

# Fitting ENM Random Forest

Emilio Berti

# How Random Forest works

1. Create  $n$  decision trees.
2. Average or majority voting across all trees.

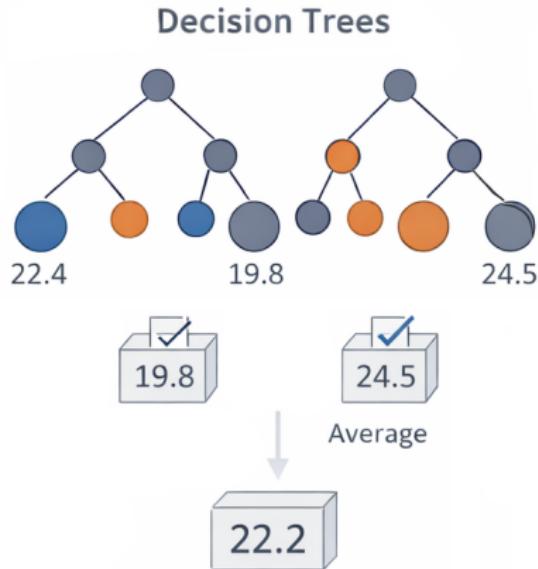


Figure 1: Conceptual representation of Random Forest algorithm

## Load virtual species

Already created, just load it.

```
d <- read.csv("../data/virtual.csv")  
head(d)
```

	BI001	BI012	occ
1	9.987156	897	1
2	9.333240	472	1
3	9.853782	663	1
4	7.313562	1025	1
5	6.567688	689	1
6	11.305479	807	1

## Fitting Random Forest

RF is a powerful machine-learning algorithm that has been applied extensively for ENM.

In *R*, RF is achieved using the package `randomForest` with the following specifications.

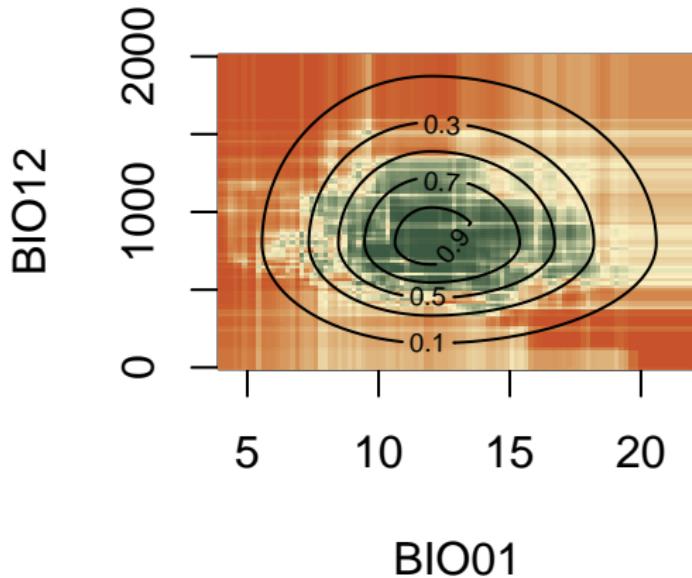
```
library(randomForest)

enm <- randomForest(
  occ ~ BI001 + BI012,
  data = d
)
```

### Note

A warning can be silenced with `as.factor(occ)`, which makes `randomForest()` consider a classification. We do not want this.

## Random Forest inferred niche



Predictions from the RF model (colors) kinda match the actual niche (black lines).

The niche inferred by RF is very jagged.

This is due to how RF works internally and can only be partly fixed.