



# CART Trees and Random Forests - Jean-Michel POGGI Master 2 Course in Statistics

## Universidad de la República – Facultad de Ingeniería, Montevideo, Uruguay

## February 2025

Guide for reading the paper:

Bel, L., Allard, D., Laurent, J. M., Cheddadi, R., & Bar-Hen, A. CART algorithm for spatial data: Application to environmental and ecological data. Computational Statistics & Data Analysis, 53(8), 3082-3093, 2009

The idea is to produce a report (of 5 to 10 pages) presenting and discussing the paper given in reference. Please find below some suggestions.

## 1. The problem

- 1. Regression or classification?
- 2. Geostatistics or point processes approaches?

### 2. How to handle the spatial dependence to modify CART?

- 1. Drawbacks of the usual CART in this specific situation?
- 2. Summarize the two proposed approaches.
- 3. Discuss pros and cons.

### 3. CART variant implementations

- 1. Discuss the ways to implement the variants starting from a given implementation of CART
- 2. Compare the two methods on simulated dataset

### 4. Real data example

- 1. Problem presentation
- 2. Discuss the application of the different variants w.r.t. classical CART
- 3. Conclude from the ecological point of view

### 5. Discussion

- 1. Discuss the interest/contributions of this paper
- 2. Suggest some extensions to deal with spatial data from a different point of view (optional)