# Integrative factorial methods to explore the relationships between genotypes, phenotypes and climate in Holstein cows

Denis Laloë, Filippo Biscarini, Salvatore Mastrangelo, Gabriele Senczuk, Christian Persichilli, Giuseppe Conte, Raffaella Finocchiano, Jan Thijs van Kaam, Roberta Ciampolini, Martino Cassandro

25th Congress
ASPA2023
Monopoli (Bari, Italy), June 13-16, 2023

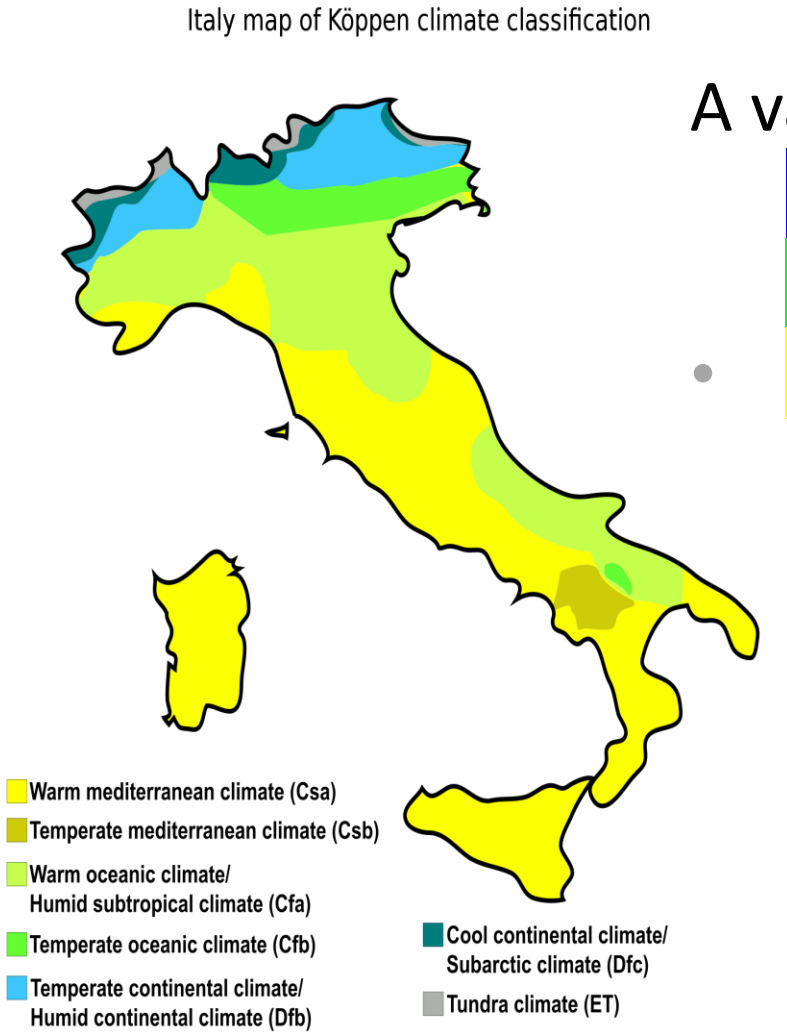
# *Integrative factorial methods to explore the relationships between genotypes, phenotypes and climate in Holstein cows*

## *A first step: Synthesizing lactation curves*

D Laloë, F Biscarini, S Mastrangelo, G Senczuk, C Persichilli, G Conte, R Finocchiano, JT van Kaam, R Ciampolini, M Cassandro

INRAE
Consiglio Nazionale delle Ricerche
IBBA
Istituto di Biologia e Biotecnologia Agraria

# Introduction
# The map of italian climates



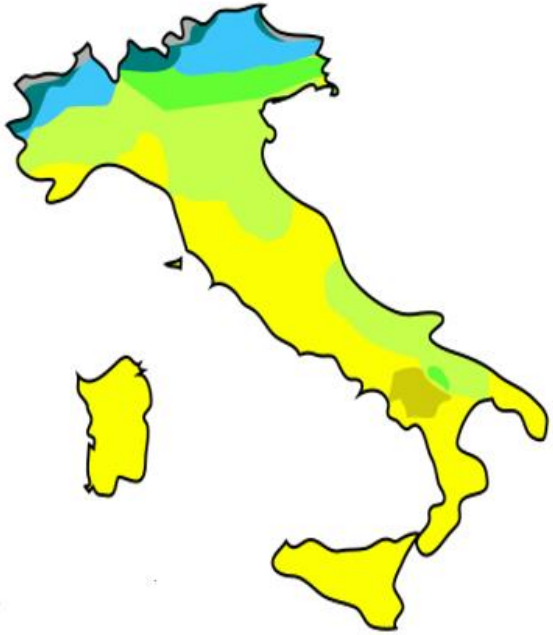

Commons.wikimedia.org

A variety of climates

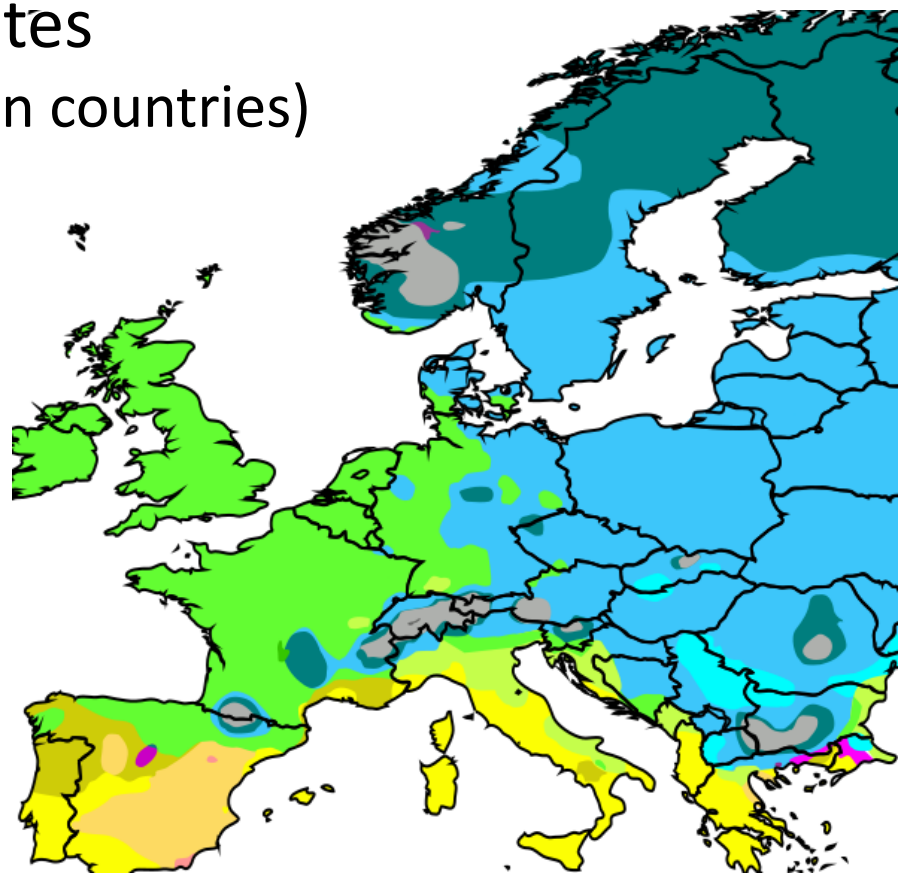
- Continental
- Oceanic / Temperate
- Mediterranean

# Introduction
## The map of italian climates

### A variety of climates
(compared to other european countries)



*Commons.wikimedia.org*



*en.wikimedia.org*

# Introduction
# Accounting for environment

The fine-scale (day) classical approach:
Reaction norms model on test-day records : test-day phenotypes linked to day-recorded meteo variables

*Bernabucci, U., Biffani, S., Buggiotti, L., Vitali, A., Lacetera, N., & Nardone, A. (2014). The effects of heat stress in Italian Holstein dairy cattle. Journal of dairy science, 97(1), 471-486.*

*Landi, V., Maggiolino, A., Cecchinato, A., Mota, L. F. M., Bernabucci, U., Rossoni, A., & De Palo, P. (2023). Genotype by environment interaction due to heat stress in Brown Swiss cattle. Journal of Dairy Science, 106(3), 1889-1909.*

*Vinet, A., Mattalia, S., Vallée, R., Bertrand, C., Cuyabano, B. C., & Boichard, D. (2023). Estimation of genotype by temperature-humidity index interactions on milk production and udder health traits in Montbeliarde cows. Genetics Selection Evolution, 55(1), 1-17.*

An other larger-scale approach:
Effect of general environment (through geography)

1. Functional PCA (fpca) on lactation curves (milk yield, fat, protein) -> a few synthetic variables
2. Links of fPCA results with geography / climate

# Methods: Functional Principal Components Analysis

**Data recorded at different discrete times-> Curve**

**Functional pca:**

- **Irregular measurement times, irregular number of measurements**
- **Extract the common temporal characteristics of a set of curves**
- **Essential modes of temporal variation: Principal Functions**
  - **% of variance**
  - **interpretation**

An extension of classical PCA (Macciotta et al (1015)

| | PCA | FPCA |
|---|---|---|
| Dimension | $P< \infty$ | $\infty$ |
| Mean | E(X) : scalar | E(X(t)) : function |
| Covariance | $Cov(x_p, x_q) = \Sigma$ | Cov (X(s),X(t) : surface |
| Eigenvectors / Eigenfunctions | $v_1$ , $v_2$ ,… | $f_1(t)$ , $f_2(t)$ ,… |


*Macciotta et al(2015). Genome-wide association analysis in Italian Simmental cows for lactation curve traits using a low-density (7K) SNP panel. *Journal of Dairy Science*, *98*(11), 8175-8185.

Yao F., Müller H.G., and Wang J.-L. (2005). Journal of the American Statistical Association 100 (470): 577–90. https://doi.org/10.1198/016214504000001745

Zhou Y et al (2022). _fdapace: Functional Data Analysis and Empirical Dynamics, R package

Arnal, M., Robert-Granié, C., & Larroque, H. (2018). Diversity of dairy goat lactation curves in France. *Journal of dairy science*, *101*(12), 11040-11051.

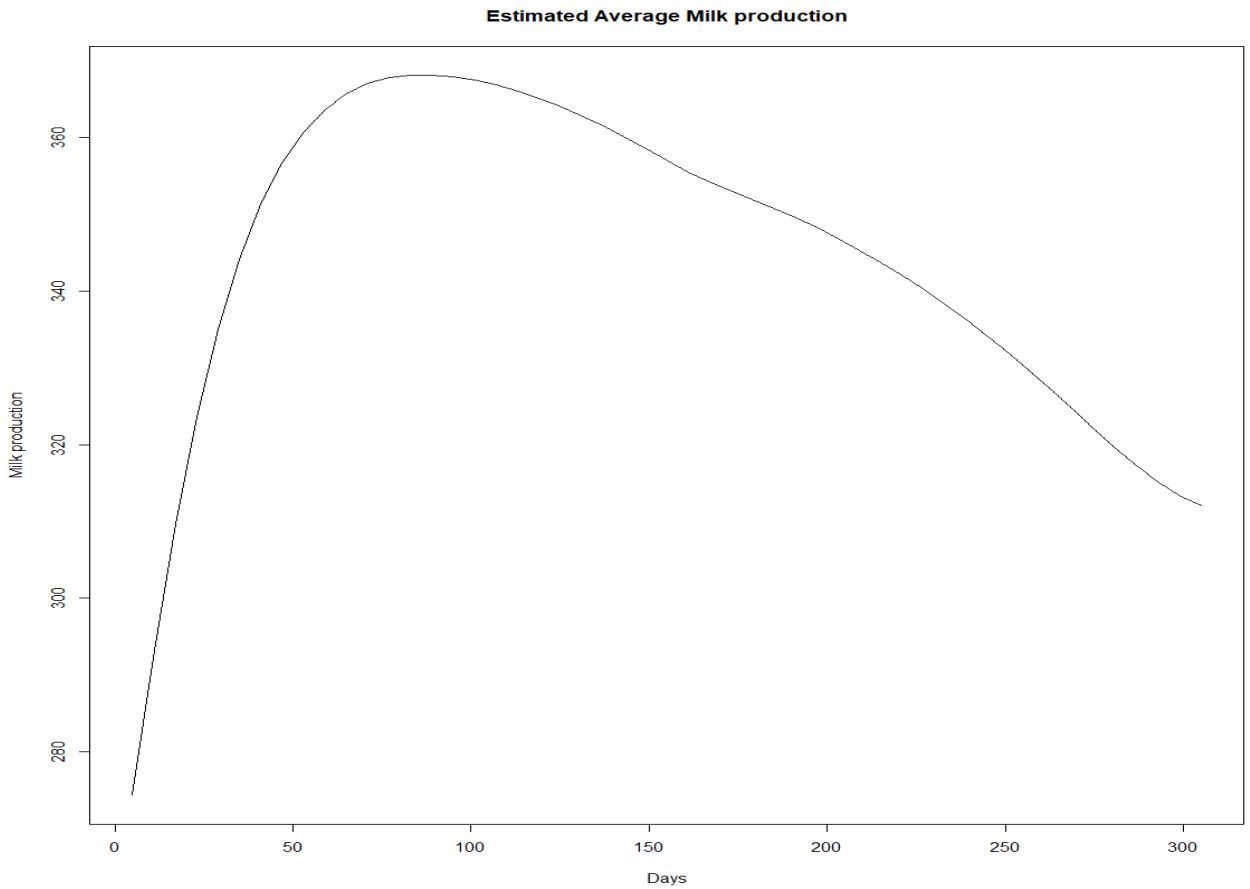
# Data (ANAFI): Milk records

- Editing : First lactations, in the same herd , with > 4 records.

| | Number |
|---|---|
| Records | 120245 |
| Lactations / Cows | 15545 |
| Herds | 881 |

| Number of records / lactation | Number |
|---|---|
| 5 | 898 |
| 6 | 1788 |
| 7 | 4162 |
| 8 | 4140 |
| 9 | 2845 |
| 10 | 1664 |
| 11 | 48 |

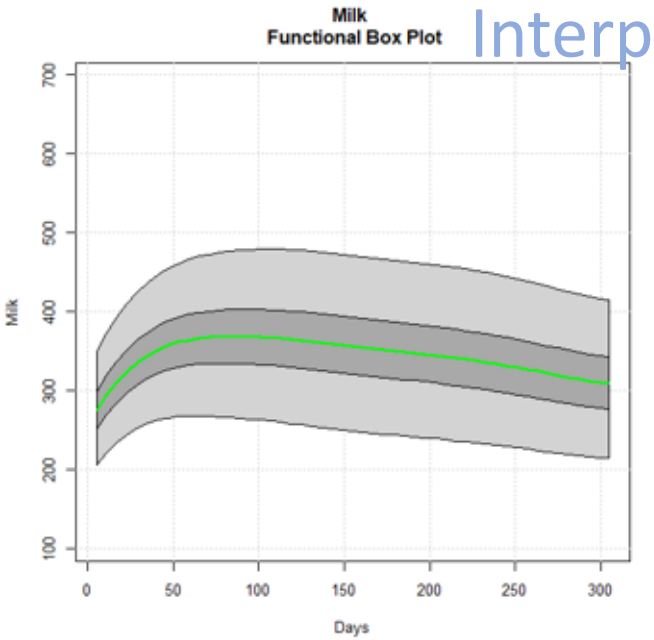
# Functional PCA Milk yield
# Average curve

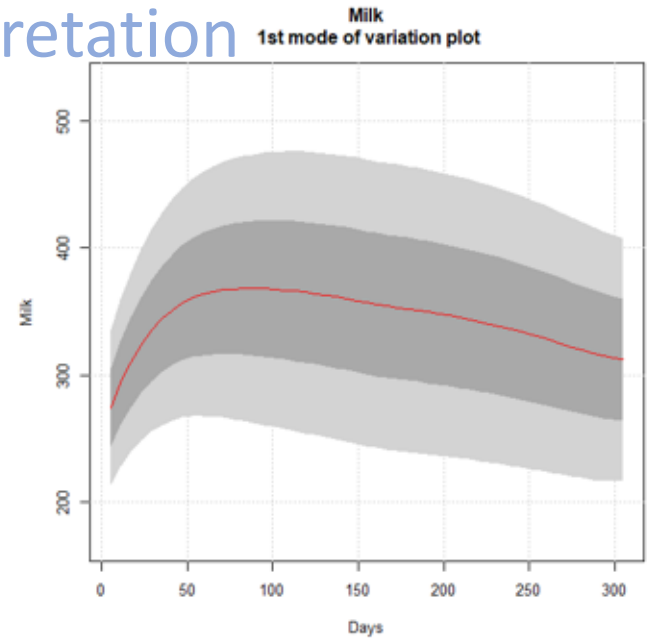


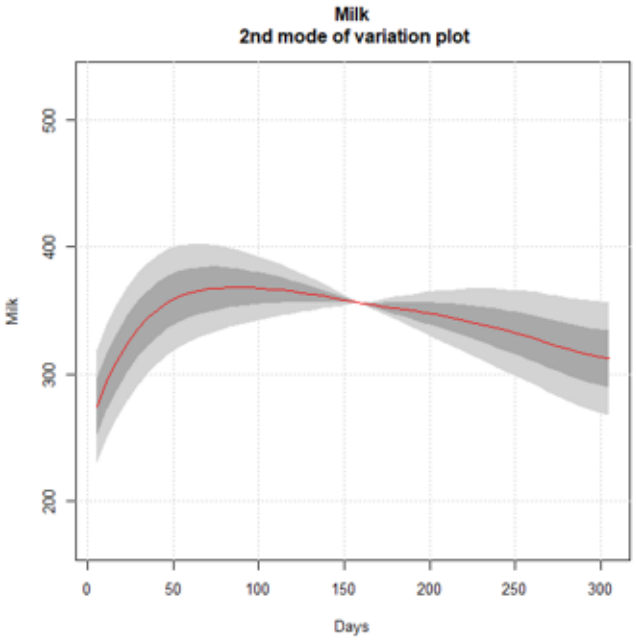
# Functional PCA Milk yield Interpretation

$$\bar{x}(t) \pm Q_\alpha \sqrt{d_i}\xi_i(t)$$

| | Milk Yield |
|---|---|
| 1 | 89,5 |
| 2 | 8,0 |
| Total(1+2) | 97.5 |




Variations around mean curves

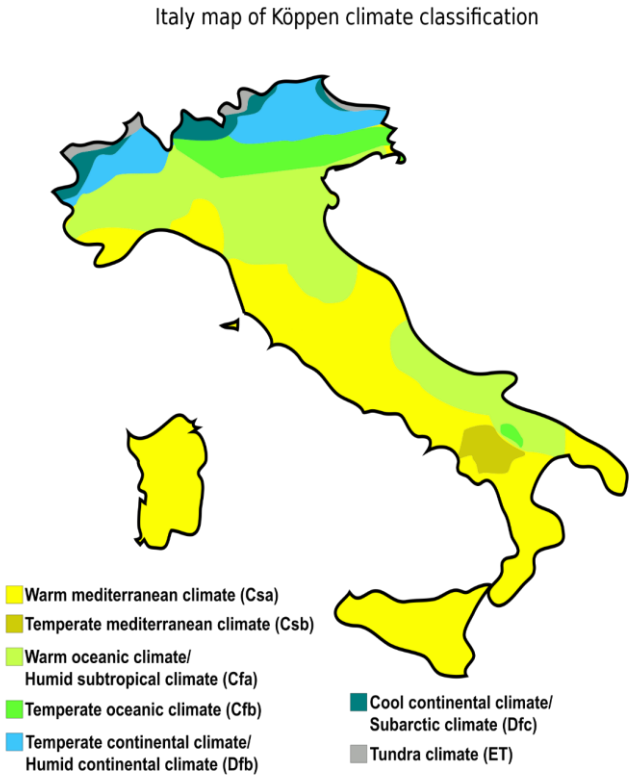
**K=1 : Level of production *(Size)***
**K=2: Persistency *(Shape)***

# Functional PCA applied to Milk yield, Protein (%), Fat(%)

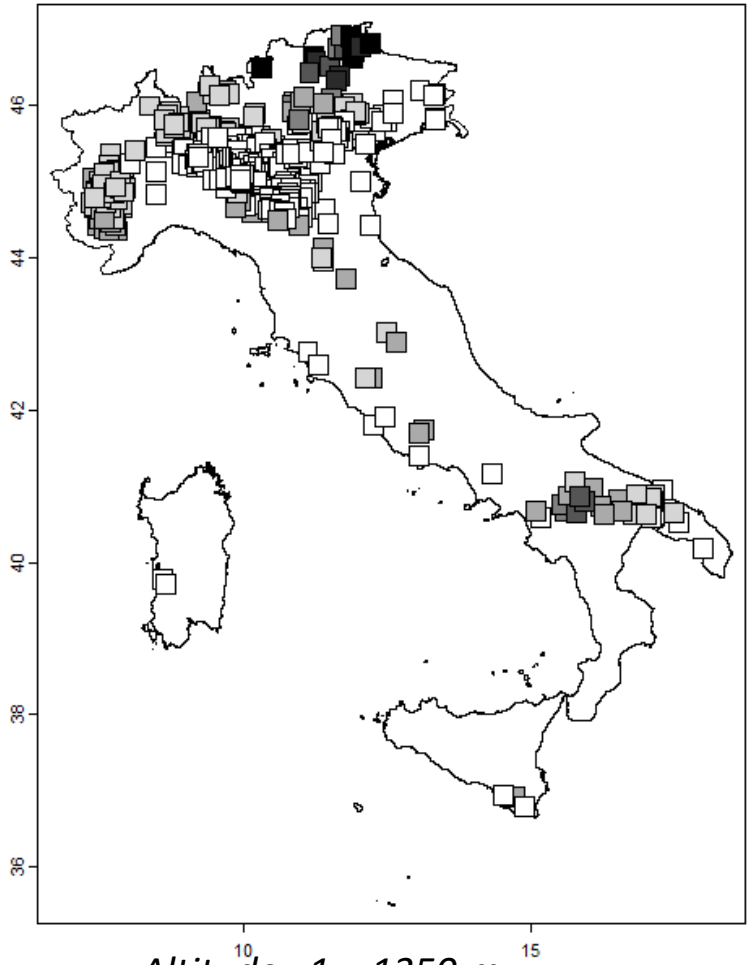
% of variance explained by the first 2 functional PCs

| | Yield | Protein % | Fat % |
|---|---|---|---|
| 1 | 89,5 | 85,2 | 88,8 |
| 2 | 8,0 | 11,7 | 7,3 |
| Total (1+2) | 97.5 | 96.9 | 96.1 |

# Data (ANAFI)
# Geography




Location of herds

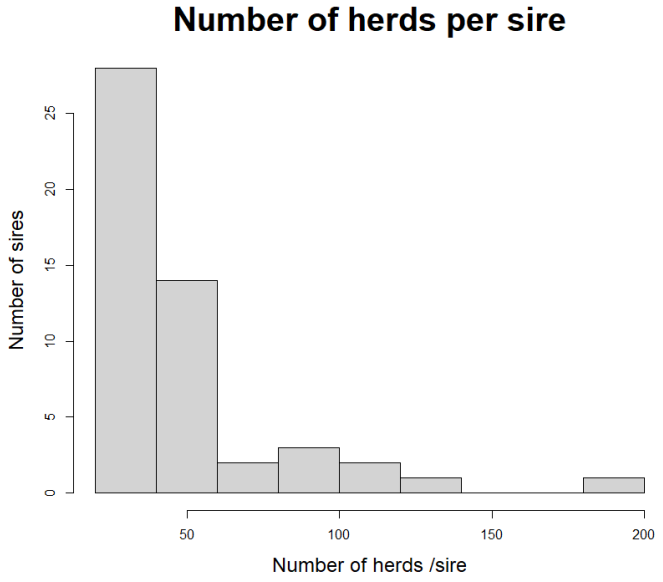


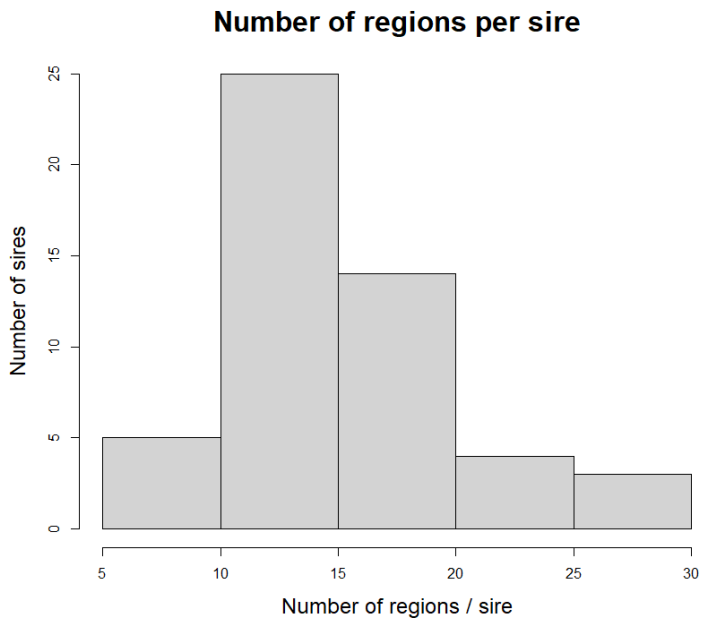

*Altitude : 1 – 1350 m*
*The darker, the higher*

# Adaptation to different environments
# Sire * geography interaction

2. Selection of « ubiquist » sires
   Sires that serve in > 25 herds

51 sires;
659 herds
42 provinces
6368 lactations



# A first attempt for the characterization of an interaction Geography * Genotype

A mixed linear sire model (R package « nlme »)
Fixed effets : Year calving, Month calving, latitude, longitude, altitude
Random effects: Sire, interactions Sire*latitude, Sire*longitude, Sire*altitude

* ML test Presence vs Absence interactions -> Significant (except when NS indicated)
* From 6 % to 25 % of the sire variance

| | Year | Month | Latit | Long | Altit | $\sigma^2_{sire}$ | $\sigma^2$ sire*lat | $\sigma^2$ sire*long | $\sigma^2$ sire*alt |
|---|---|---|---|---|---|---|---|---|---|
| Milk | NS | * | NS | * | NS | 0.056 | 0.014 | 0.006 | 0.014 |
| Prot% | * | * | * | * | NS | 0.108 | 0.011 | 0.007 | *0.003 (NS)* |
| Fat% | NS | * | * | NS | * | 0.068 | 0.006 | 0.006 | *0.004 (NS)* |

Pinheiro J, Bates D, R Core Team (2023). _nlme: Linear and Nonlinear Mixed Effects Models_. R package version 3.1-162, <https://CRAN.R-project.org/package=nlme>.

# Conclusion

For almost each trait*geographical coordinate, the interaction is significant, exhibiting a genotype*environment interaction

A (over?) simplified modelling
Environment (climate, breeding system, herd) synthetized by geography
Should be completed with more focused genetic studies (genetic / genomic studies )

Nevertheless,
Functional PCA allows us to synthesize a lactation curve into a few components;
- Can be seen as an extension of PCA
- Helps to relate milk yield to global climate parameters;
- could be used jointly with climate projections or scenarii (# genomic vulnerability)

A complementarity with reaction –norm approach


Thanks for Your Attention!!