lambda\left(\mathrm{y},\mathrm{t}\right)\mathrm{dydt}\right\} \tag{2.8}$$

where $\exp\{.\} \mapsto \mathrm{e}^{\{.\}}$, $\beta$ is the scale parameter, and $\mu$ is the location parameter. Location ($\mu$) has the effect to shift the *pdf* to left or right along 'x' axis, thus, if location value is changed the effect is a movement of *pdf* to the left (small value for location), or to the right (big value for location). Scale has the effect to stretch ($\beta > 1$) of compress ($0 < \beta < 1$) the *pdf*, if scale parameter is close to zero the pdf approaches a spike.

Figure **??** shows *pdf* with location ($\mu$) = 100 and scale ($\beta$) = 40, using equation (2.1).

### 2.6.3 Recurrence Intervals

$$\int_{\mathrm{yN}}^{\infty}\lambda\left(\mathrm{y},\mathrm{t}\right)\mathrm{dydt} = \frac{1}{\mathrm{N}} \tag{2.9}$$

###Declustering Samples ###Threshold Selection

$$U = F(Y)$$
$$W = -log(1 - U)$$

### 2.6.4

# Chapter 3

# Methodology

Placeholder

## 3.1 Input Data Selection and Standarization

### 3.1.1 Data Selection

### 3.1.2 Data Standarization

#### Anemometer height - 10 m

#### Surface Roughness - 0.03 m

#### Averaging Time - 3-s gust

## 3.2 Fit data to a POT - Poisson Process

### 3.2.1 Velocities at Return Periods

## 3.3 spatial Interpolation

## 3.4 Footnotes and Endnotes

## 3.5 Bibliographies

## 3.6 Anything else?

# Conclusion

If we don't want Conclusion to have a chapter number next to it, we can add the `{-}` attribute.

**More info**

And here's some other random info: the first paragraph after a chapter title or section head *shouldn't be* indented, because indents are to tell the reader that you're starting a new paragraph. Since that's obvious after a chapter or section title, proper typesetting doesn't add an indent there.

# Appendix A

# The First Appendix

This first appendix includes all of the R chunks of code that were hidden throughout the document (using the `include = FALSE` chunk tag) to help with readibility and/or setup.

**In the main Rmd file**

**In Chapter 3:**

# Appendix B

# The Second Appendix, for Fun

# References

Placeholder