



# A homogeneous database at mediterranean basin scale and machine learning analysis to understand the potential

Michel Moulery, Claude Napoleone, Davide Martinetti, Maria Esther Sanz  
Sanz

## ► To cite this version:

Michel Moulery, Claude Napoleone, Davide Martinetti, Maria Esther Sanz Sanz. A homogeneous database at mediterranean basin scale and machine learning analysis to understand the potential. 4. Open Science Meeting 2019, Apr 2019, Bern, Switzerland. 2019. hal-02104999

**HAL Id: hal-02104999**

**<https://hal.archives-ouvertes.fr/hal-02104999>**

Submitted on 19 Apr 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# A homogeneous Database at Mediterranean basin scale and Machine Learning analysis to understand the potential representativeness variables

Michel Mouléry<sup>1</sup>, Claude Napoléone<sup>1</sup>, Davide Martinetti<sup>2</sup>, Esther Sanz Sanz<sup>1</sup>

<sup>1</sup> INRA PACA, Unité Ecodéveloppement, Avignon, France.

<sup>2</sup> INRA BIOSP, Unité biostatistique et processus spatiaux, Avignon, France.



4<sup>th</sup> Open Science Meeting 2019  
Transforming Land Systems for People and Nature  
April 24 – 26, 2019 · Bern, Switzerland



DIVERCROP : Land system dynamics in the Mediterranean basin across scales as relevant indicator for species diversity and local food systems

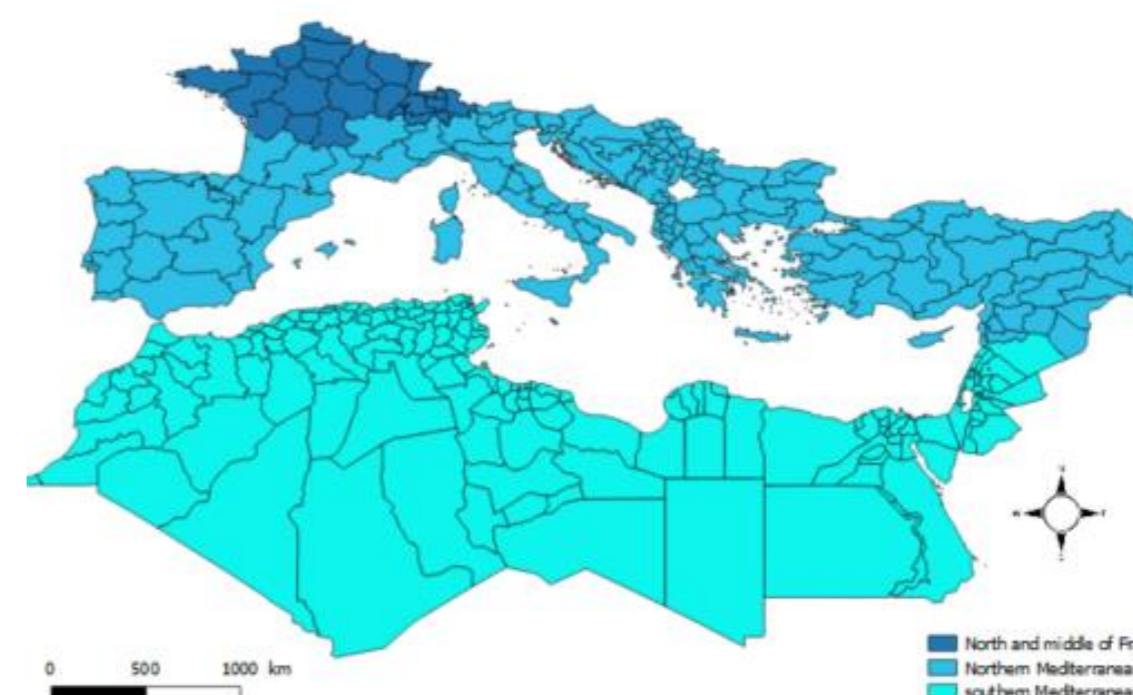
<http://divercropblog.wordpress.com/>

ID 263

**1 Research question : Which are the variables best representative for the crops production in Mediterranean? (wheat case study)**



**2**



**3 The mediterranean homogeneous spatial database**

**LAND COVER / LAND USE**

CCI - ESA (Louvain) – 300 m résolution (2005) – raster images  
natural vegetation, Tree cover  
Scrubland, Urban areas  
Bare areas, Water bodies

**MAPSPAM MODEL - 42 CROPS**

(2005) – Raster images  
Area, production, yield, crop suitability, irrigation, rainfield  
Pixels 5 min ~ 8- 10 Km resolution

**POPULATION (URBAN, RURAL)**

LANDCAN Database (2005, 2015), Hyde Database

**PROTECTED AREAS**

World Database on Protected Areas (vector)

**SOILS TYPES & QUALITY**

Harmonized world soil database

**GROUND**

DTM: Global 30 Arc-Second Elevation (GTOPO30) USGS (slope) – raster image

**LIVESTOCK**

Gridded Livestock of the world

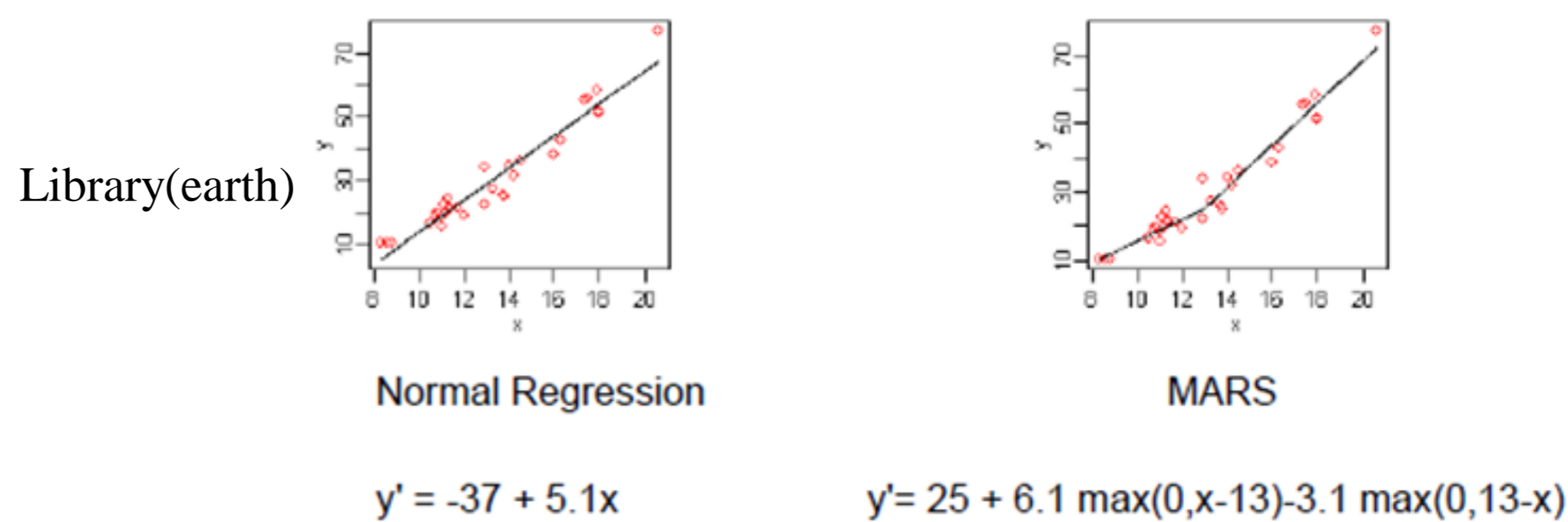


Peter Horstbol Møller

**4**

## Methodology & Tools

Multivariate adaptive regression splines



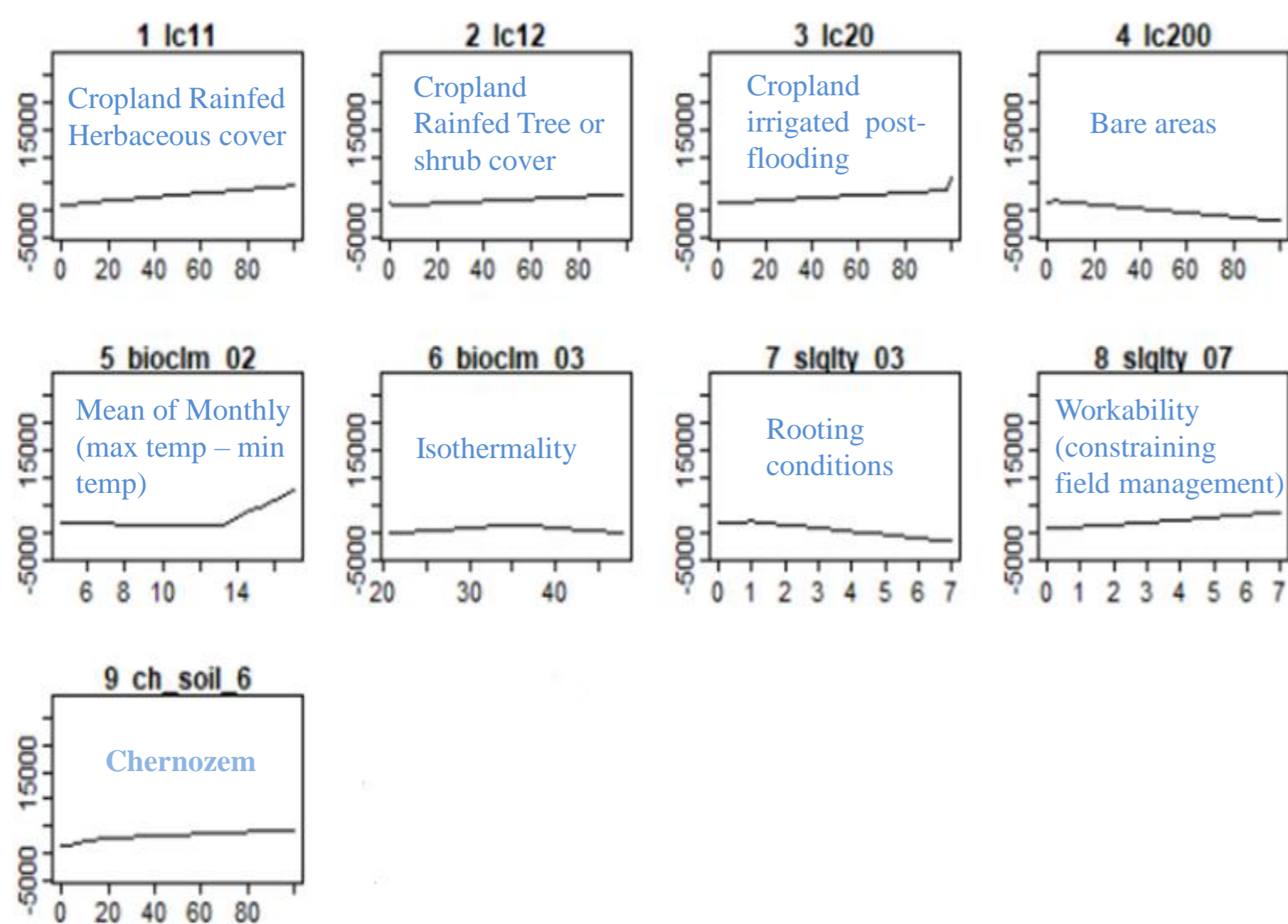
- A) Join all informations in a dataframe
- B) Choose dépendent variables (42 crops)
- C) Choose Independent variables (other variables, urban, population, livestock, altitude, soils ..)
- D) Launch model with function Earth with R

```
mod1_nord <- earth(x=nx, y=ny[,1]) # y = dependantes x non dependantes
```

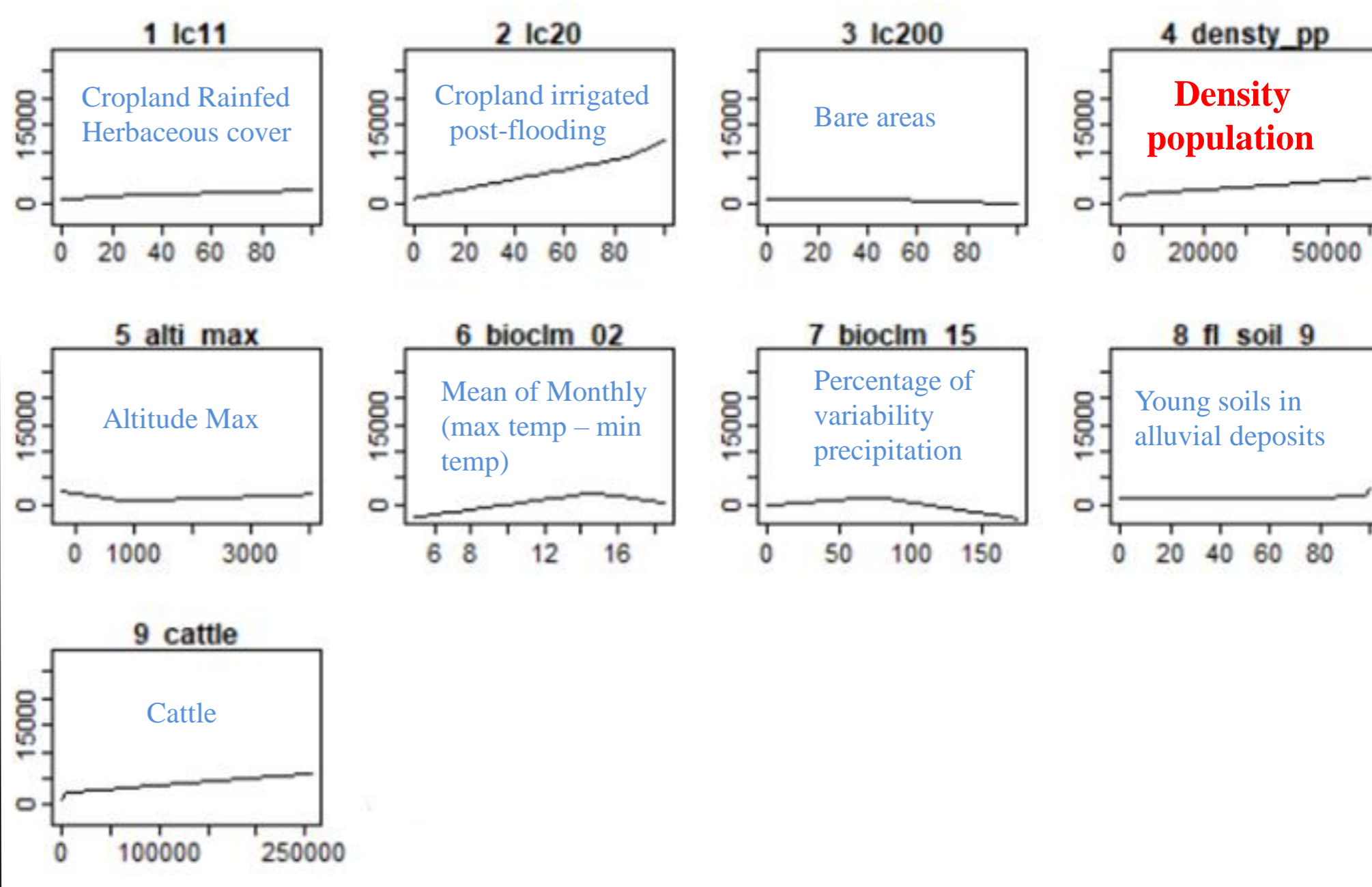
**5**

## Results on wheat production

### NORTHERN MEDITERRANEAN



### SOUTHERN MEDITERRANEAN



**One example of counterintuitive result :**

➔ Population is a good incentive in the south of the basin (wheat may be a production close to the city) contrary to the north, where there is no effect (production areas and cities are not in the same location)

Thanks to: Johanna Fusco (Inra – Institut National de la recherche Agronomique, Avignon), Ricardo Villani (Institute of Life Sciences, Italia), Leonith Hinojosa-Valencia (Louvain University), Tiziana Sabbatini (Scuola Superiore Sant'Anna [Pisa]), Alberte Bondeau (IMBE - Institut méditerranéen de biodiversité et d'écologie marine et continentale).

This database is owned by INRA (France), issued from the DIVERCROP research project (N° ANR-16-ARM2-0003-01) funded by Arimnet2 program (FP7 – ERA-NET no. 618127; Mediterranean Agriculture); see <https://hal.archives-ouvertes.fr/hal-01907477>. For all dissemination and use, contact [marta.debolini@inra.fr](mailto:marta.debolini@inra.fr) or [claudio.napoleone@inra.fr](mailto:claudio.napoleone@inra.fr)

Régression multivariée par spline adaptative. (2018, décembre 3). Wikipédia, [http://fr.wikipedia.org/w/index.php?title=R%C3%A9gression\\_multivari%C3%A9e\\_par\\_spline\\_adaptative&oldid=154474849](http://fr.wikipedia.org/w/index.php?title=R%C3%A9gression_multivari%C3%A9e_par_spline_adaptative&oldid=154474849).

Assessment of wheat yield gap in the Mediterranean (Mustafa Pala et al, 2011), ICARDA, [geoagro.icarda.org](http://geoagro.icarda.org) (Assessment\_of\_wheat.pdf)