Norm of reaction for maritime pine

Victor Papin, Laurent Bouffier, Léopoldo Sanchez

To cite this version:

Victor Papin, Laurent Bouffier, Léopoldo Sanchez. Norm of reaction for maritime pine. B4EST Training course ”Phenotypic plasticity and norms of reaction modelling”, B4EST project, Oct 2021, On Line meeting, France. hal-04717539

HAL Id: hal-04717539
https://hal.inrae.fr/hal-04717539v1

Submitted on 1 Oct 2024

Case study 3:
Norm of reaction for maritime pine

Victor Papin, INRAE (UMR BIOGECO Pierroton)
Laurent Bouffier, INRAE (UMR BIOGECO Pierroton)
Léopoldo Sanchez, INRAE (UMR BioForA Orléans)

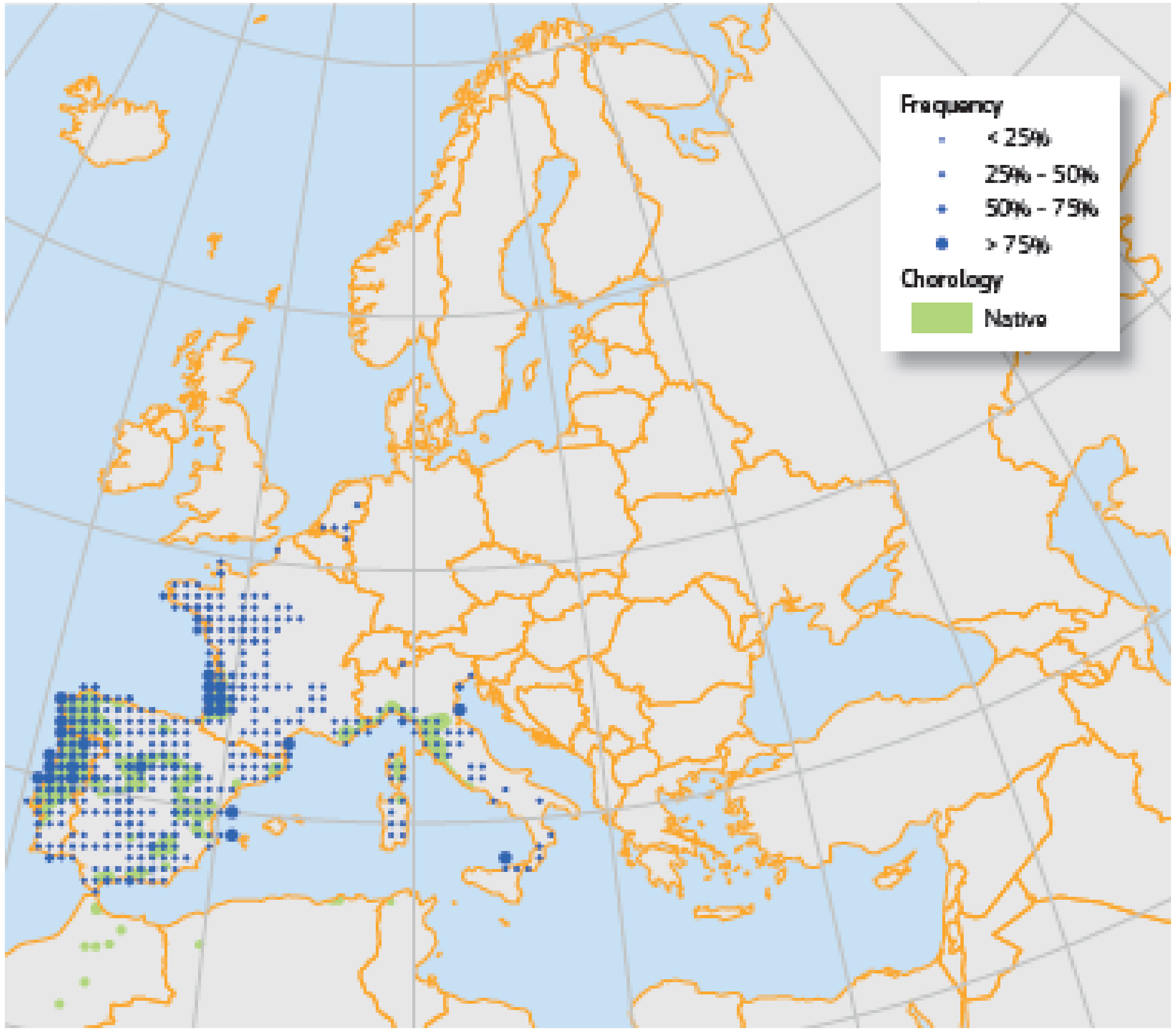
Outlines

- Maritime pine (*Pinus pinaster*)
- GxE in the French maritime pine breeding context
- Genetic component of individual norm of reaction (NoR)
- Concept of Random Regression (RR)
- Application on maritime pine based on wood longitudinal data

Maritime pine (*Pinus pinaster*)

- Natural distribution area (green) and field observations (frequency of occurrences)

Maritime pine clonal archive

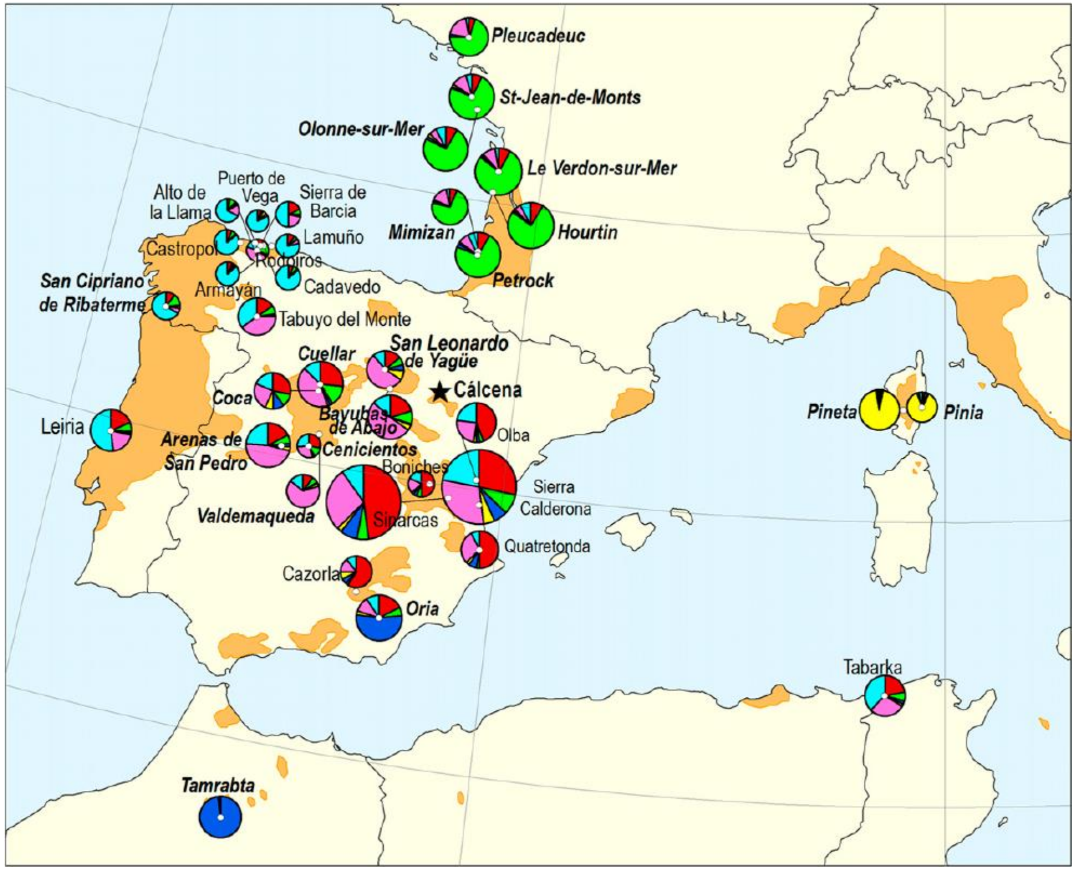


Abad Viñas, R., Caudullo, G., Oliveira, S., de Rigo, D., 2016. ***Pinus pinaster* in Europe: distribution, habitat, usage and threats.** In: San-Miguel-Ayanz, J., de Rigo, D., Caudullo, G., Houston Durrant, T., Mauri, A. (Eds.), *European Atlas of Forest Tree Species*. Publ. Off. EU, Luxembourg, pp. e012d59+

Maritime pine (*Pinus pinaster*)

- High genetic structuration:

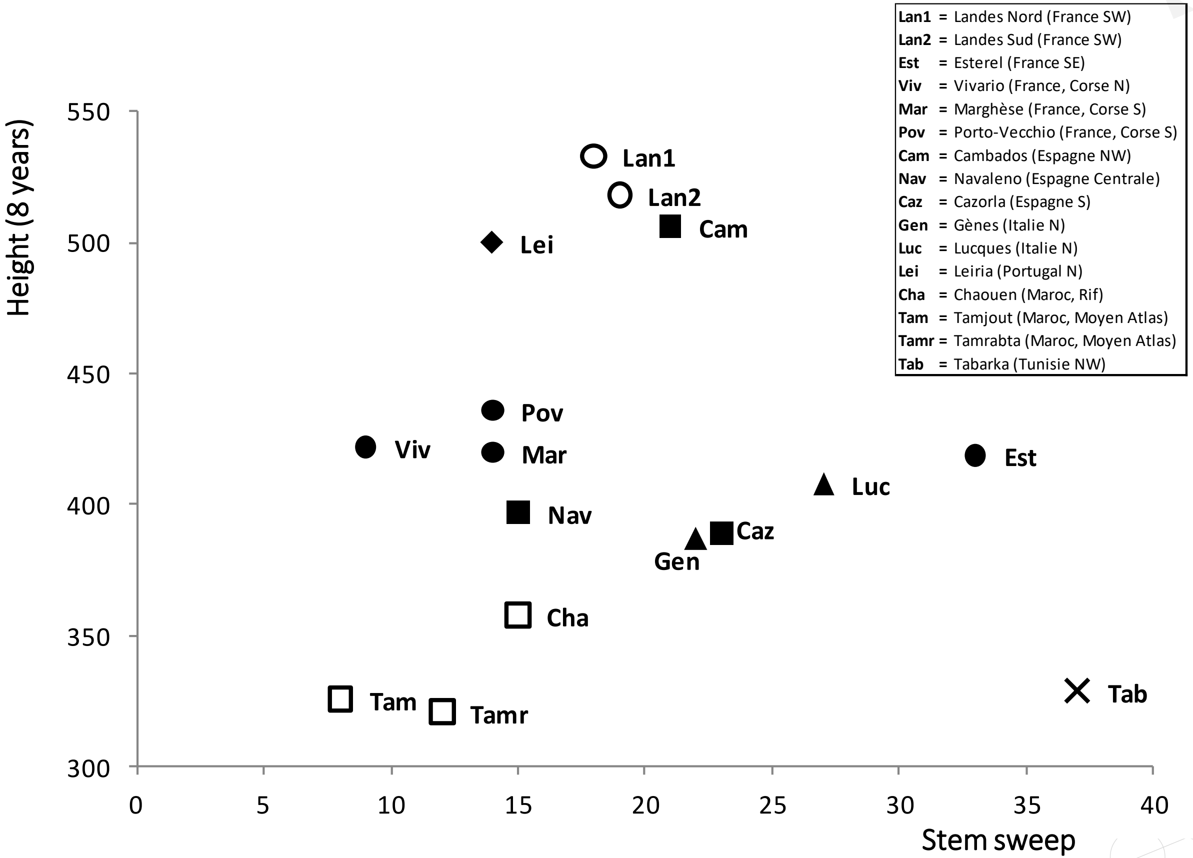
6 main genepools identified from 36 natural populations



From Jaramillo-Correa *et al.* (2015)

- Phenotypic variability between provenances:

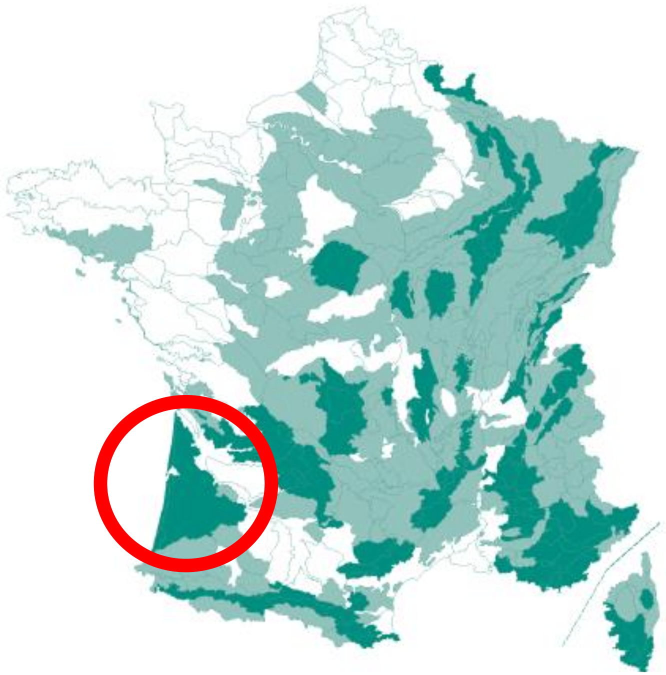
results from a progeny trial established in France



Lan1 = Landes Nord (France SW)
Lan2 = Landes Sud (France SW)
Est = Esterel (France SE)
Viv = Vivario (France, Corse N)
Mar = Marghèse (France, Corse S)
Pov = Porto-Vecchio (France, Corse S)
Cam = Cambados (Espagne NW)
Nav = Navaleno (Espagne Centrale)
Caz = Cazorla (Espagne S)
Gen = Gènes (Italie N)
Luc = Lucques (Italie N)
Lei = Leiria (Portugal N)
Cha = Chaouen (Maroc, Rif)
Tam = Tamjout (Maroc, Moyen Atlas)
Tamr = Tamrabta (Maroc, Moyen Atlas)
Tab = Tabarka (Tunisie NW)

From Harfouche (1995)

Maritime pine breeding in France



- **“Landes de Gascogne” forest:** 0.8 million hectares (7% of the French forest - 24% of French wood harvest): ➔ high productivity in a limited area
- plantations with **improved seedlings** except in sand dunes (natural regeneration)
- **One breeding zone**
  - Similar climatic conditions (dry summers and wet winters)
  - Low variation in site types (poor soils with sandy podzol)

- Breeding population mainly based on the **local provenance** (Landes provenance)

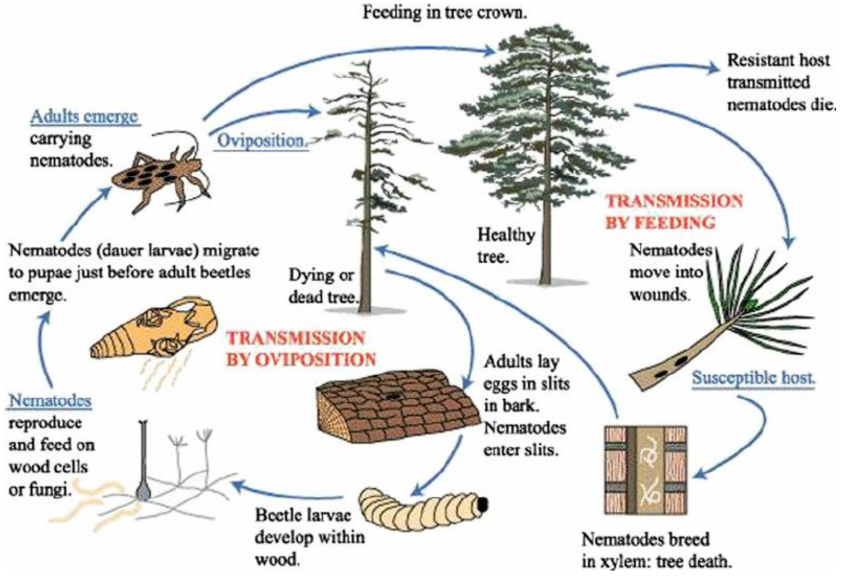
- **Generalist varieties** produced in open seed orchards ➔ additive effect

In that breeding context: why studying GxE ?

- Even if environmental variation is limited in the current breeding zone, varieties could be designed per site type to increase genetic gains

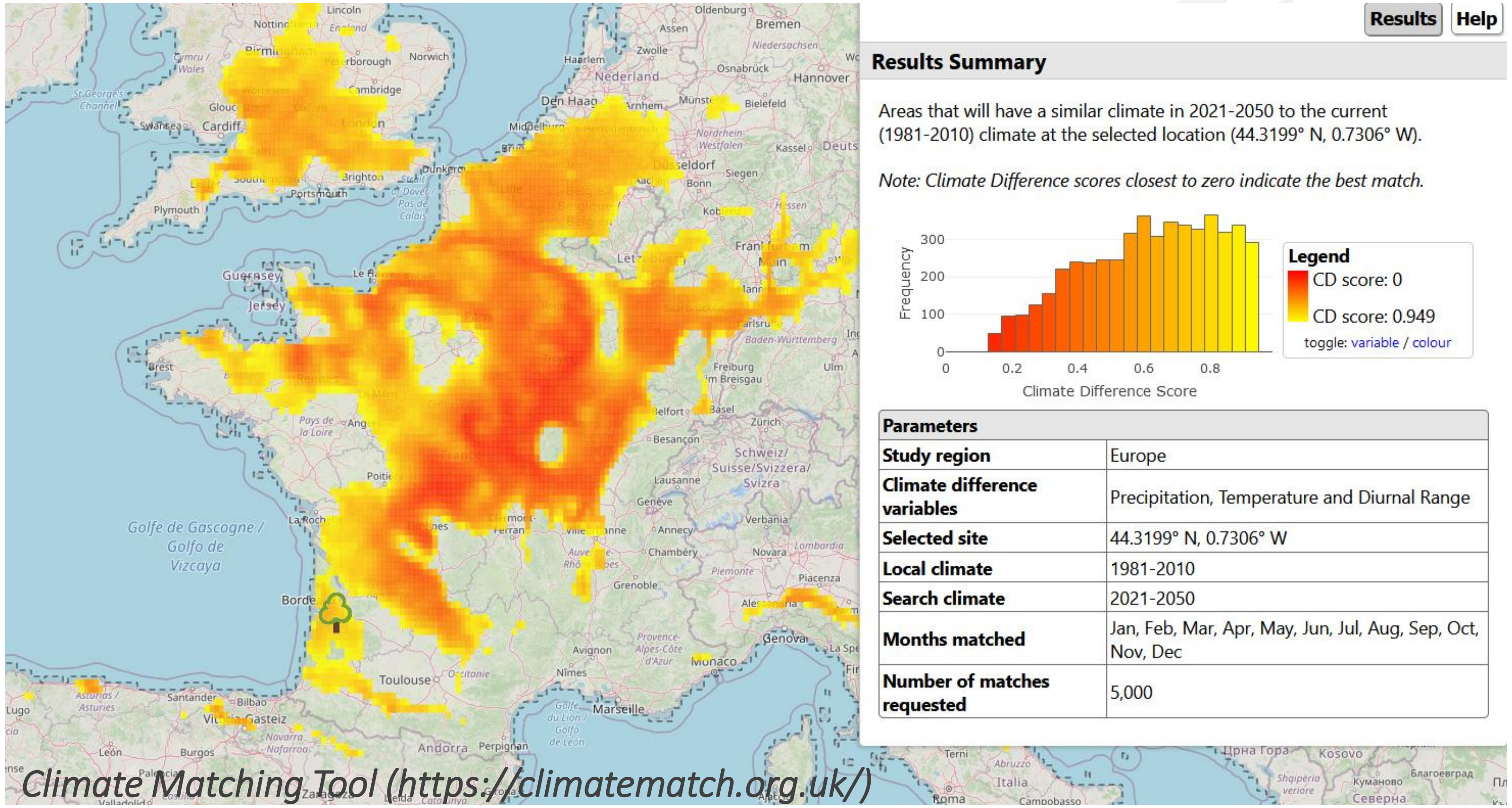
  ➔ Need to better understand GxE to design specialized varieties

- Anticipation for futur environments: climate change and new pests (pine wood nematode)
  - Current breeding population must be evaluated in drier environments ➔ establishment of progeny trials south of the current breeding zone
  - Introduction of non local provenances in the breeding population ➔ plus trees selected have to be caracterized in the current breeding zone



Pine wood nematode

- Opportunity to expand maritime pine plantations north of the current breeding zone (evaluation of the breeding populatio in new site types)



Results Summary

Areas that will have a similar climate in 2021-2050 to the current (1981-2010) climate at the selected location (44.3199° N, 0.7306° W).

Note: Climate Difference scores closest to zero indicate the best match.

Parameters	
Study region	Europe
Climate difference variables	Precipitation, Temperature and Diurnal Range
Selected site	44.3199° N, 0.7306° W
Local climate	1981-2010
Search climate	2021-2050
Months matched	Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct, Nov, Dec
Number of matches requested	5,000

Climate Matching Tool (https://climatematch.org.uk/)

➔The breeding population (Landes provenances + other genepools) has to be caracterized over a large range of environments to anticipate future changes

Genetic component of individual norm of reaction

- **Phenotypic plasticity** = phenotypic trajectory over an heterogeneous environmental landscape

- Breeding is focused on the selection of superior genotypes to establish seed orchards ➔ phenotypic plasticity must be studied at the **individual level** (≠ population level)

- Phenotypic plasticity implies multiple phenotypic states per genotype:
  - **site-related plasticity**: measurements repeated over sites on the same genotype but different ramets (clonal trial)
    - ✓ Independent measurements (simple statistical methodologies) but site selection can be tough (the environmental range covered by the sites is generally different from the range of interest for breeding)
  - **time-related plasticity:** measurements repeated over time for the same individual
    - ✓ Autocorrelation between measurements (specific statistical methodologies) but the environmental range explored is similar to the production conditions

- Phenotypic plasticity can be modelled statistically as a function of one or several explanatory environmental variables. The resulting function is then called **norm of reaction** (NoR). The choice of environmental variables (climatic variables, soil variables, indexes…) is a key issue. Ecophysiology knowledge is required to select the relevant variables.

- Genotypes from a given population might show different NoRs over the same environmental cline, resulting for instance in intercrossing curves. This is typically a genotype by environment interaction ➔ **Genotypic by environment interaction measures variation in NoR**.

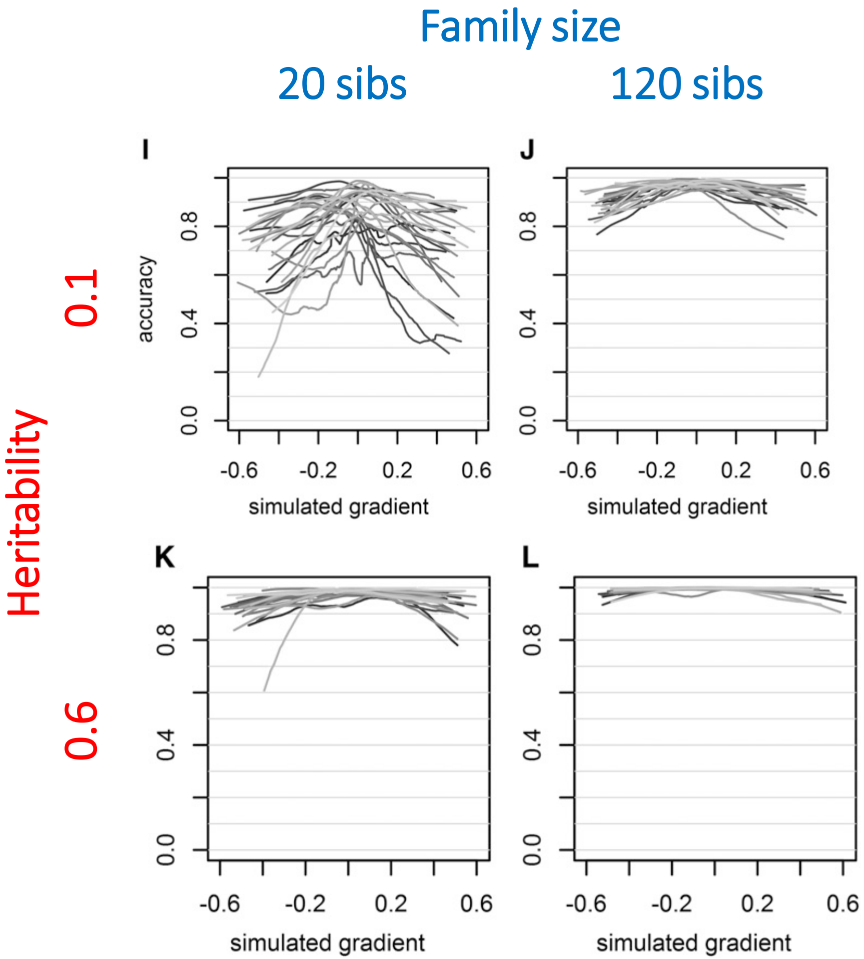
Maritime pine trials available to study GxE

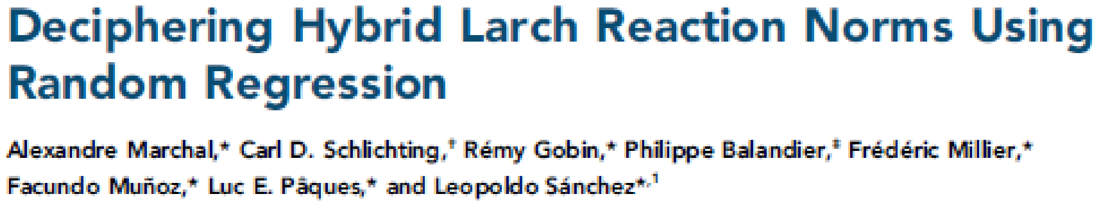
- Clonal trials established from micro-cuttings: large genetic variability (natural populations from various genepools) established in several contrasted sites

- In the French breeding context:
  - No clonal trials avalaible
  - Multisite progeny trials which explore the environmental variation of the current breeding zone (no site with extreme climates)
  - Only since recently:
    - exploration of new sites north or south of the breeding zone
    - exploration of larger genetic variability (trials with various provenances)

Genetic NoR based on familial data

- Estimation of site-related plasticity
- When no clonal data available, site-related plasticity can be estimated with familial data




Deciphering Hybrid Larch Reaction Norms Using Random Regression

Alexandre Marchal,* Carl D. Schlichting,† Rémy Gobin,* Philippe Balandier,‡ Frédéric Millier,*
Facundo Muñoz,* Luc E. Pâques,* and Leopoldo Sánchez*,1

Simulations (Metagene software):

- Full diallel design (10 parents) to generate full-sib families
- Only one sib per site (environmental gradient)
- 4 scenarios simulated which differ by the quantity (family size) and quality (h^2) of information
- Accuracy = correlation between the parents' predicted additive performance and their true additive performance

➔ Simulation results:

High accuracy with either high number of sibs or high heritability

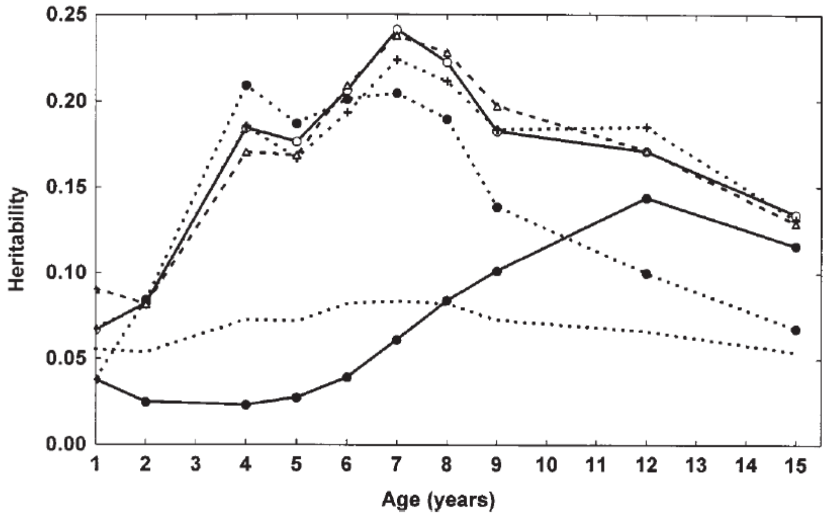
Genetic NoR based on longitudinal data

- **Estimation of time-related plasticity**
- **Longitudinal data**: refers to repeated observations of the same variables on the same individuals over a given period of time
- **Longitudinal data can be a way to explore a large range of environmental conditions**
- **Limitations**: repeated measures are linked by autocorrelation and ontogeny can impact the phenotype studied
- **Example**: time series for growth

Variance modelling of longitudinal height data from a *Pinus radiata* progeny test

Luis A. Apiolaza, Arthur R. Gilmour, and Dorian J. Garrick

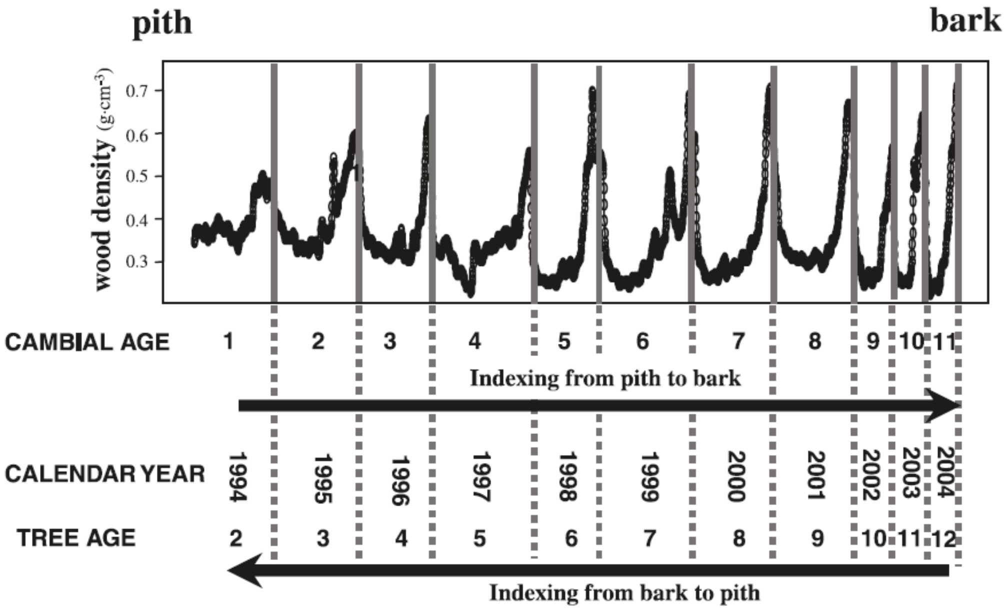
Can. J. For. Res. **30**: 645–654 (2000)



Longitudinal data based on wood

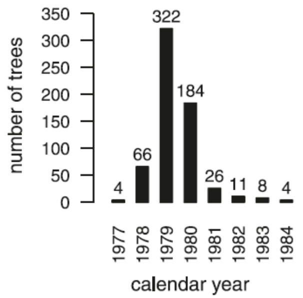
- Wood = record of tree anatomical reaction to environmental changes

During wood formation, the cambium reacts to environmental changes by adjusting the anatomy of the newly formed xylem cells. This continuous modification leaves a permanent anatomical trace in the wood that follows seasonal changes and weather events (➔ rings in a temperate climate)

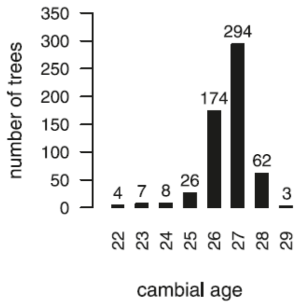


From Bouffier *et al.* 2008

(a) Cambial age = 2



(b) Calendar year = 2004



- In experimental trial (plantation): calendar year and tree age are confonding
- Indexing choice is important: cambial age indexing vs. calendar year indexing
- Analyses can focus on variability between rings or within rings

Genetic parameters from NoR

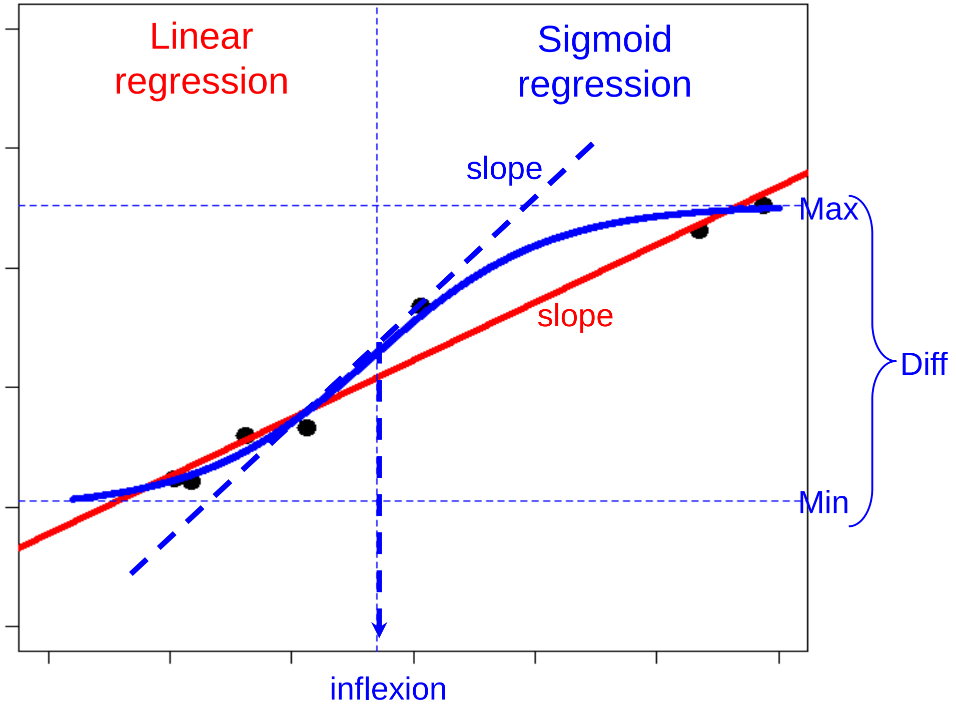
- Same quantitative framework to estimate the genetic parameters (genetic variance, genetic correlation, h²…):

➢ Pedigree / genomic data

➢ Estimation in several steps:

1. Fitting a NoR per genotype
2. Extraction of variables from each NoR (slope, inflexion point, max…)
3. Estimation of genetic parameters as for traditional traits

➢ Estimation in a single step:
mixed model (trajectory and genetic effect in the same model) ➔ random regression



Biological interpretation

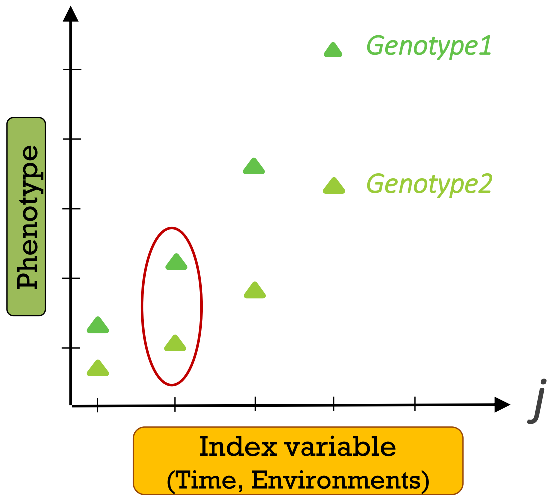
- **Slope** = speed of wood density increase in response to drought constraint
- **Inflexion** = susceptibility of tree to drought constraint
- **Min** = wood density when there is no water limitation
- **Max** = wood density when water limitation is maximal
- **Diff** = plasticity level for wood density

The concept of random regression

The random regression as an extension of the classical linear mixed model

Univariate model

Each level of the index variable = 1 trait

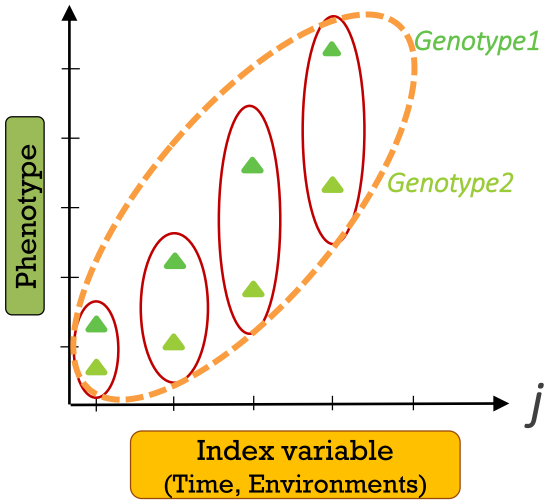


$$y_i = \mu + \beta_i + \alpha_i + \varepsilon_i$$

μ : mean
β_i : fixed effect
α_j : random effects

Multivariate model

Consideration of each level of the variable index as correlated traits

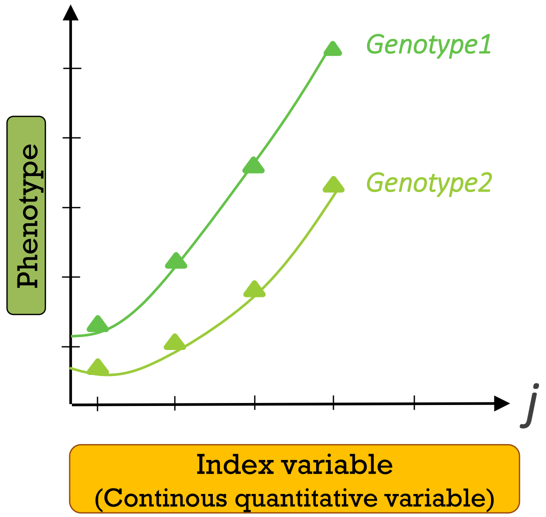


$$y_{ij} = \mu_j + \beta_{ij} + \alpha_{ij} + \varepsilon_{ij}$$

μ : mean
β_i : fixed effect
α_j : random effects

Random regression

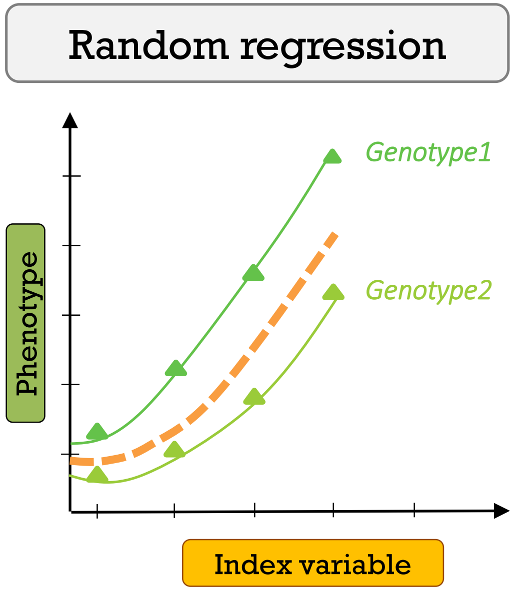
Adjustment of a trajectory according to a meaningful index



$$y_{ij} = M(t_j) + F_{ij} + A_i(t_j) + P_i(t_j) + \varepsilon_{ij}$$

$M(t_j)$: mean trajectory
F_{ij} : fixed effects
$A_i(t_j), P_i(t_j)$: random effects

Elements of random regression models



$$y_{ij} = \underbrace{M(t_j) + F_{ij}}_{\text{Fixed part}} + \underbrace{A_i(t_j) + P_i(t_j) + \varepsilon_{ij}}_{\text{Random part}}$$

$M(t_j)$: mean trajectory
→ of the whole population,
of a subgroup…

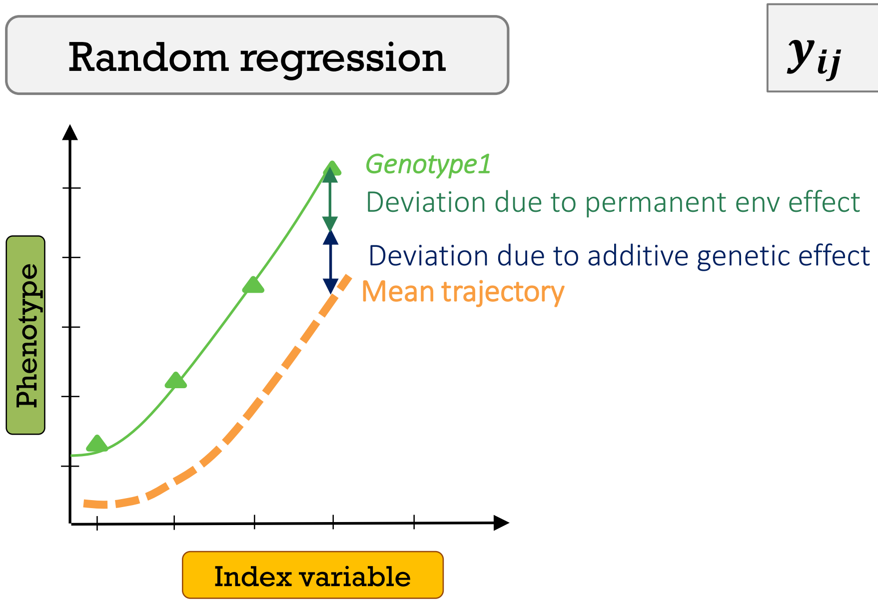
F_{ij} : fixed effects
→ Site, Treatment…

$A_i(t_j)$: genetic additive effect
→ Part of the deviation from mean trajectory due to genetic additive effects
→ Variance constrained by relationship matrix

$P_i(t_j)$: Permanent environmental effect
→ Part of the deviation from mean trajectory due to permanent environment effect

ε_{ij} : residuals (variance can be homogeneous or heterogeneous)

Elements of random regression models



$$y_{ij} = \underbrace{M(t_j) + F_{ij}}_{\text{Fixed part}} + \underbrace{A_i(t_j) + P_i(t_j) + \varepsilon_{ij}}_{\text{Random part}}$$

$M(t_j)$: mean trajectory
→ of the whole population,
of a subgroup…

F_{ij} : fixed effects
→ Site, Treatment…

$A_i(t_j)$: **genetic additive effect**
→ Part of the deviation from mean trajectory due to genetic additive effects
→ Variance constrained by relationship matrix

$P_i(t_j)$: Permanent environmental effect
→ Part of the deviation from mean trajectory due to permanent environment effect

ε_{ij} : residuals (variance can be homogeneous or heterogeneous)

Trajectory modeling in random regressions (1)

$$\boldsymbol{y_{ij}} = \boldsymbol{M(t_j)} + \boldsymbol{F_{ij}} + \boldsymbol{A_i(t_j)} + \boldsymbol{P_i(t_j)} + \boldsymbol{\varepsilon_{ij}}$$

$$\boldsymbol{y_{ij}} = \sum_{r=1}^{k_m} m_{ir}\,\varphi_r(t_j) + F_{ij} + \sum_{r=1}^{k_\alpha} \alpha_{ir}\,\varphi_r(t_j) + \sum_{r=1}^{k_\gamma} \gamma_{ir}\,\varphi_r(t_j) + \boldsymbol{\varepsilon_{ij}}$$

3 elements in each sum :

$\varphi_r(t_j), \varphi_r(t_j), \varphi_r(t_j)$

rth basis funtion evaluated for t_j

k_m, k_α, k_γ

Polynomial order for modeling

$m_{ir}, \alpha_{ir}, \gamma_{ir}$

regression coefficients for mean trajectory, additive genetic, permanent environment effects

Trajectory modeling in random regressions (2)

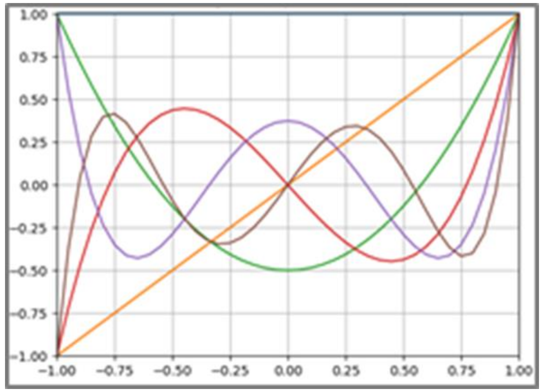
*For additive genetic effect (**α**) with a 2-order plonymial modeling (**k_α**=3) :*

$$\sum_{r=1}^{k_\alpha} \alpha_{ir}\, \varphi_r(t_j) \;=\; \alpha_{i1}\varphi_1(tj) + \alpha_{i2}\varphi_2(t_j) + \alpha_{i3}\varphi_3(t_j)$$

$\boxed{\varphi_1(t_j)}$ *Basis function : Legendre polynomials*

Results for each individual i

*With **k_α**=3 and j ϵ [1:20]*



2-order Legendre polynomial coefficients			
j=1	0,71	-1,22	1,58
j=2	0,71	-1,10	1,11
j=3	0,71	-0,97	0,69
…	…	…	…
j=17	0,71	0,84	0,32
j=18	0,71	0,97	0,69
j=19	0,71	1,10	1,11
j=20	0,71	1,22	1,58

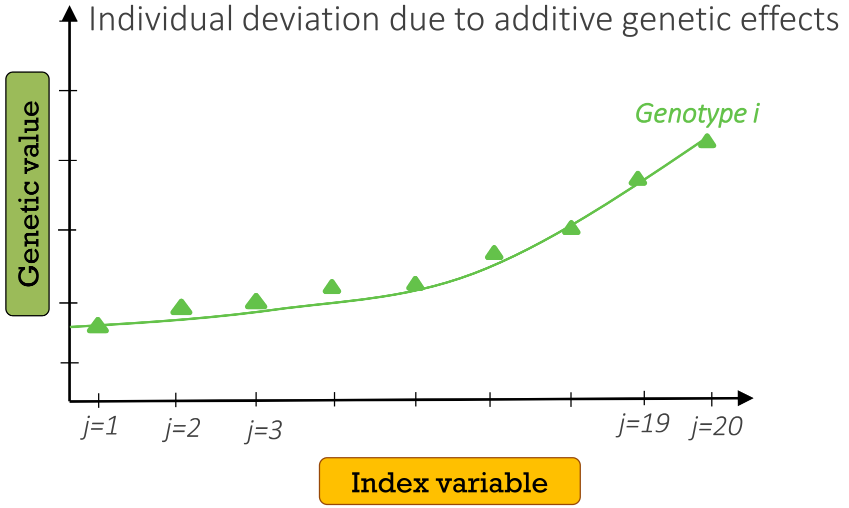
$$\begin{pmatrix}\alpha_{i1}\\ \alpha_{i2}\\ \alpha_{i3}\end{pmatrix}$$

3 regressions coefficents for the additive genetic effects

Results of a random regression

For each individual i, for the additive genetic effect :

$$\begin{bmatrix} \alpha_{i1} \\ \alpha_{i2} \\ \alpha_{i3} \end{bmatrix}$$



Possible to extract one breeding value for each individual at each value of the index variable

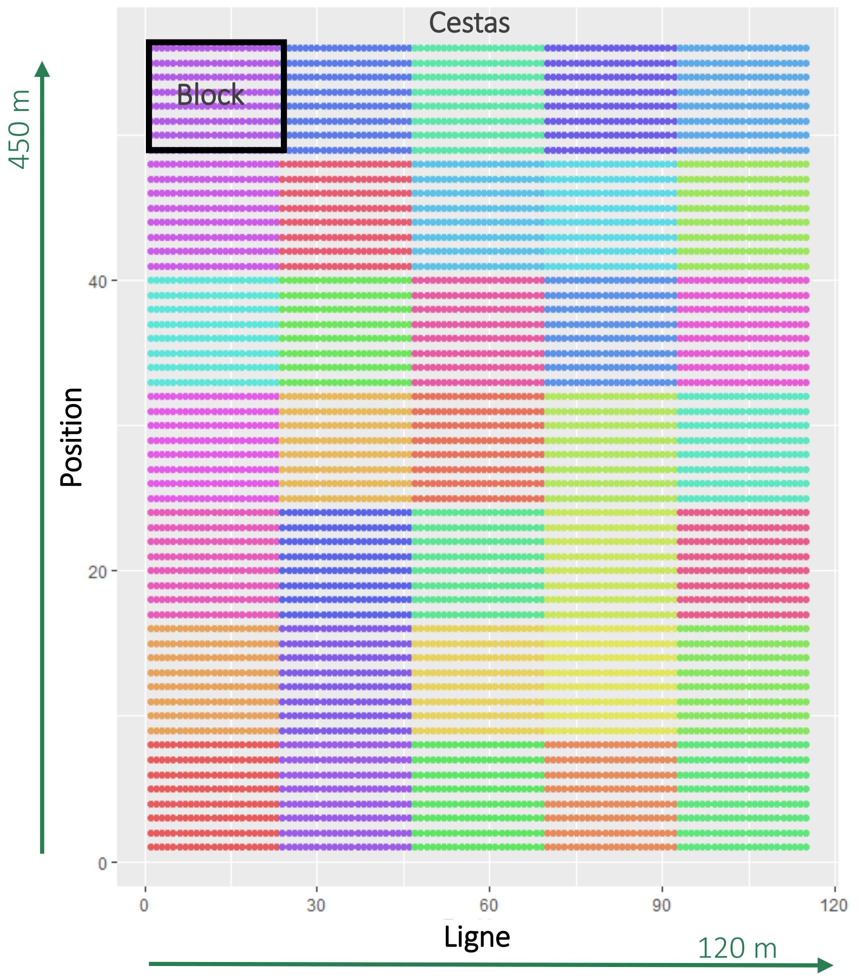
Breeding value	
j=1	a1
j=2	a2
j=3	a3
…	…
j=17	a17
j=18	a18
j=19	a19
j=20	a20

- Same procedure for mean trajectory, permanent environmental effects, genetic variance

Norm of reaction for Maritime pine

Experimental design and measurments

Experimental design for Maritime pine (6300 trees)

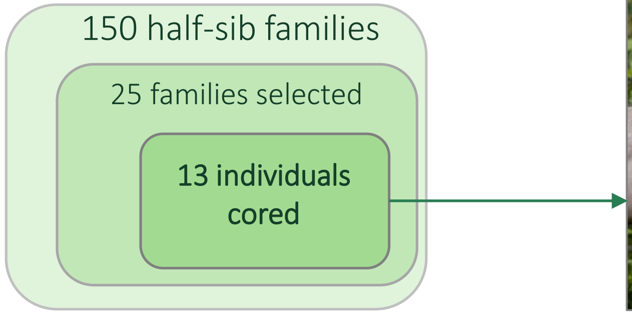


Experimental design

- 2 locations : Cestas (humid) & Escource (dry)
- Installation in **1996** : **26** years old trees
- **150** half-sib families with **35** individuals/family → **6300** trees per site
- **Complete** block design (1 individual of each family per block)

Phenotypic measurments :

- Classical growth measures at different ages (height, circumference, straightness)
- Core sampling of **325** trees per site :


150 half-sib families

25 families selected

13 individuals cored

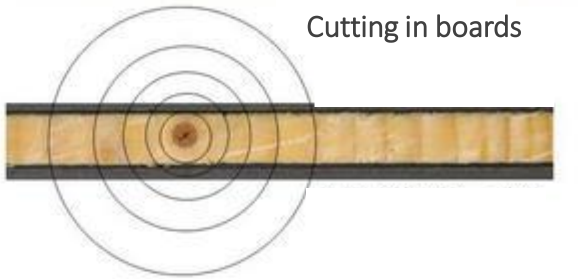

© Thibaut VERGOZ/IMBE/CNRS Photothèque

Environmental measurments :

- Conventional annual climate measurements (temperature, rainfall…)

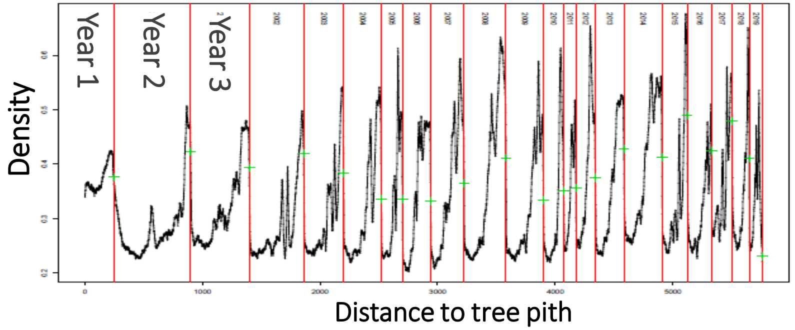
Experimental design and measurments

Core sampling

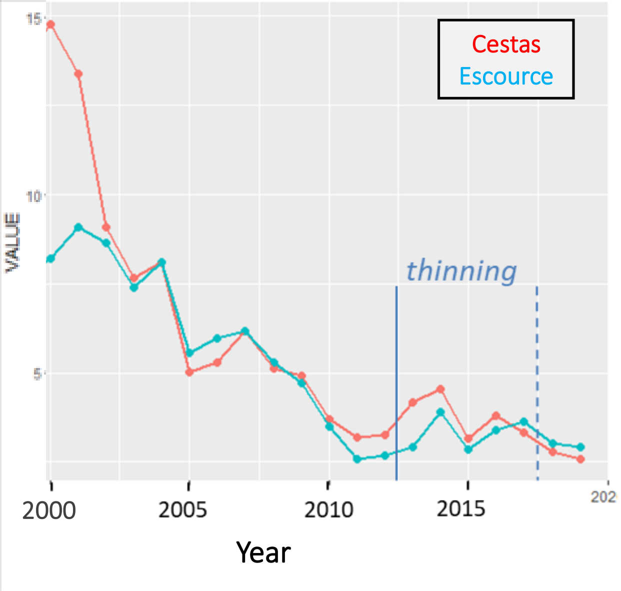


Cutting in boards

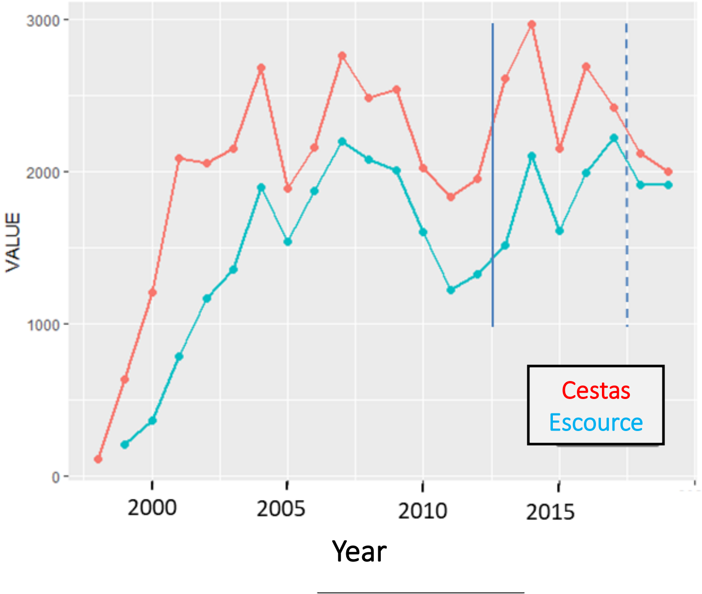
Density profile after X-ray radiography



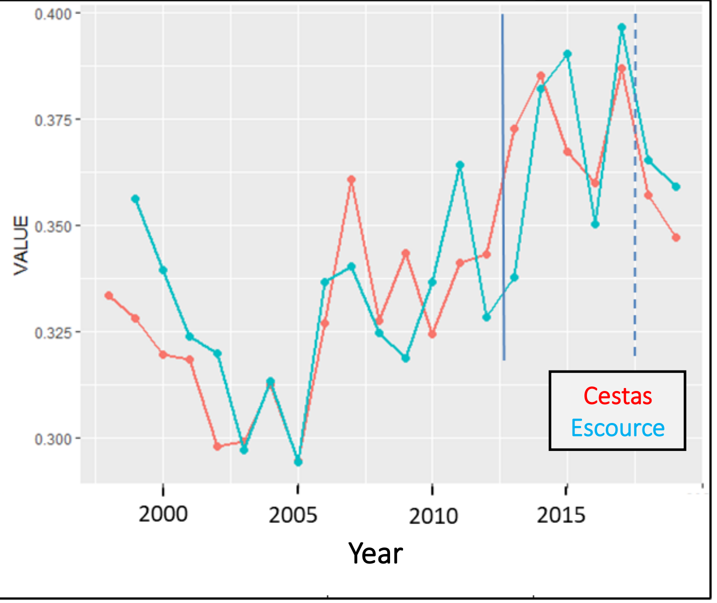
Ring width



Ring surface



Ring mean density



Norm of reaction for maritime pine : trait/year

- We want to model **one trajectory for each tree of each site** : *650 trajectories*
- First modelisation are done according to the **year** : *index variable = year & j є [1999:2019]*

For individual i :

$$\boldsymbol{y_{ij}} = \sum_{r=1}^{k_m} m_{ir}\,\varphi_r(t_j) + F_{ij} + \sum_{r=1}^{k_\alpha} \alpha_{ir}\,\varphi_r(t_j) + \sum_{r=1}^{k_\gamma} \gamma_{ir}\,\varphi_r(t_j) + \boldsymbol{\varepsilon_{ij}}$$

Inter-ring phenotypic trait :

ring width for the
indiviudal i,
at the the year j

Mean trajectory as fixed effect:

- The mean trajectory of the ring width acrros the year for the whole trees
  - One mean trajectory per site

Other fixed effect

Polymix effect (2 different polymix used to generate the HS families)

Additive genetic random effect

- Based on one individual i
- Effect variance constraine by a kinship matrix (constructed with pedigree information)

Permanent environmental random effects

Overall effects due to the individual i but not from additive genetic origin

Residuals

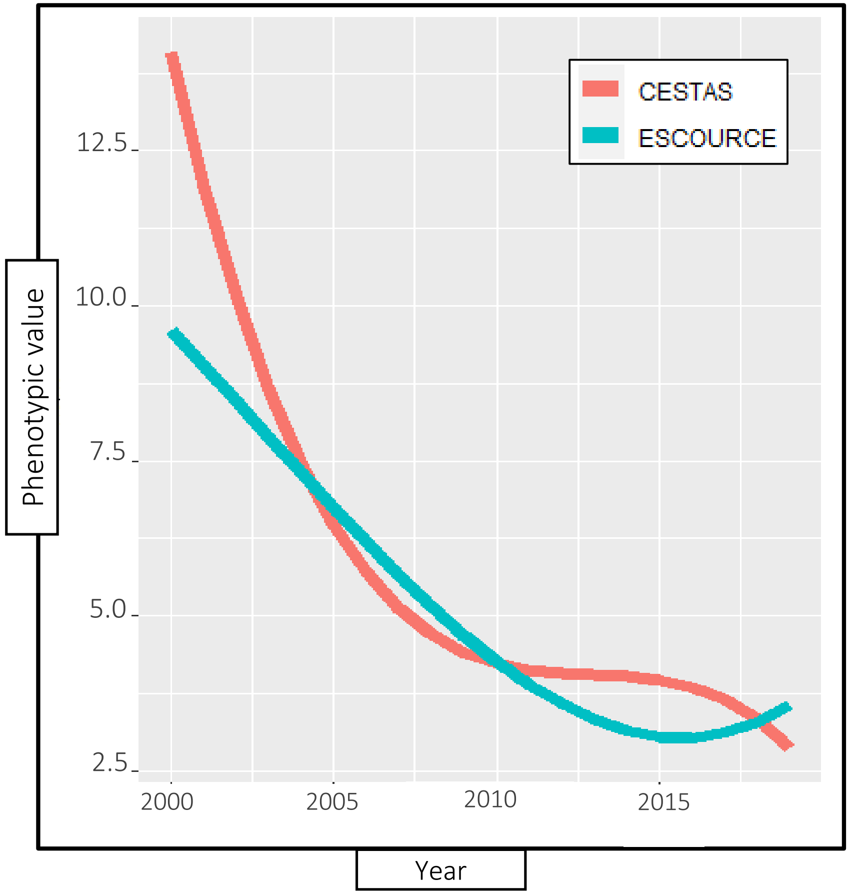
a heterogeneous residual variance for each year

- Mean trajectories, additive genetic effect and permanent environmental effects are modeled by **2-order Legendre polynomials**

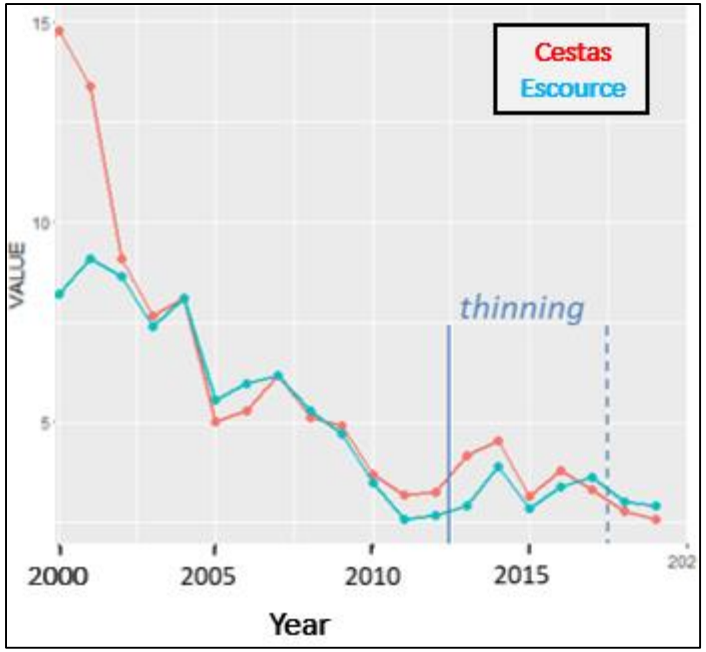
Norm of reaction for maritime pine : trait/year

<u>Fixed part of the model :</u>

Mean trajectories according to each site



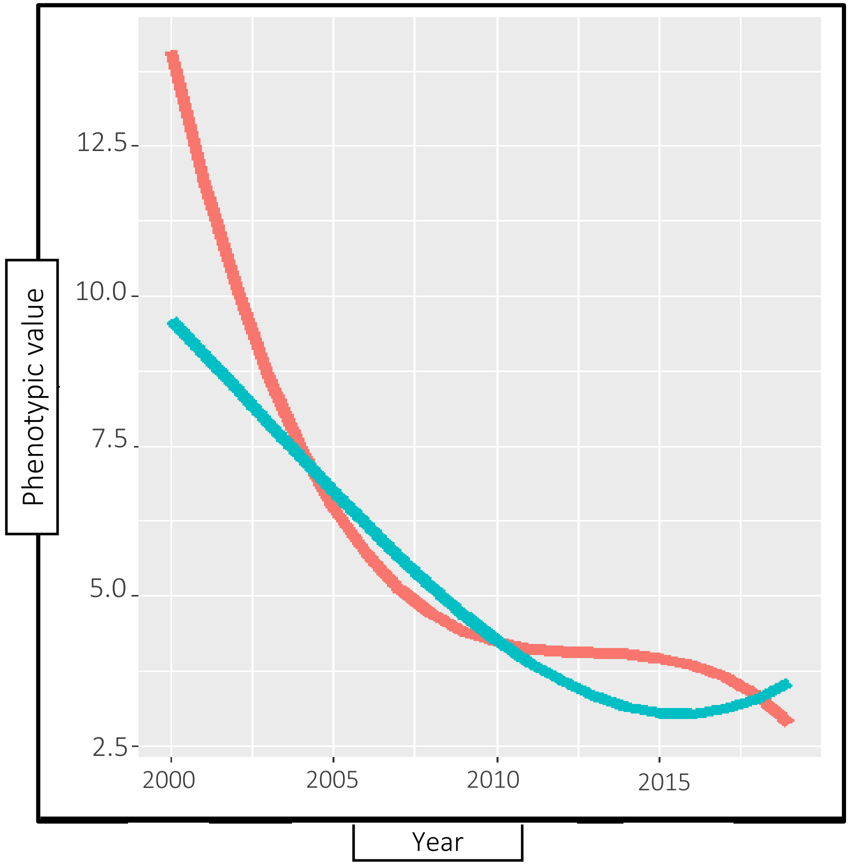
<u>Mean raw data</u>



Norm of reaction for Maritime pine – Papin, Bouffier, Sanchez-Rodriguez

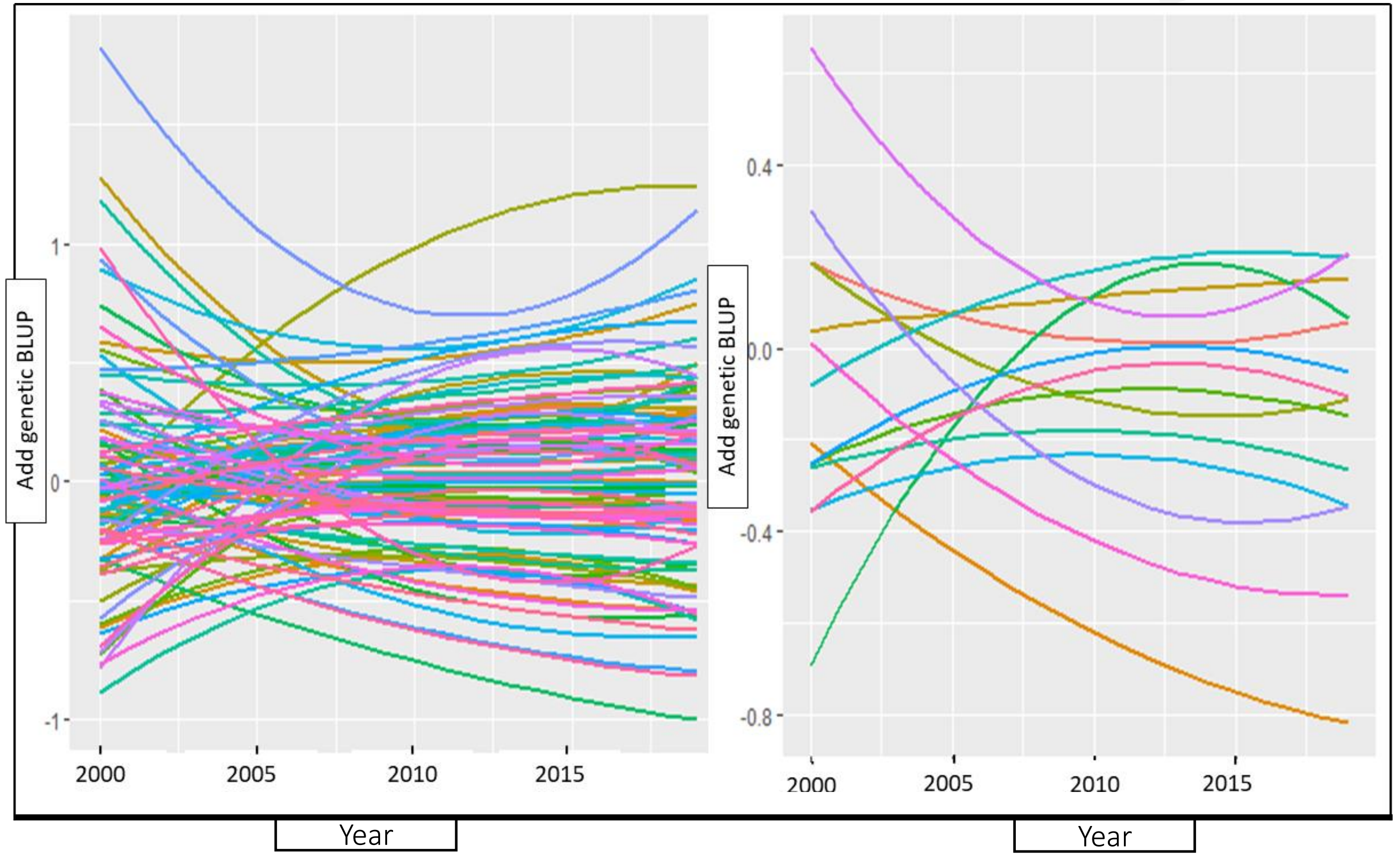
Norm of reaction for maritime pine : trait/year

Fixed part of the model :
Mean trajectories according to each group



Random part of the model :
individual trajectories for additive genetic effect

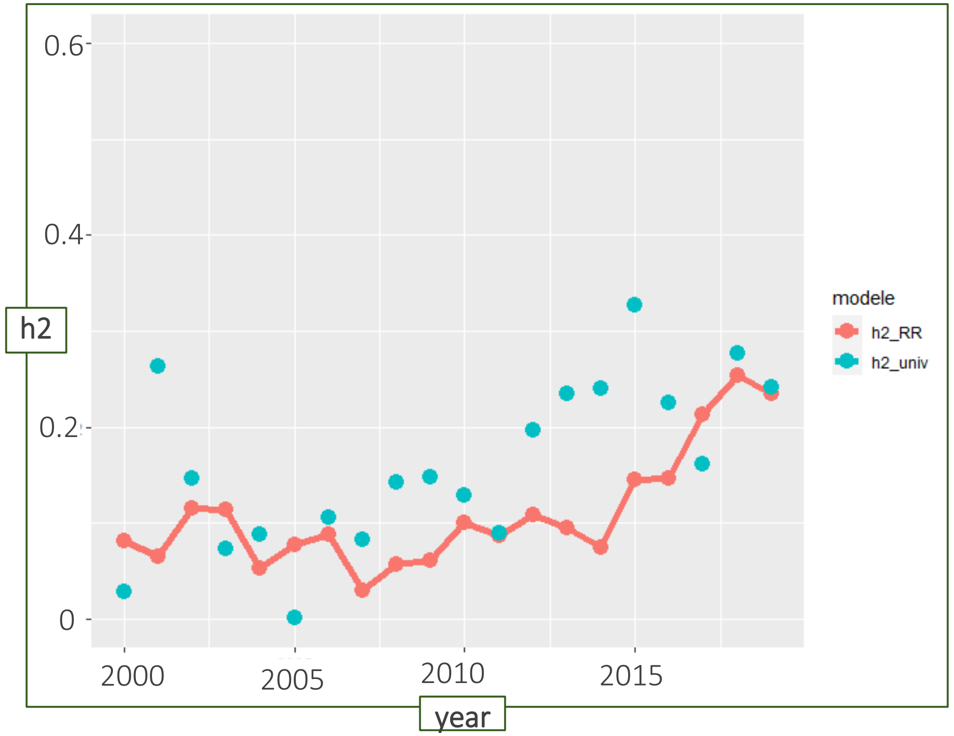
Random part of the model :
Family trajectories for additive genetic effect



Norm of reaction for maritime pine : trait/year

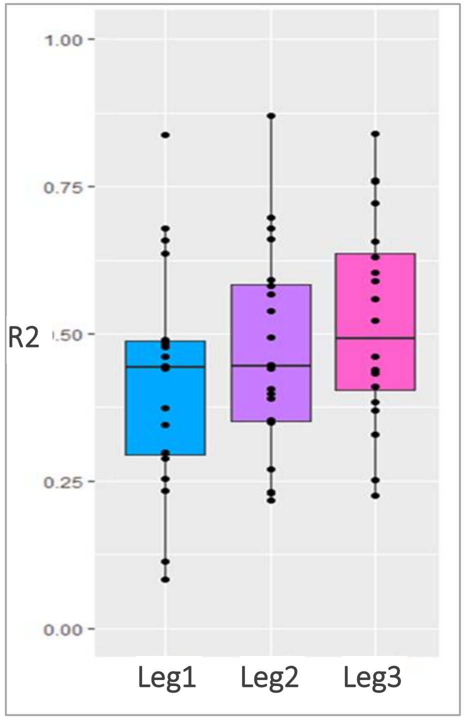
Evolution of heritability across the year

(estimated with RR model and independent univariate model)



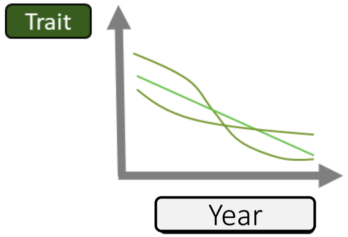
Model quality : R² (coefficient of determination)

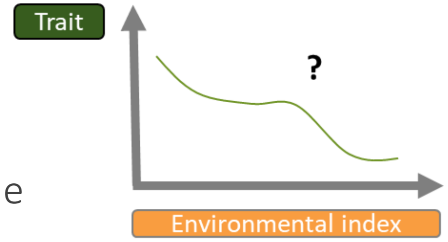
$$R2 = \frac{Variance\ explained\ by\ the\ model}{Total\ variance}$$



Norm of reaction for maritime pine : trait / env index

- In previous analysis, index variable = year





- But in a context of breeding and rapid climatic change, our objective is <u>to predict the response of trees in a specific environment</u> : new index variable = environmental variable

<u>Method :</u>

Martonne aridity index, for a month :

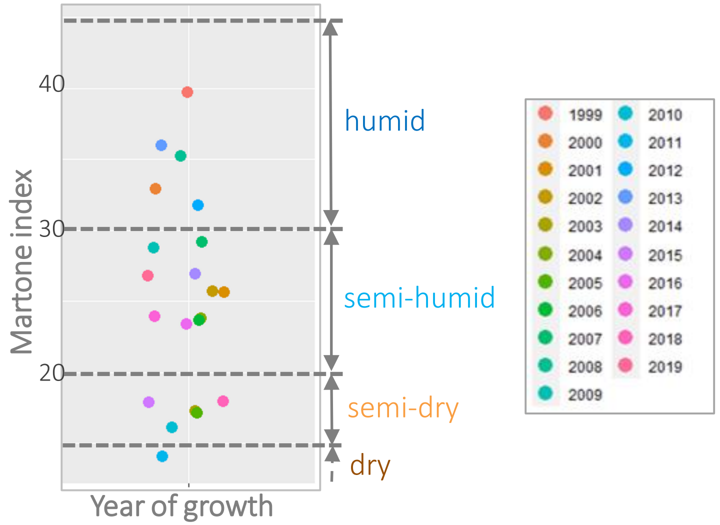
$$I = \frac{12p}{t + 10}$$

t : average monthly temperature
p : total monthly rainfall

available through INRAE weather databases

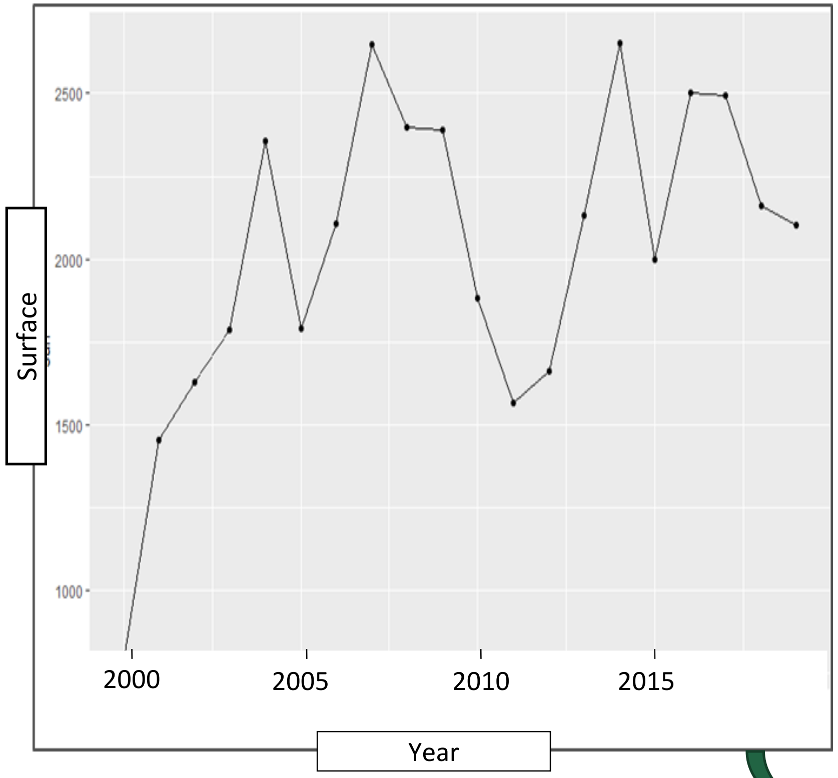
For a year, mean of monthly Martonne index on growth season (April – September)

Characterization of each year

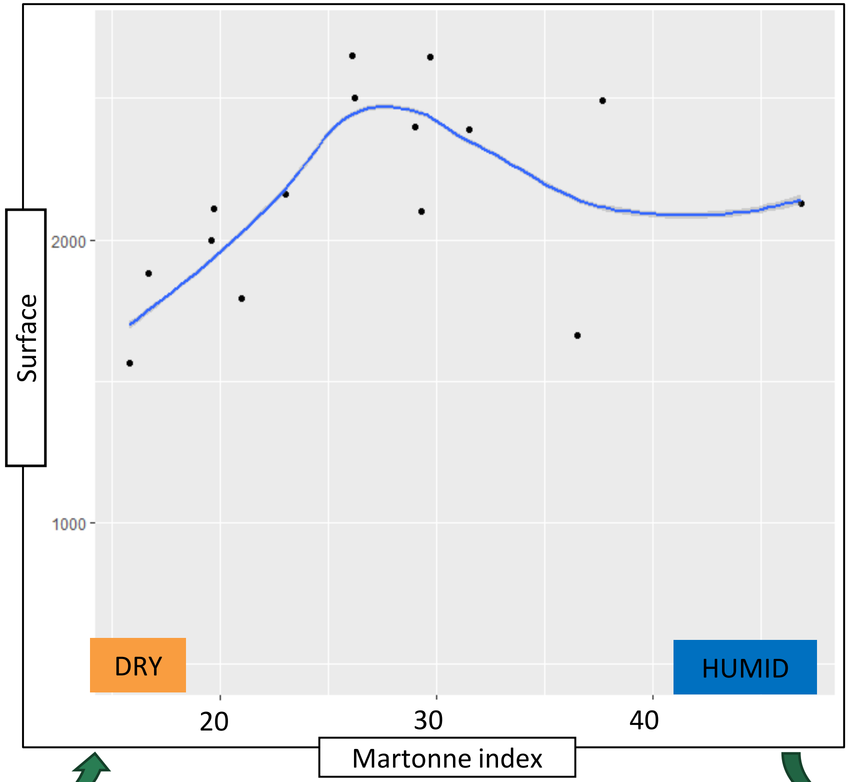


Norm of reaction for maritime pine : trait / env index

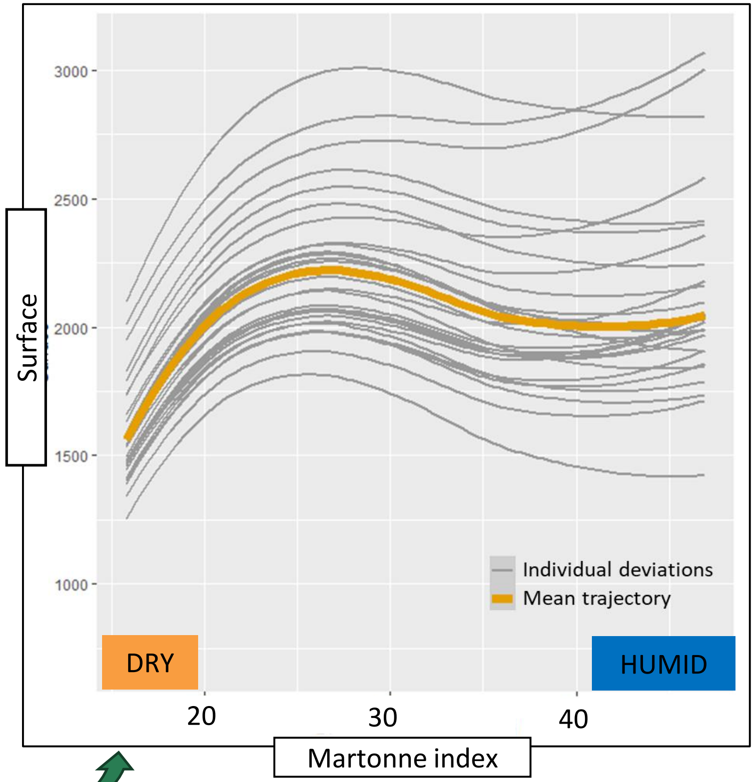
Ring surface across years



Ring surface across aridity index



Ring surface modelisation by RR

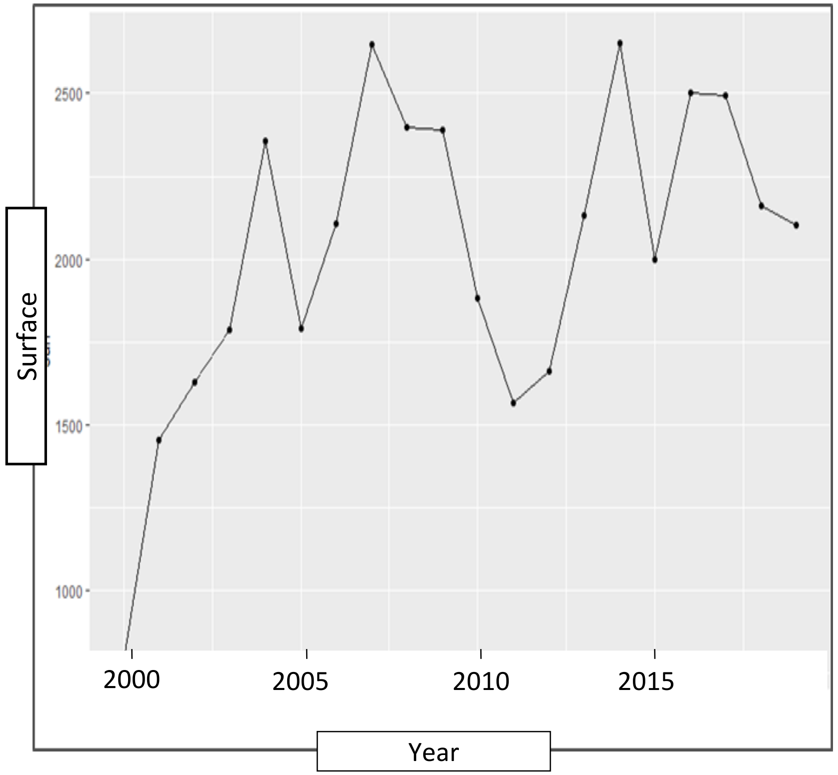


re-organization

Modeling

Norm of reaction for maritime pine : trait / env index

Ring surface across years



Raw trajectories may contain hidden effects and not only responses to environmental conditions of the year

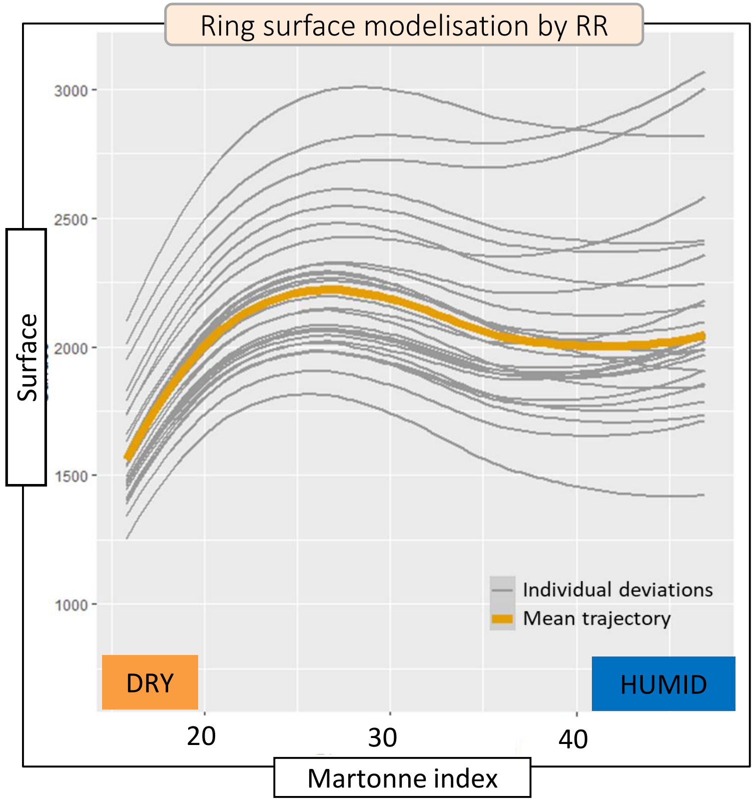
For example the effect of age, of a disease attack…

In our case, we have a constraint of **<u>the age effect</u>** (all trees were planted in the same year)

Possible improvements :

- Focus on 2005-2019
- 3-step approach :
    - Perform a 1st RR on year to estimate the mean trajectory
    - Correcting phenotypes with this mean effect
    - Perform a 2nd RR env_index with corrected data

Norm of reaction for maritime pine : prospects



- Environmental index :
  - Improve year characterization (soil humidity, water balance, impact of previous years, others indexes…)
  - Compare the relevance of the different environmental variables in the RR models
- Genetic data :
  - Replace pedigree information by genotypic data
- Phenotypic trait
  - Model other inter-ring traits (mean density, percentage of final wood…)
  - Consider intra-ring traits (conversion from a distance to time, model through a continous environmental variable

Main objectives :

- ➢ Identify environmentally sensitive traits & identify environmental variables that induce differential responses in trees
- ➢ Characterize the diversity of tree reactions & identify families of interest for selection