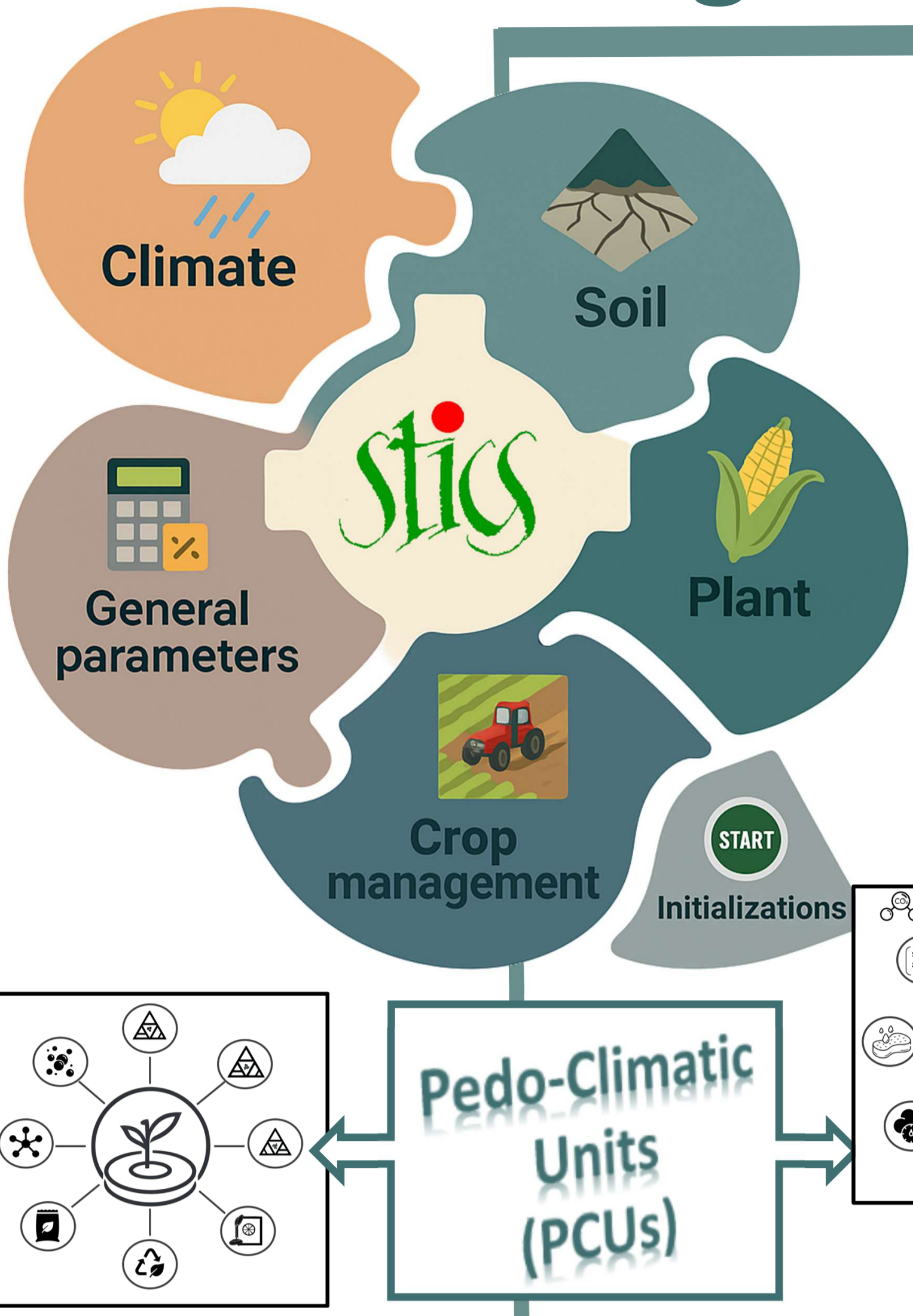


How to use European soil and climate data to simulate agriculture with crop models for bioeconomy scenarios ?

Hugo SENEGES, H el ene RAYNAL, Julie CONSTANTIN,
Univ Toulouse, Toulouse INP, EIP, INRAE, AGIR, Castanet-Tolosan, France,
Hugues CLIVOT, UMR FARE, URCA-INRAE INRAE Reims
November 2025



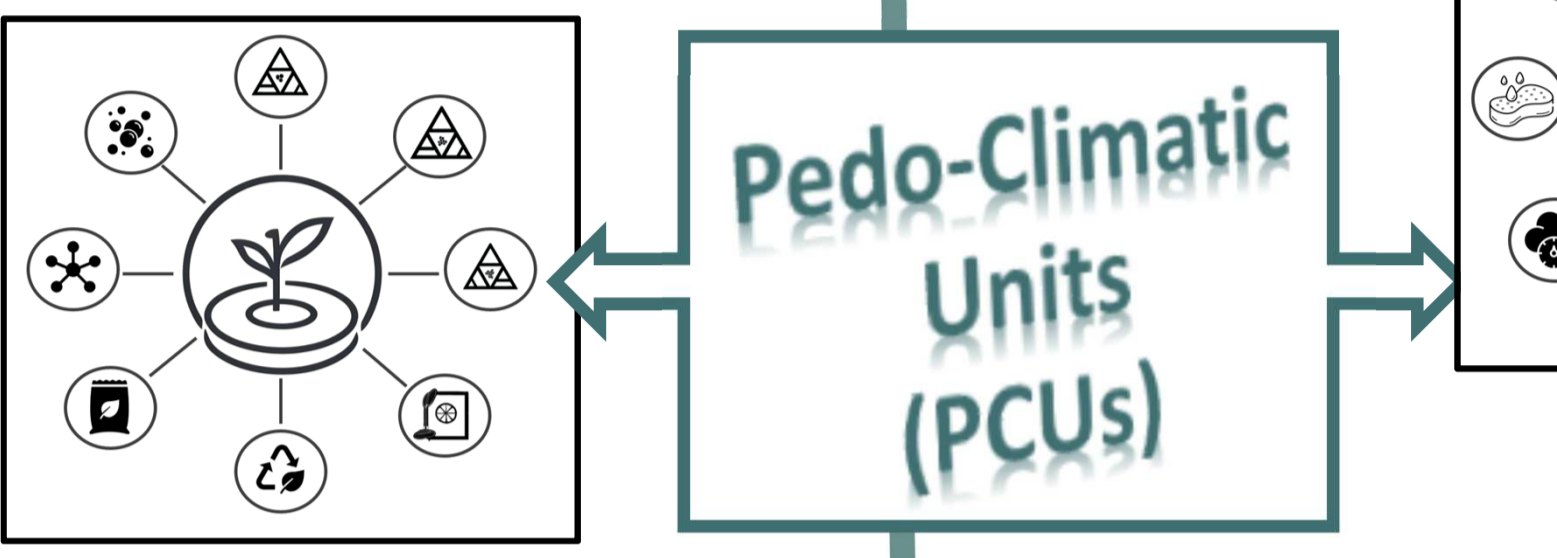
Resume :

- ❖ Integrated Assessment and Modelling (IAM) research project aiming at better bioeconomy territorial strategies including:
 - quantification of crop residues and mobilizable cover crops biomasse,
 - diversified production systems, allocation of raw materials and finished products ...
- ❖ **European extension** (and adaptation) of the former 4p1000 study, based on the **STICS** model
- ❖ As a **first step** : combination of climate and pedological data to feed the STICS model in Europe

Climate chronicle and spatial division :



- ❖ Variables of Interest (**Vols**): temperature, radiation, precipitation, wind, pressure, humidity and snow
- ❖ Chosen product: **Agri4Cast** (JRC), a daily chronicle by **25 km² grid cell**. Comparing it with other datasets shows:
 - Better sensitivity to heat waves events, precipitation, or strong winds.
 - Less than 1% of mean relative difference in values between Agri4Cast and control stations (per variable and per day)



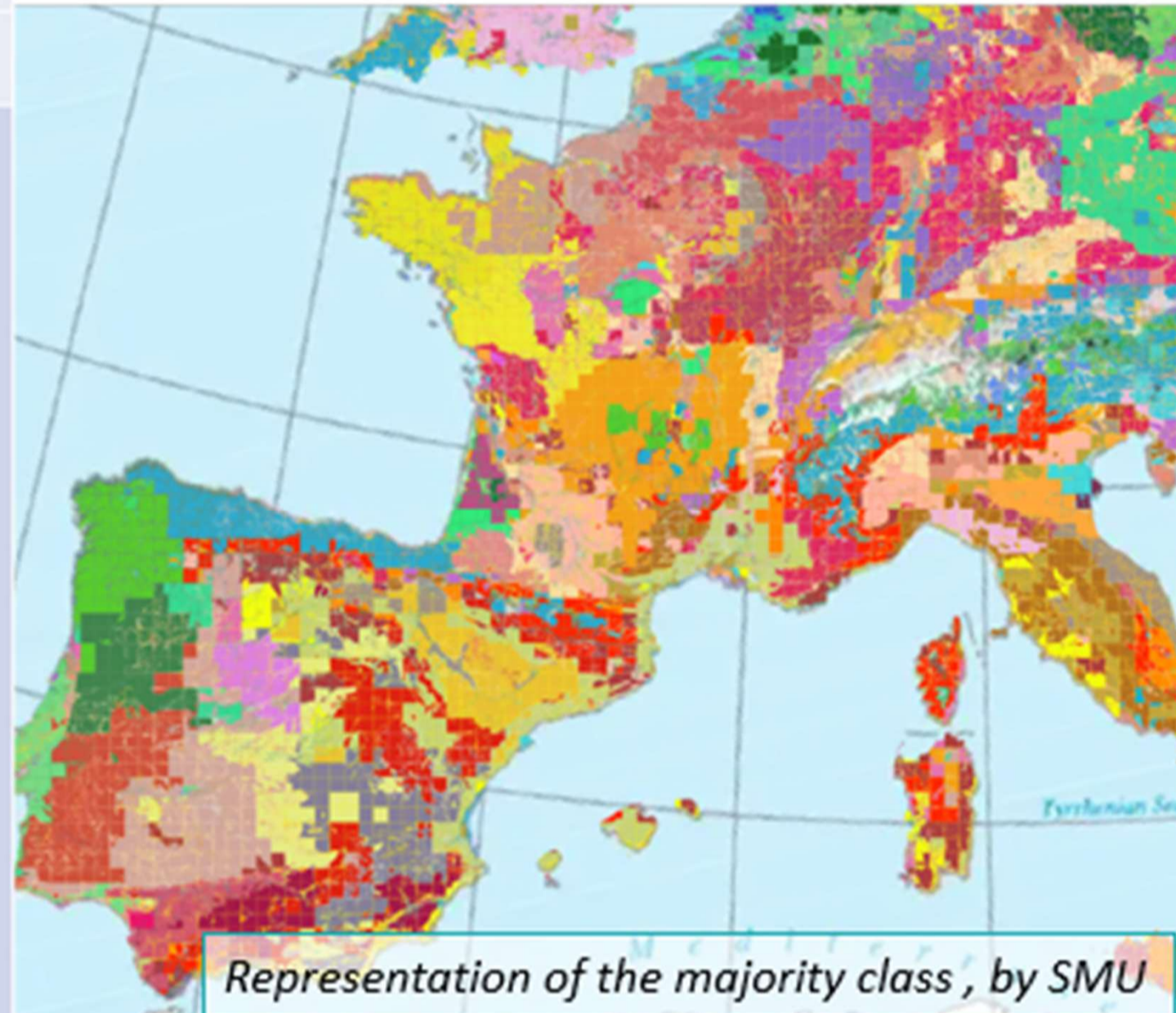
Former Soil Atlas of Europe (JRC)

- All types of soils.
- "World Reference Base" nomenclature (pedological determination key).
- 25 soils groups (mapped at scale 1:1 million) and 155 soil types (**unlocalized**).
- ➔ Delineation of groups of soil types (called Soil Mapping Units). Specialized and standardized nomenclature

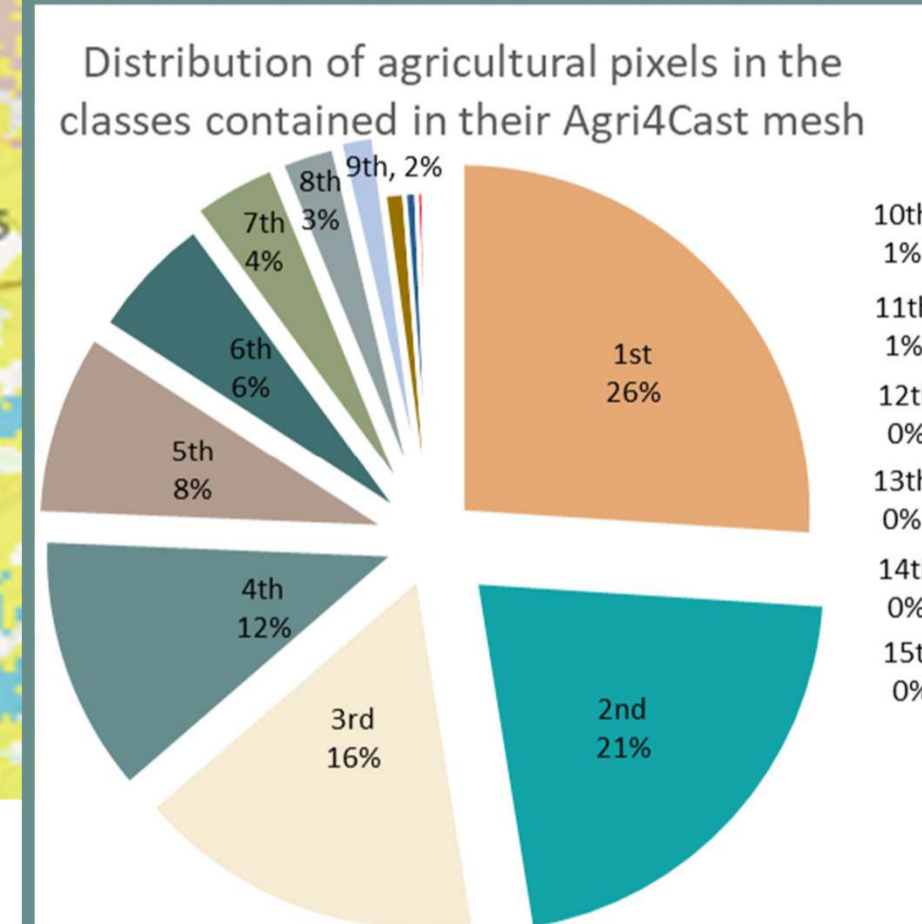


New K-means classified PCUs

- Focus on cultivated soils.
- 80 class unsupervised clustering (17 Vols space).
- Classification on textures, structure and composition data. Resolution at 250m
- ➔ Categorization, designed to ease semantically relevant sampling, based on floating Vols. Unnamed and *sui generis* nomenclature



For each of 7,700 climate cell of the AG4C grid, we select as much as 15 pixels. Each one represent its pedological class among its cell, as long as this class contains at least 12 pixels. The selection is based on the overall distance between pixel's and class center's Vols value (shortest distance)



- ### Soil product : ISRIC
- ❖ **Vols**: texture, pH, SOC, nitrogen, bulk density and pebbles
 - ❖ Chosen product: **SoilGrids**, only European scale information system for horizons below 30cm, and at a resolution of **250m**. JRC's Lucas program is relevant for topsoil and include a "CaCo3" data that completes our data set
 - ❖ Qualification shows a mitigation of extreme values and a poor spatial accuracy. For several Vols, the margin of error can reach 60%
 - ❖ Most unprecise Vols: soil depth, pebble content and porosity. Propagation to the variables derived by pedotransfer function (humidities, porosity...)

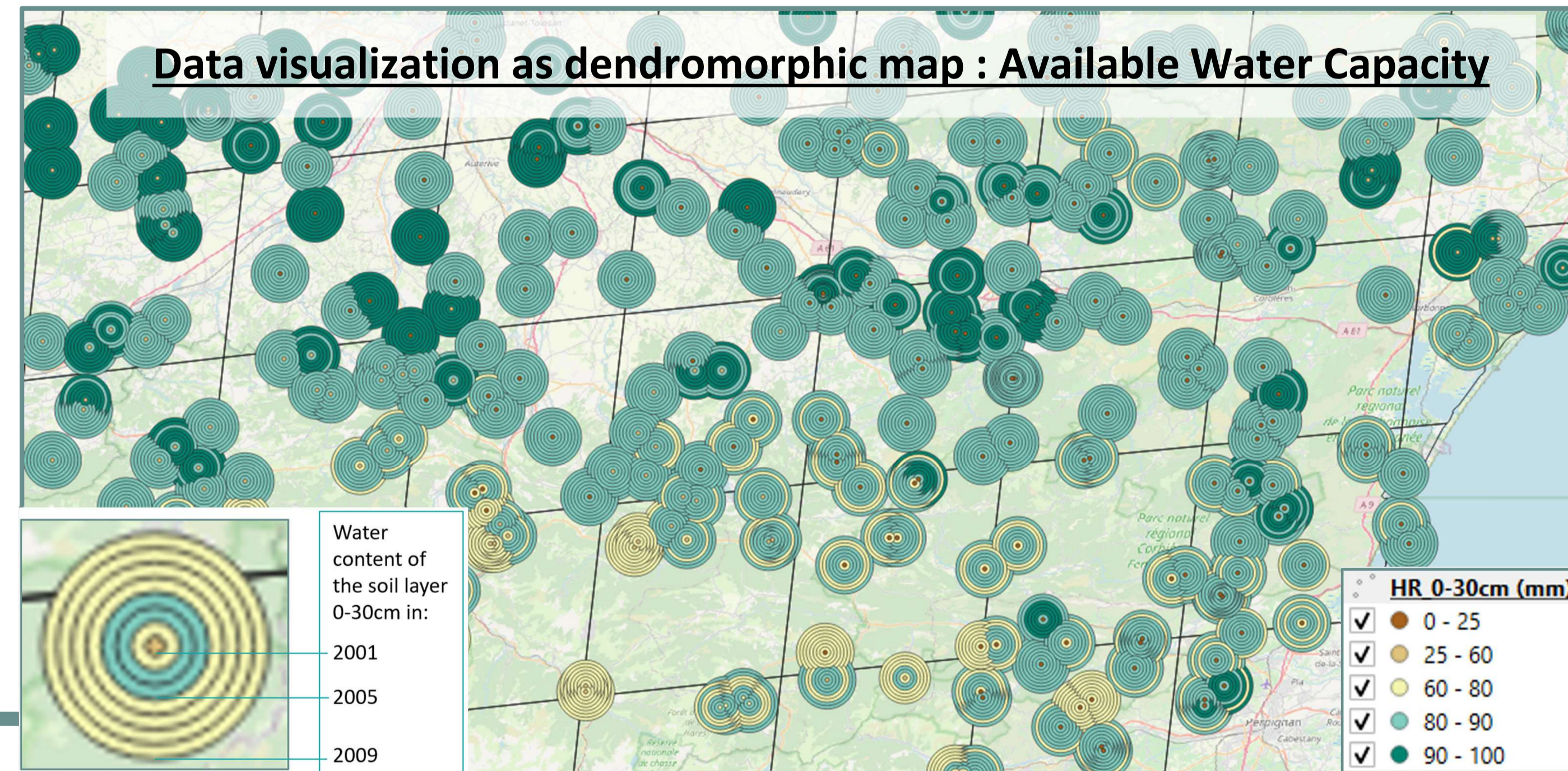
Sampling and classification by k-means :

New Dataset

- ❖ **Optimization** of the simulation cost implies grouping pedo-climatic units (PCUs) by climatic grid cell
- ❖ Method: Clustering the soil product's pixels into **80 classes**, inside a 17 dimension (Vols) space. Most relevant soil-types filtered at 12 pixels by PCU
- ❖ **Qualification**: 47,400 PCUs, out of 28 million agricultural pixels (without grassland) ➔ a sampling rate of 1 for 580
Average deviation at the value of the class center = 12,5%
Standard deviation = 7.5 percentage points.

Preliminary test on France :

- Why France only: existing reference data (on crop management and soils), 4p1000 study's legacy...



- ### Implementing model simulation:
- Inputting the PCUs
 - registration into a table of simulation hyper-parameters (area, cultivars, crop rotations, location...)
 - lead operations sequence

Results:

- 7450 PCUs modelled (15% of Europe's PCUs) for the 2000's, with a fictional maize cover everywhere
- Analyzing several outputs:

Large scale application plane :

- ❖ **Parallelized simulation** jobs sent to distant cluster "ROMEO" (Reims' University) via containerization and SSH connections
- ❖ Independent simulation of each PCU by a parallelized worker inheriting context, parameters and scenario
- ❖ Continuous improvement scheme: various scenarios or hypothesis can be tested, and the simulations outputs can be refined
- ❖ What we still investigate (at the time of November 2025) : data on **cropping system management** in some European countries. Prospective scenarios also need to be defined

