Infrared spectroscopy : the second life of the spectra

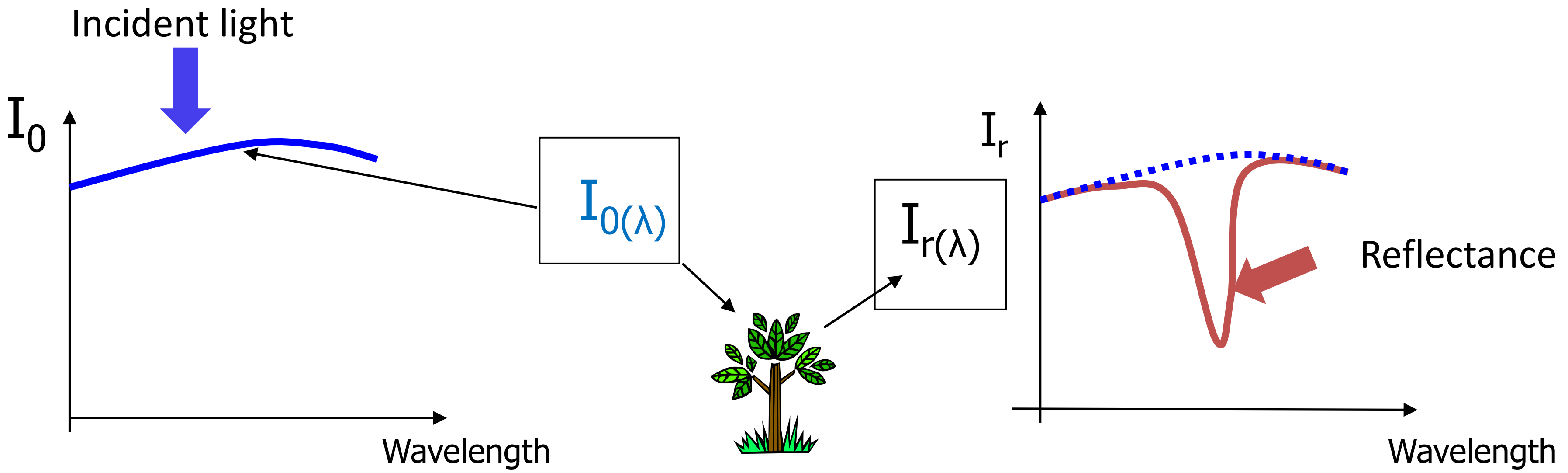
Héloïse Villesseche, Elsa Ballini[1], Ryad Bendoula[2], Martin Ecarnot[3], Nathalie Gorretta[4], **Pierre Roumet[3]**

[1]Momie UMR Phim, [2]Comic UMR Itap, [3]Ge2pop Umr Agap, Montpellier

[4] Inrae DAR Diagonal, Paris

Collaborations :

What's spectroscopy ? A light matter interaction



$$\% \ R=(I_{r(\lambda)}/I_{0(\lambda)})*100$$

Reflectance : signature of chemical and physical properties of the targeted matrix (leaf, grain, roots …)

Measurement scales : Organs (Nirs)
Pixel (hyperspectral imaging –IHS-)

How to analyse spectra data ?

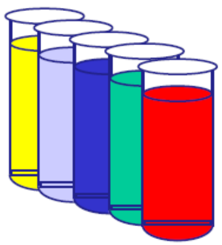
<u>First life of the spectra : the supervised analysis</u>

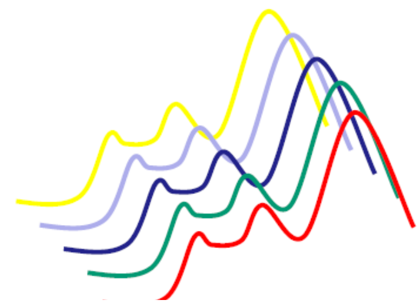
= To calibrate the spectra signal to infer some matrix properties

Set of common sample set

Reference lab. Analysis

Spectra data





Different steps before to have an accurate and robust model :

- Calibration
- Validation
- Routine Analysis

Some applications : Technological properties of grains

Semolina & Pasta values

Spectra + reference analysis (>20 years)

Signal calibration

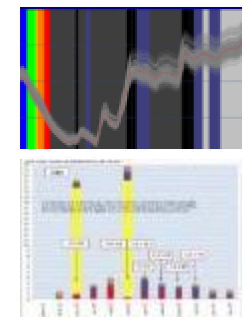
1st processing

2nd processing

+

Scan NIrs

Wet analysis



<u>Milling value :</u>

Semolina yield, vitreoussness, TKW, Test Weight FHB contamination, Issue and flour rates : $0{,}78 < R^2_{valid} < 0{,}98$

<u>Pasta value</u> :

GPC : $R^2_{valid}=0{,}98$

Color Indexes $R^2_{valid}=0{,}72$.

(Compan et al., 2013; Jaillais et al. 2015) .

Hyperspectral Imagery (IHS) : Clustering

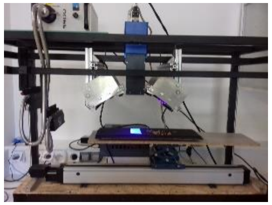
Objective : to predict the percentage of each component in binary genotype mixtures

28 binary mixtures (X 5 rep.)

Spectra

Nir + Sorter

Spectra + grain shape



IHS

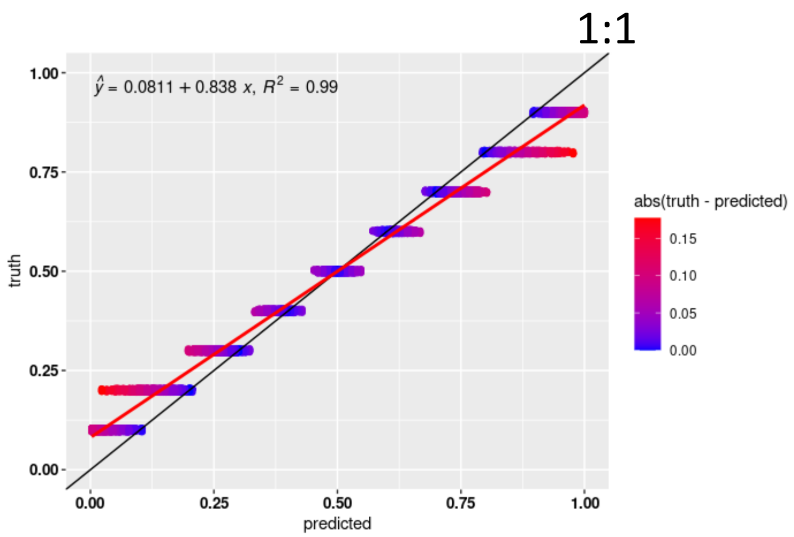
Hyperspectral Imagery (IHS) : Clustering

Objective : to predict the percentage of each component in binary genotype mixtures

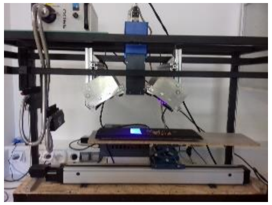
28 binary mixtures (X 5 rep.)

<u>*Spectra*</u>

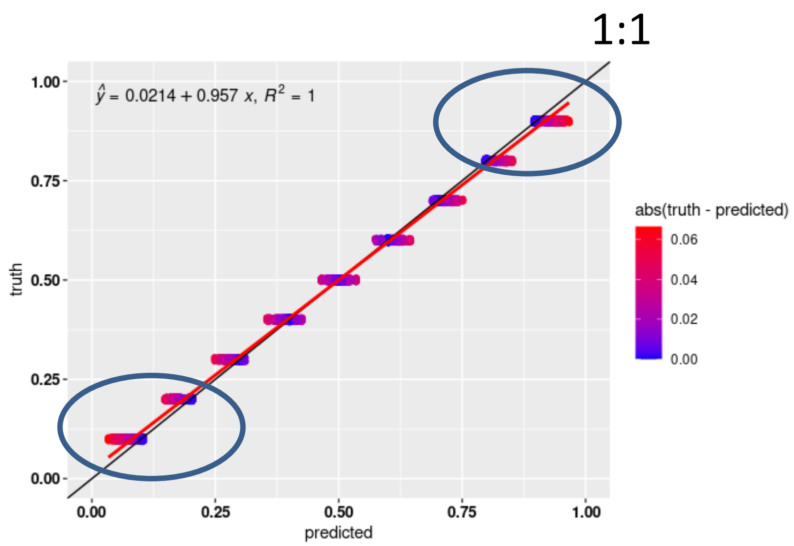
Grain distributor +
Nir_sorter



<u>*Spectra + grain shape*</u>



IHS



IHS : High prediction quality

(M2, A. Belny, 2021)

First spectra life : the supervised analysis

Linking spectra data with qualitative distribution or quantitative variations of various traits

- Modelling based both on linear (PLS…) or non linear algorithms (SVM, NN)

⇒ *it works with ± loss of information*

- To decipher complex phenomenons in elementary components

⇒ *Multiple losses + not possible to take into account interactions between components*

An alternative : the unsupervised modelling

Second spectra life : the unsupervised analysis

<u>Hypothesis</u>: Spectral data provide a global picture of the state of the matrix;
Any change (biochemical content, physical properties) will alter the spectral signal.

<u>Consequences</u>

1. These spectral changes = alerts for the studied phenomenon
2. No calibration to characterize the targeted phenomenon

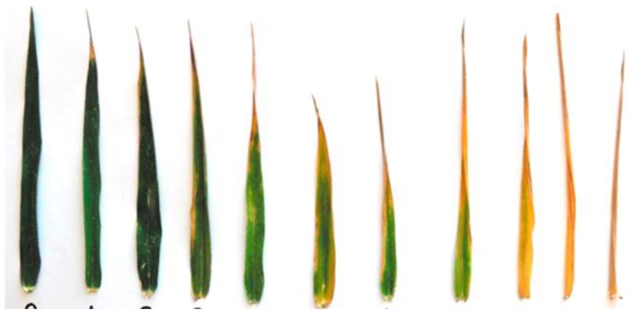
<u>Which targets</u>?

To document time course of complex traits and looking for some discrepencies

Two examples : Foliar senescence and plant pathogen interaction.

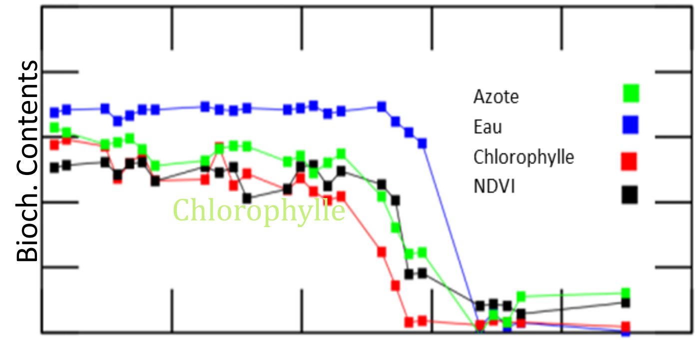
Flag leaf senescence

Several traits involved. Focus on water, nitrogen, chlorophyll contents



Flowering

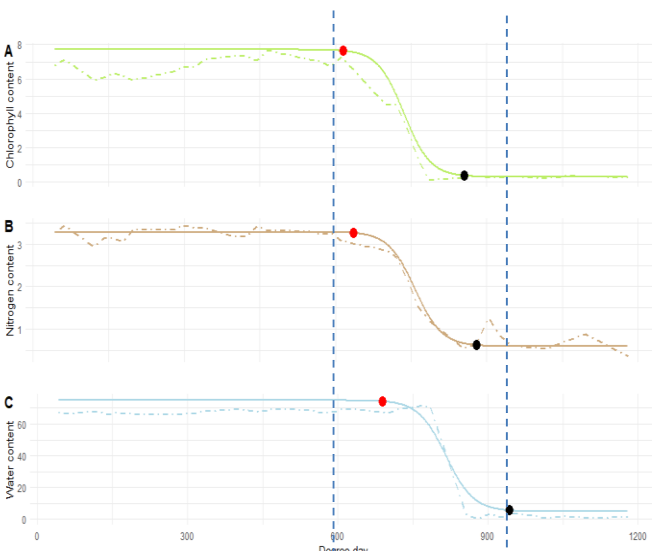
Phys. Maturity



°Gdd

Flowering 400 800 Phys. Maturity

T0 T1



Chlorophyll

Nitrogen

Water

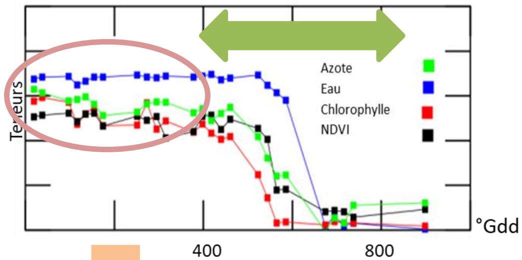
Flowering

Phys. Maturity

- N, H_2O & Chlorophyll : Similar time courses (plateau/drop/plateau) modeled with a logistic curve with 4 parameters.
- Time lag between time courses
- Two key points : T0 and T1 for the beginning and the end of senescence

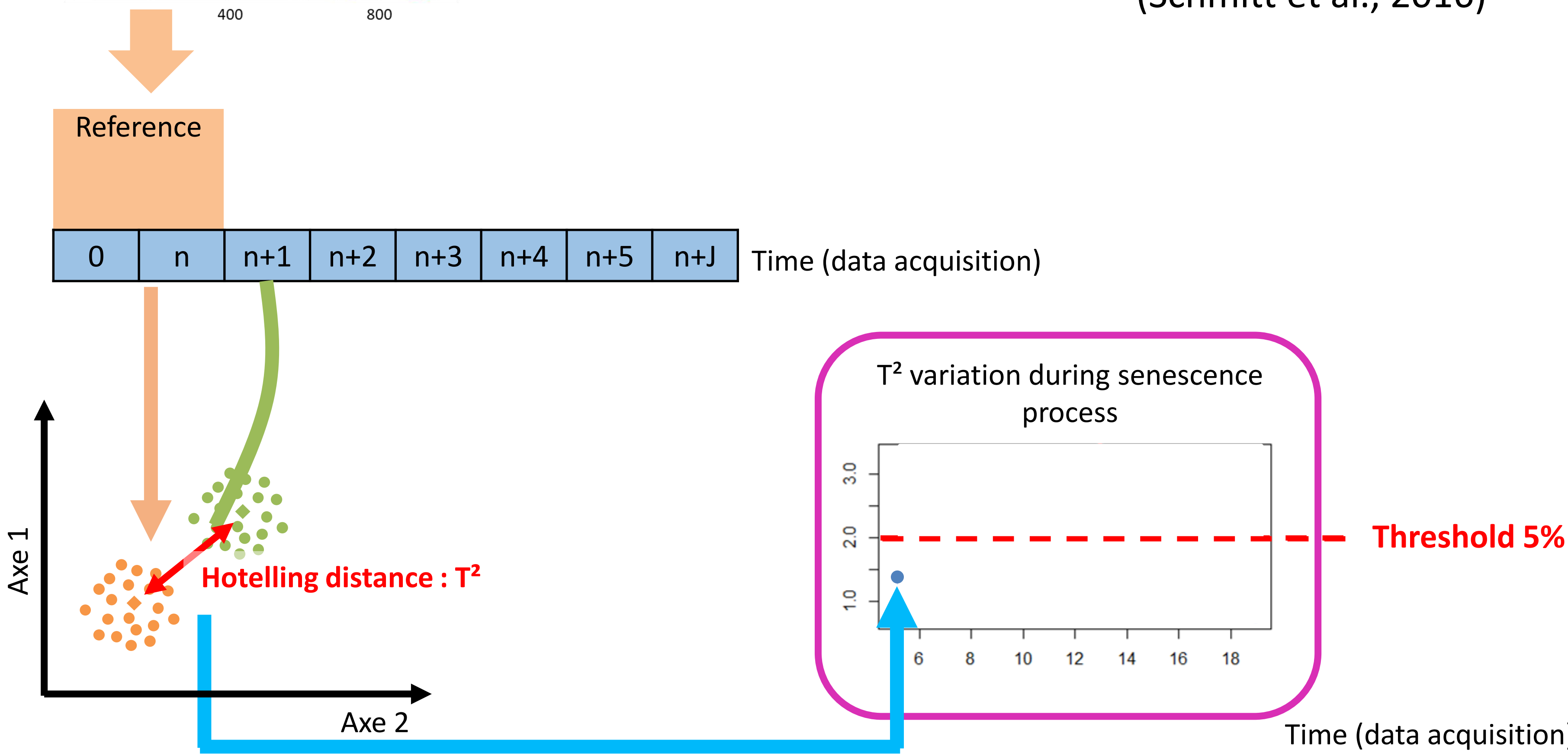
Flag leaf senescence

Analysis based on Moving Window Principal Component Analysis (MWPCA)

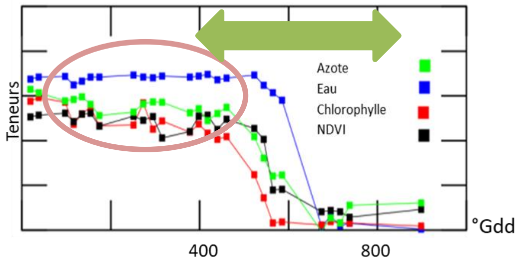


MWPCA principle
(one MWPCA per leaf)

(Schmitt et al., 2016)

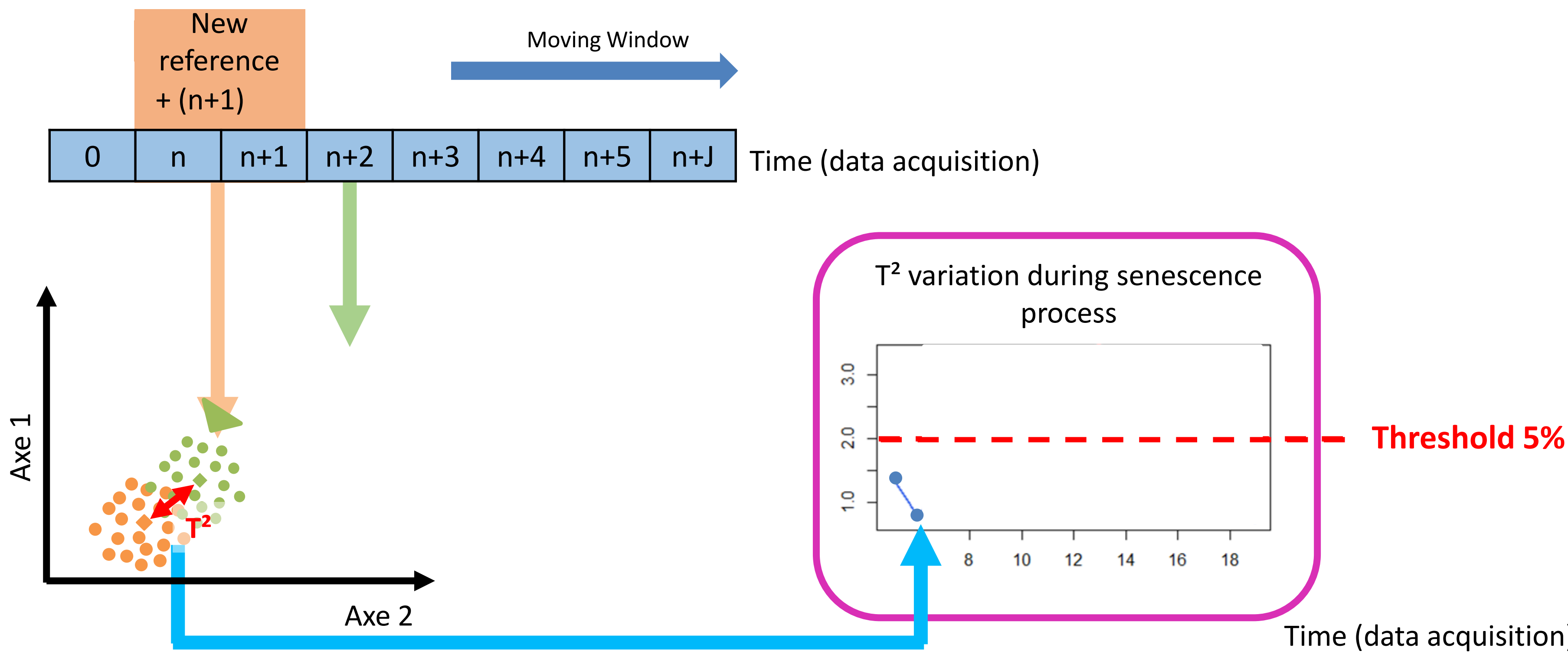


Flag leaf senescence

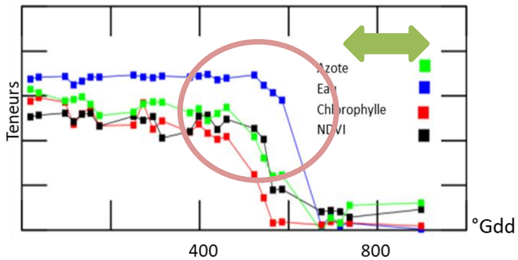


MWPCA Principle

(Schmitt et al., 2016)

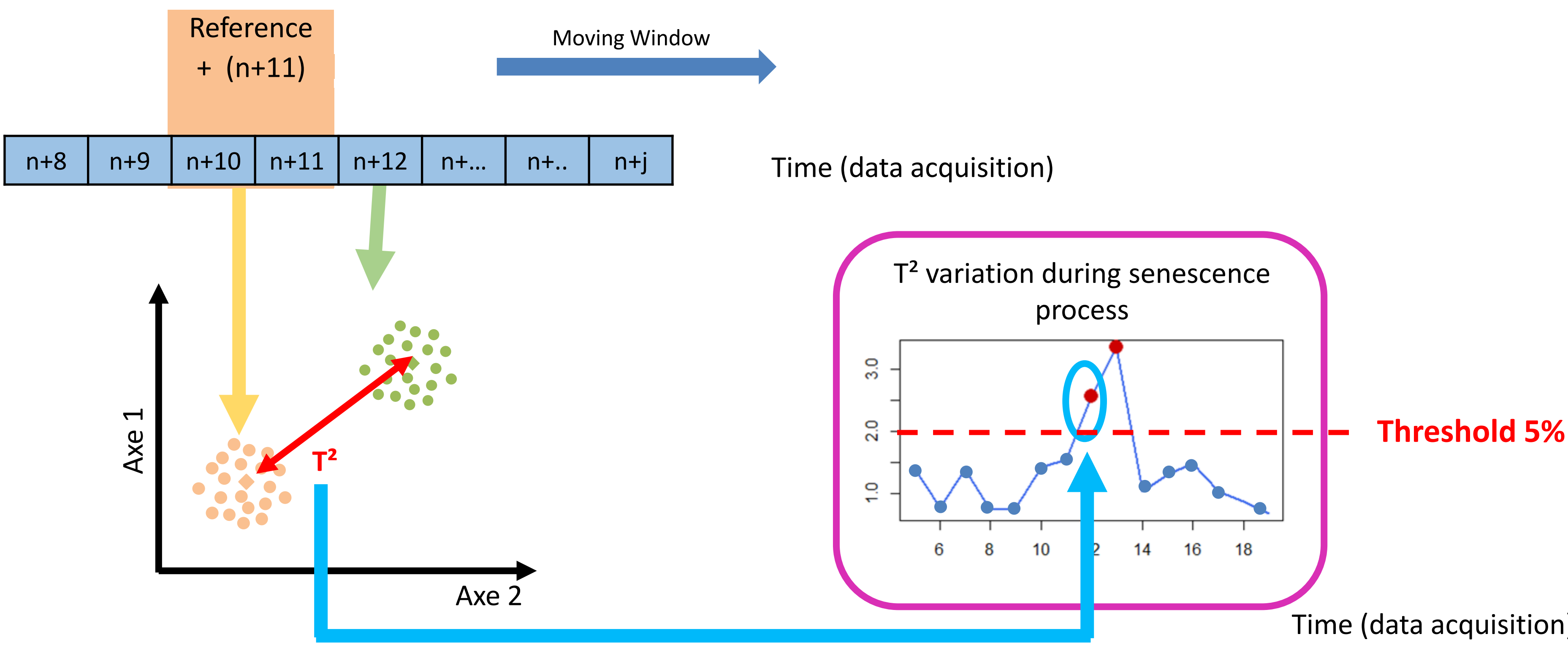


Flag leaf senescence



MWPCA Principle

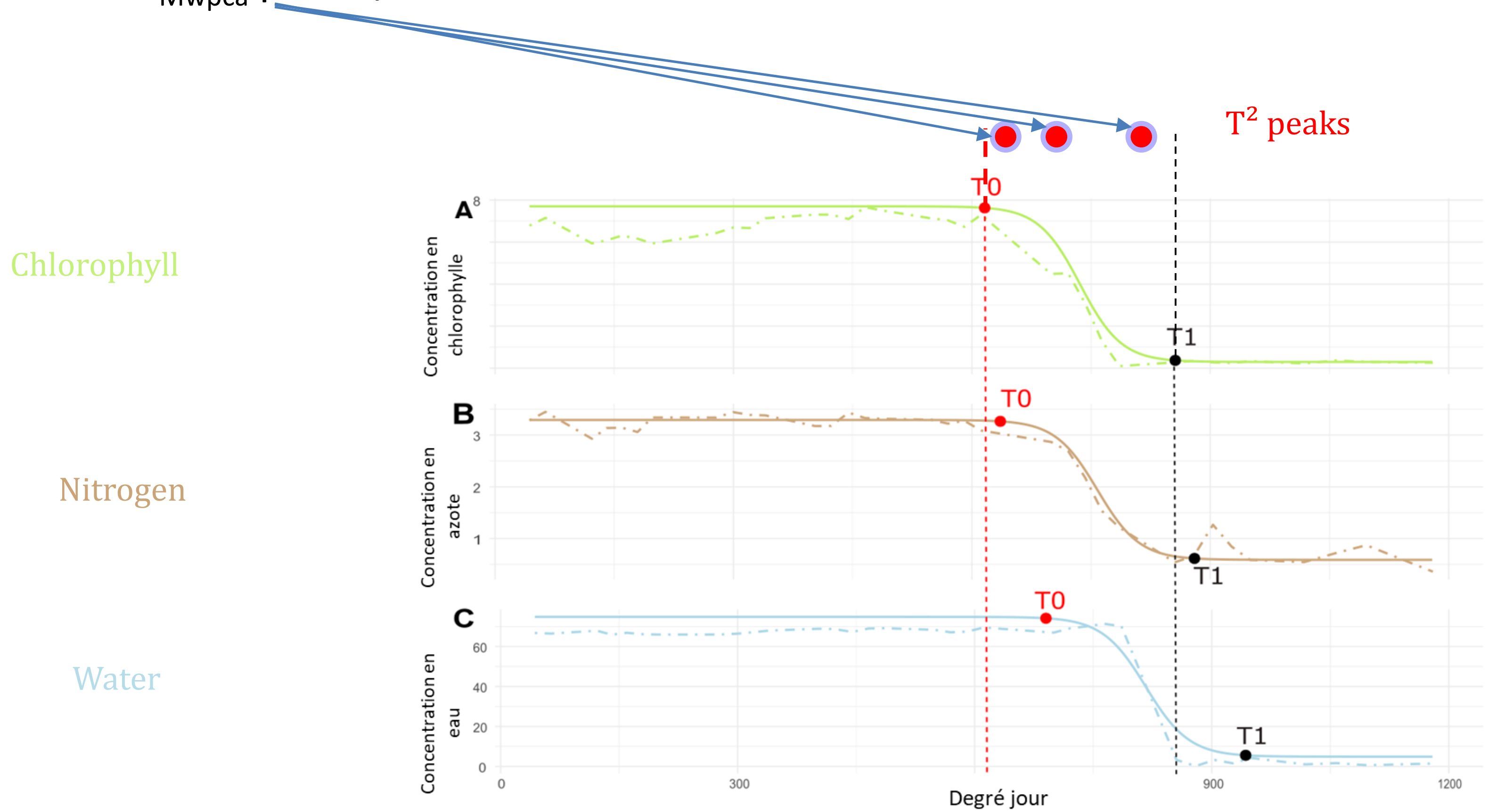
(Schmitt et al., 2016)



Flag leaf senescence

N=139 durum wheat plants : large variability for senescence parameters (genotypes, treat.)

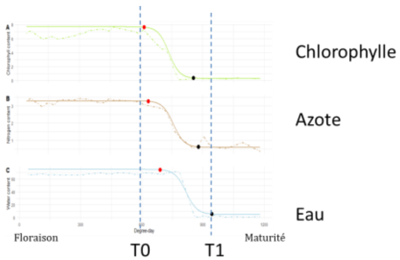
From 2 to 4 T^2_{Mwpca} peaks / leaf

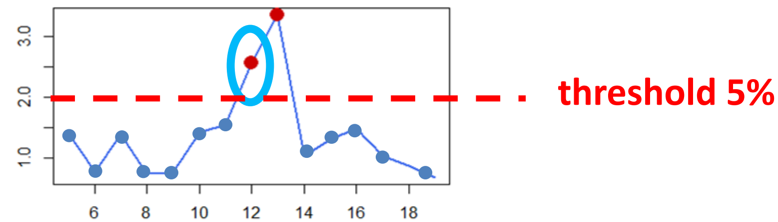


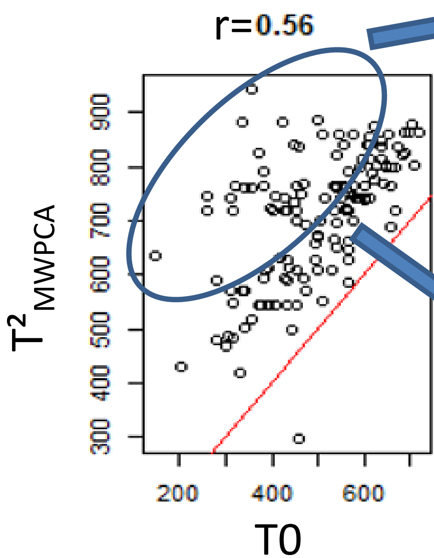
T^2_{Mwpca} Peaks matching with the senescence process

Flag leaf senescence

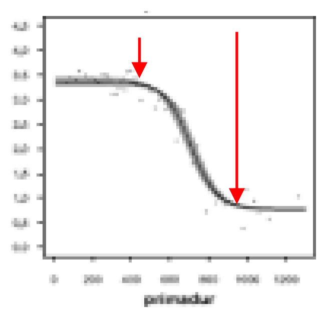
Linking T0 (beginning of senescence) and the first peak T^2_{Mwpca}



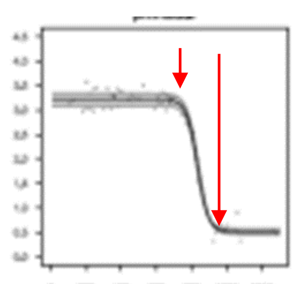




Slowy Senescence



Speedy Senescence



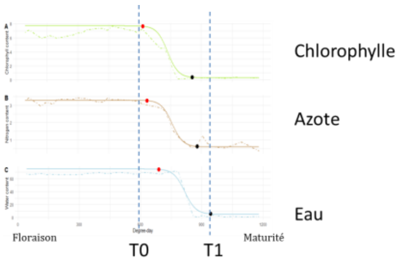
T^2_{Mwpca}

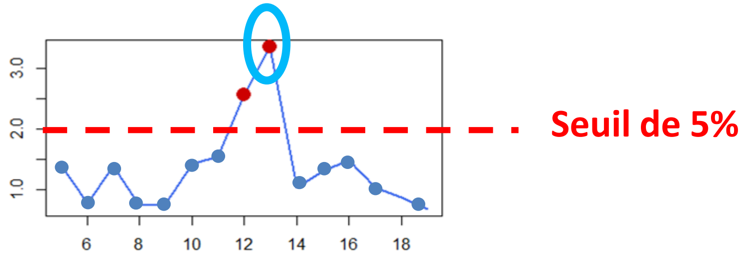
- Consistent with T0 (r=0,56***)
- Delayed Signal (+120 GDD)
- Informatif in case of rapid Senesc.
- Unformative with slow senesc.
  =limit of our method

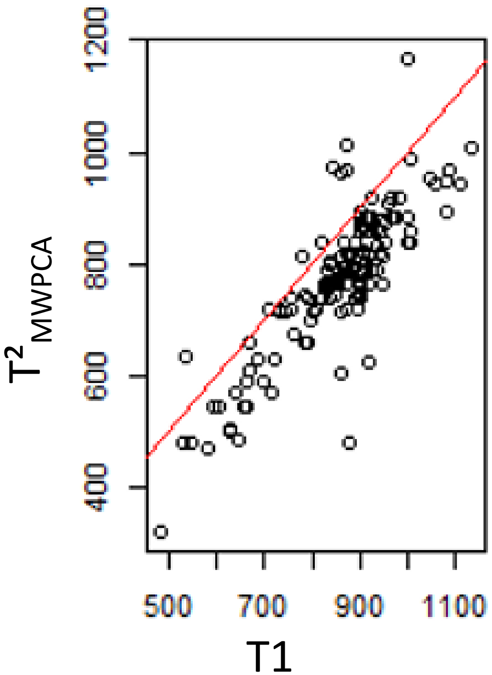
(Villessèche et al; 2022)

Flag leaf senescence

Linking T1 (End of senescence) and the last peak T^2_{Mwpca}







T^2_{Mwpca}

- Very consistent with T1 (r=0,84***)
- T^2_{MWPCA} earlier (- 70 GDD)

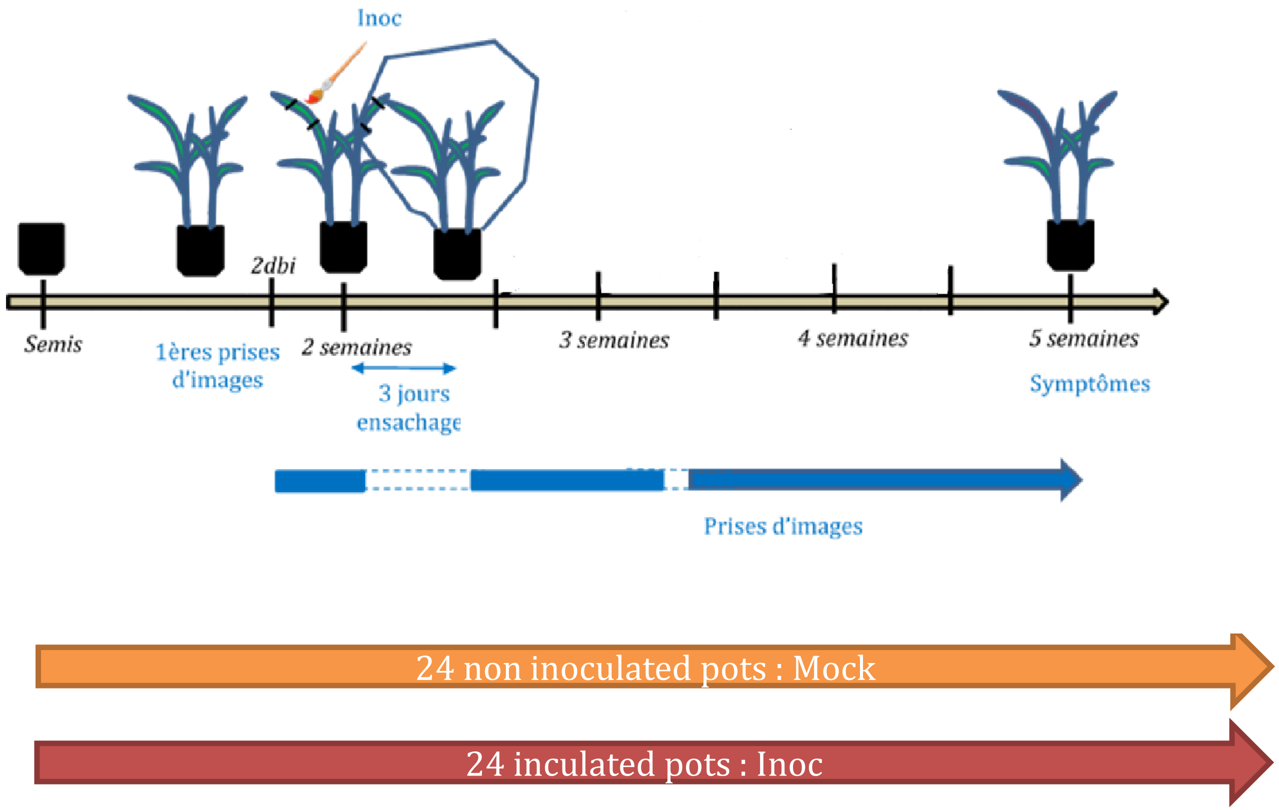
Unsupervised model : promising alternative to document leaf senescence
Caution : its ability to identify slow phenomenons

(Villessèche et al; 2022)

Plant pathogen interaction

Durum Wheat / Septoria IHS

Experimental design



2 susceptible lines : EPO_67 Pycn. -; EPO_68 Pycn. +

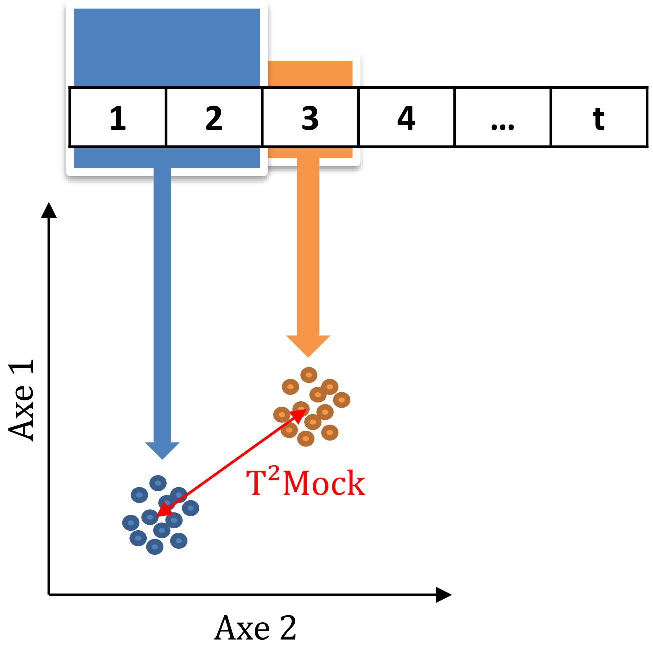
1 measurement / day from 1 dpi to 20 dpi

Plant pathogen interaction

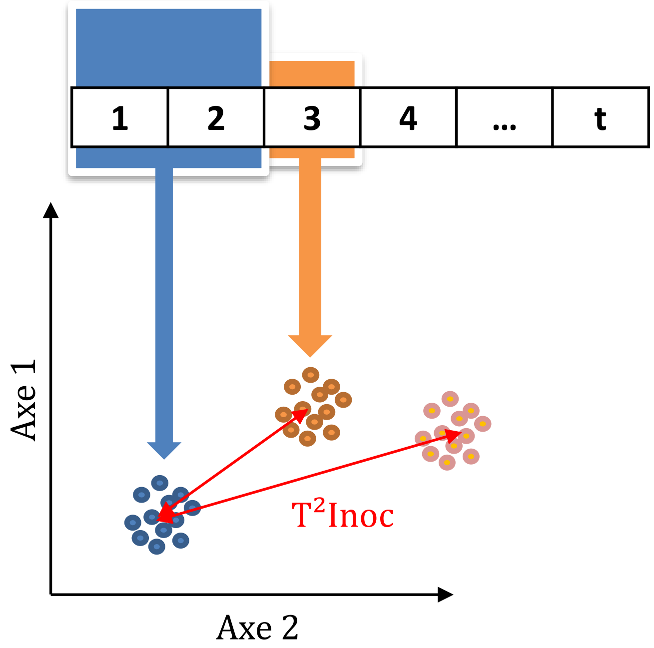
Durum Wheat / Septoria IHS

MWPCA analysis

- Mock leaf



- Inoc Leaf

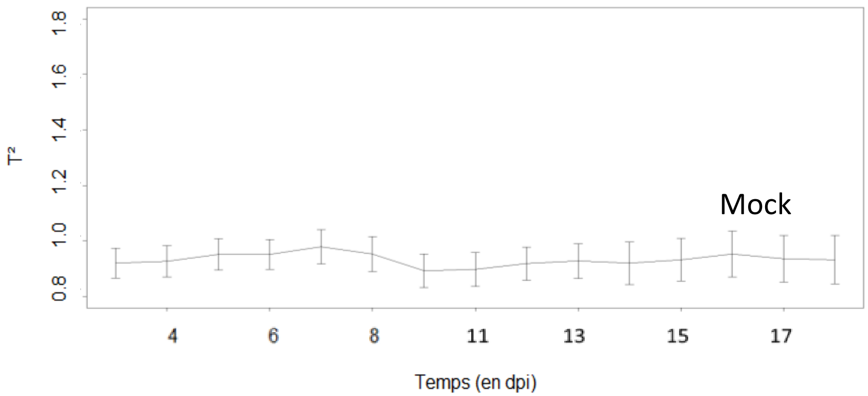


Plant pathogen interaction

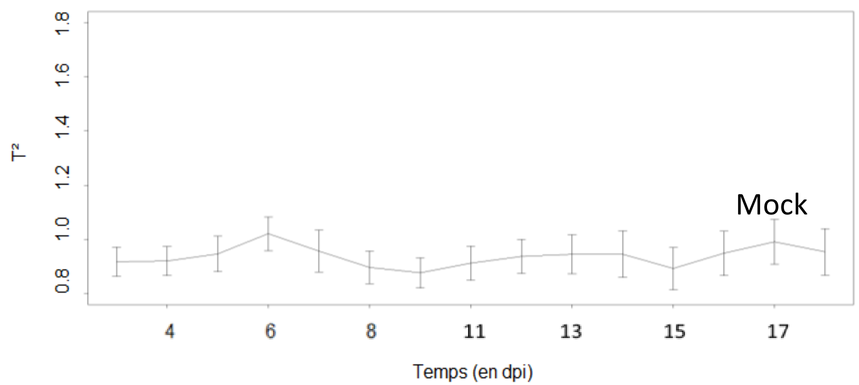
Durum Wheat / Septoria IHS

Mock leaves : mean T^2_{MWPCA} (pixels, leaves)

EPO_067 Pycn.-



EPO_068 Pycn. +



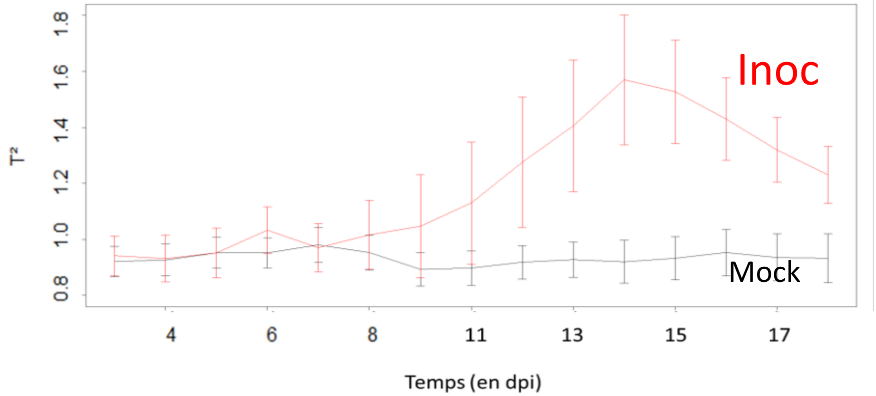
- Values from 0,8 to 1,0 ; low and stable values throughout the experiment
- No T^2 peaks : slow leaf processes
- Similar for the 2 lines

Plant pathogen interaction

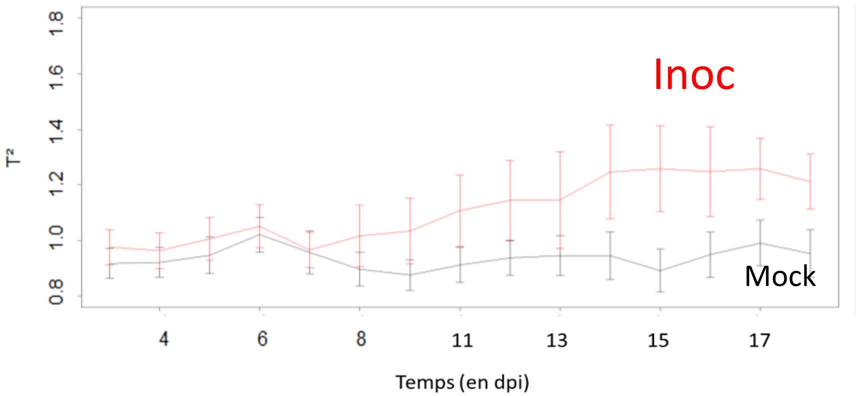
Durum Wheat / Septoria IHS

Mock and Inoc leaves : mean T^2_{MWPCA} (pixels, leaves)

EPO_067 Pycn. -



EPO_068 Pycn. +



- Clear differentiation of Mock/Inoc leaves from 10 dpi.
- Earlier than visual inspection
  - 50 % of leaves with symptoms at 13 dpi
  - 50% of leaves with inoc-like T^2 values at 9-10 dpi

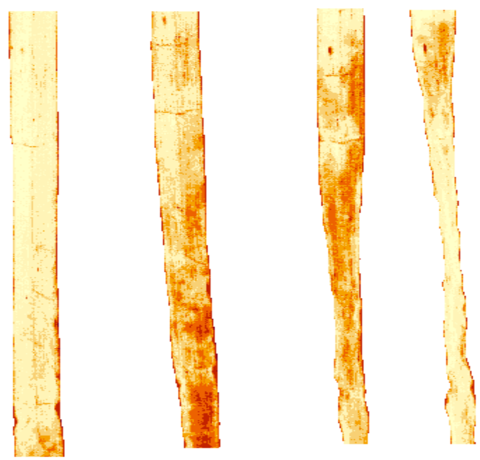
- High T^2 variability in Inoc leaves
  - Leaves with different time course
  - Heterogeneity of T^2_{MWPCA} values

Plant pathogen interaction

Durum Wheat / Septoria IHS

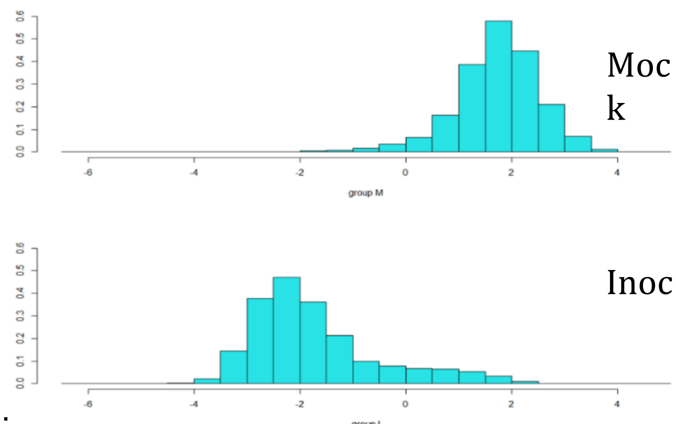
Distribution of T^2_{MWPCA} values of each leaf pixel (Inoc & Mock)

T^2_{MWPCA} Heatmap at 12 dpi



14 15 16 17 dpi

Inoc



Mock

Inoc leaves : High variability of T^2_{MWPCA} values with some values close to mock one's

To identify presence of pathogen ; Early detection ?

T^2_{MWPCA} Heat map are consistent with the necrotic areas

Different patterns according the level of resistance ?

What design for routine use ?

Spectra : multiple lifes and what the future ?

1) Supervised modelling

Signal analysis : linear and non linear models (Deep learning)

2) Unsupervised modelling

Highly complex data: Temporality / spatiality ...

Exciting new perspectives, especially for complex / integrative phenomena complex

Need for validation and degradation of models (spectral, spatial, temporality dimension

Much work on spectra but little on image

4) Strong technological dynamics

Miniaturisation and portability

Falling prices