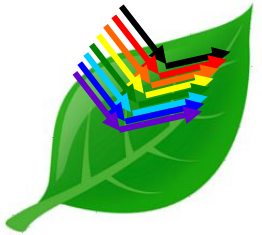
# Selecting Hyperspectral bands for Leaf Mass per Area Prediction by Means of Neural Networks



**Catherine Baskiotis, Josselin Aval, Marwa El Bouz, Ayman Alfalou**
Light Scattering Learning Group, L@bisen, ISEN, Yncréa-Ouest, Brest, France

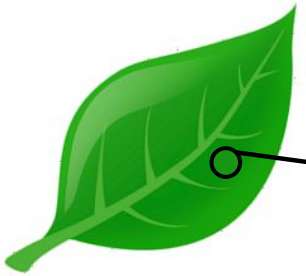
contact : catherine.baskiotis@isen-ouest.yncrea.fr

**WE.V6 Session: Surface Biophysical Parameters**
**Oral Presentation: WE.V6.5**

**International Geoscience and Remote Sensing Symposium (IGARSS),**
**Kuala Lumpur, Malaysia, 17 - 22 July, 2022**

# Leaf Mass per Area hyperspectral monitoring

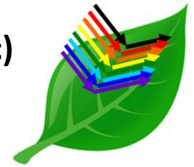
- ## Leaf Mass per Area (LMA) definition



$$\frac{\text{Total Disc Weight}}{\text{Disc Area}} = \underbrace{\frac{\text{Water Weight}}{\text{Disc Area}}}_{\text{Equivalent Water Thickness (EWT)}} + \underbrace{\frac{\text{Dry Matter Weight}}{\text{Disc Area}}}_{\text{Leaf Mass per Area (LMA)}}$$

- ## Leaf Mass per Area hyperspectral monitoring applications

  - ✓ **Precision Agriculture** (Fertilizer management, Early detection of plant diseases, etc)
  - ✓ **Global environmental watching**



# Leaf Mass per Area hyperspectral monitoring

## Open question:

**How much can we reduce the number of hyperspectral bands while keeping a good accuracy for LMA prediction ?**

## Literature answers:

✓ Physics-based PROSPECT model: ~100 hyperspectral bands [1]

*Currently offering the best accuracy for LMA prediction, Relative RMSE ~18%*

✓ Vegetation indices: 2 to 6 hyperspectral bands [2,3]

*Poor LMA prediction accuracy, Relative RMSE ~28%*

✓ Partial Least Square Regression (PLSR): ~15 hyperspectral bands [2]

*Very poor LMA prediction accuracy, Relative RMSE ~38%*

$$\text{Relative Root Mean Square Error (RMSE)} = \frac{\text{Root Mean Square Error (RMSE)}}{\text{LMA mean value over the tested database}}$$


**[1]** J.-B. Féret et al., *Remote Sensing of Environment*, **231**, pp. 110959, 2019.
**[2]** Y. Chen et al., *Remote Sensing*, **13,** pp. 3761, 2021.
**[3]** G. le Maire et al., *Remote Sensing of Environment*, **112**, pp. 3846-3864, 2008

# Study goal

- **Open question:**

**How much can we reduce the number of hyperspectral bands while keeping a good accuracy for LMA prediction ?**

- **Literature answers:**

✓ **Physics-based PROSPECT model: ~100 hyperspectral bands [1]**

*Currently offering the best accuracy for LMA prediction, Relative RMSE ~18%*

✓ **Vegetation indices: 2 to 6 hyperspectral bands [2,3]**

*Poor LMA prediction accuracy, Relative RMSE ~28%*

✓ **Partial Least Square Regression (PLSR): ~15 hyperspectral bands [2]**

*Very poor LMA prediction accuracy, Relative RMSE ~38%*

- **Study goal:** **To improve the accuracy of LMA prediction with a small number of bands (<15)**

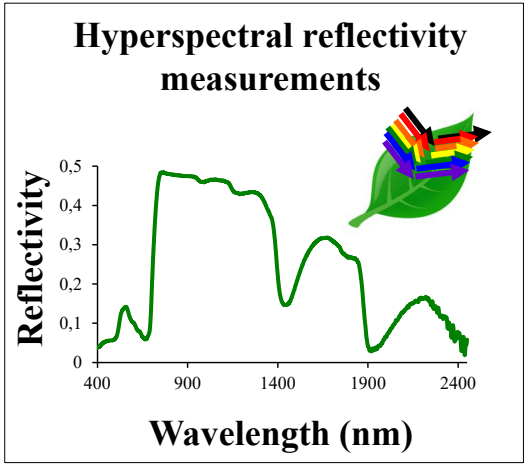
# Outline

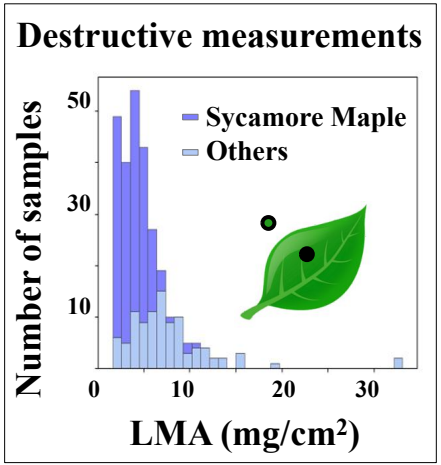
ISEN | école d'ingénieurs
ALL IS DIGITAL!

1. Introduction
2. Material
3. Band selection by means of neural networks
4. Results
5. Conclusions and perspectives

# Material

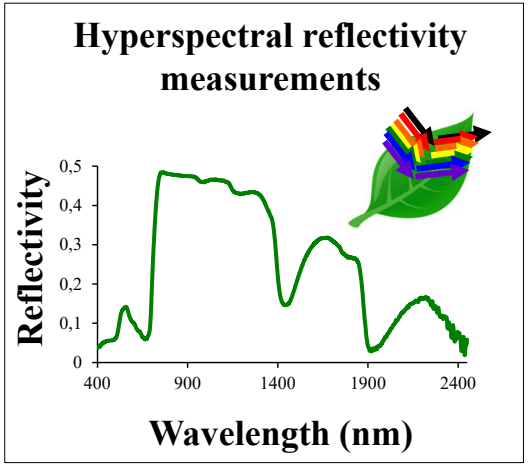
- **Publicly Available *Angers* dataset**
  - ✓ **274 samples**
  - ✓ **2050 hyperspectral bands**

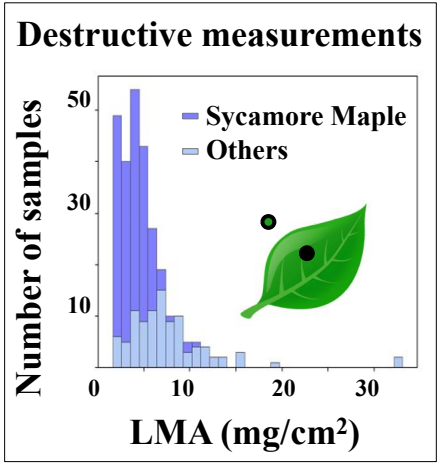


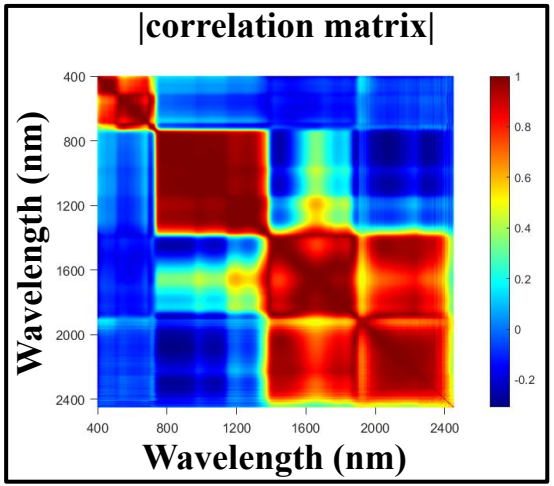
# Material

- **Publicly Available *Angers* dataset**
  - ✓ **274 samples**
  - ✓ **2050 hyperspectral bands**




- **Hyperspectral signal characteristics**



# Why use neural networks for band selection ?

ISEN ALL IS DIGITAL! | yncréa

## Limitations of AI state-of-the-art band selection methods

### ✓ Unsupervised methods

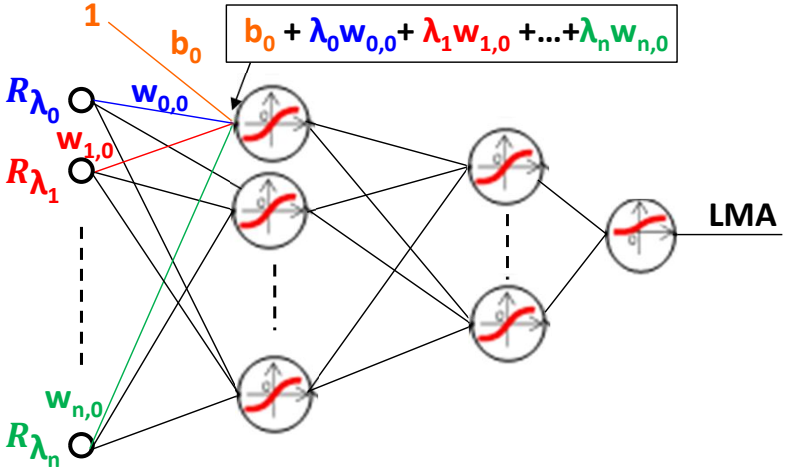
- **Main methods :** Principal Component Analysis, Distance-based methods, Similarity-based methods, Bands Mutual Information-based methods
- **Principle :** Inter-bands relationships investigation ignoring the relashionships between the bands and the output
- Proven efficiency in numerous classification problems
- Poor accuracy in regression problems

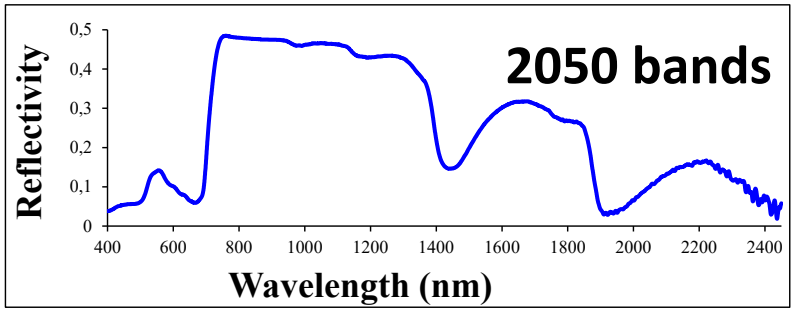
### ✓ Supervised methods

- **Main method :** Band-Output Mutual Information-based method
- **Principle :** Investigation of the relationships between bands and output
- Proven efficiency in classification problems (where the output is the class)
- No straightforward adaptation to regression problems (Binning, Kraskov's method,… )

# Band selection algorithm

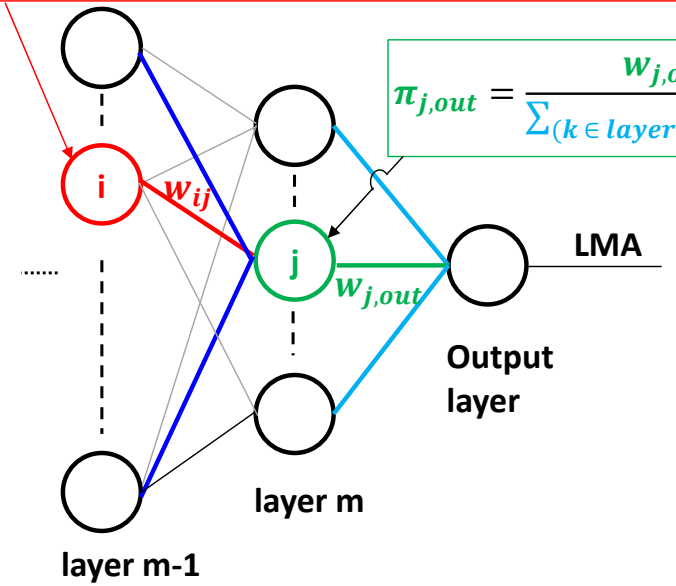
## Step 1: Band weight computation [Yacoub et Bennani, IJNS, 2000]




$$Weight(i) = \sum_{(j \in layer\ 1)} \frac{w_{ij}}{\sum_{(k \in layer\ 0)} w_{kj}}\ \pi_{j,out}$$

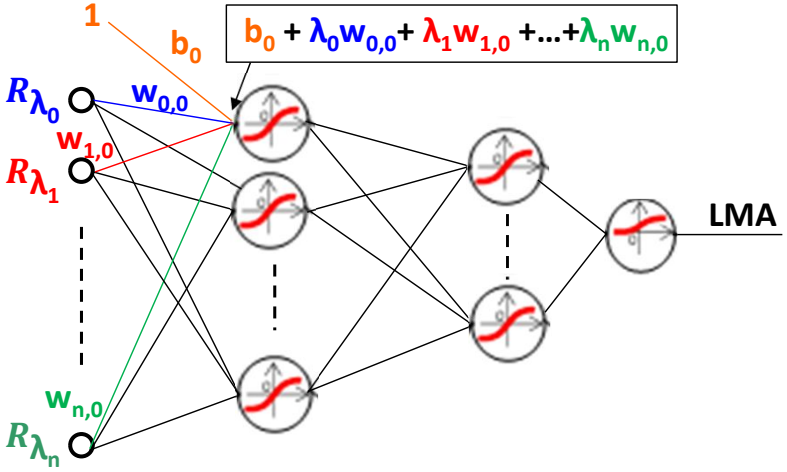
$$\pi_{j,out} = \frac{w_{j,out}}{\sum_{(k \in layer\ 1)} w_{k,out}}$$

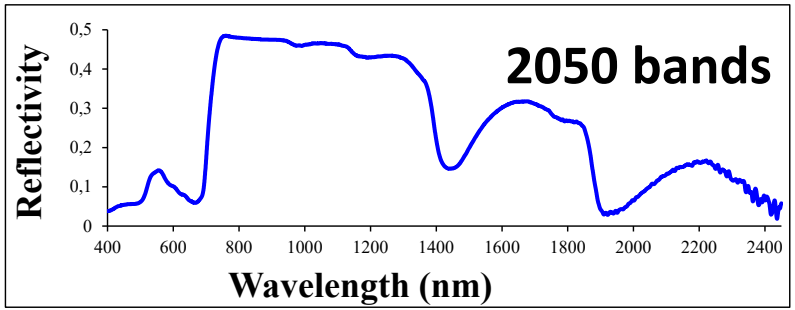


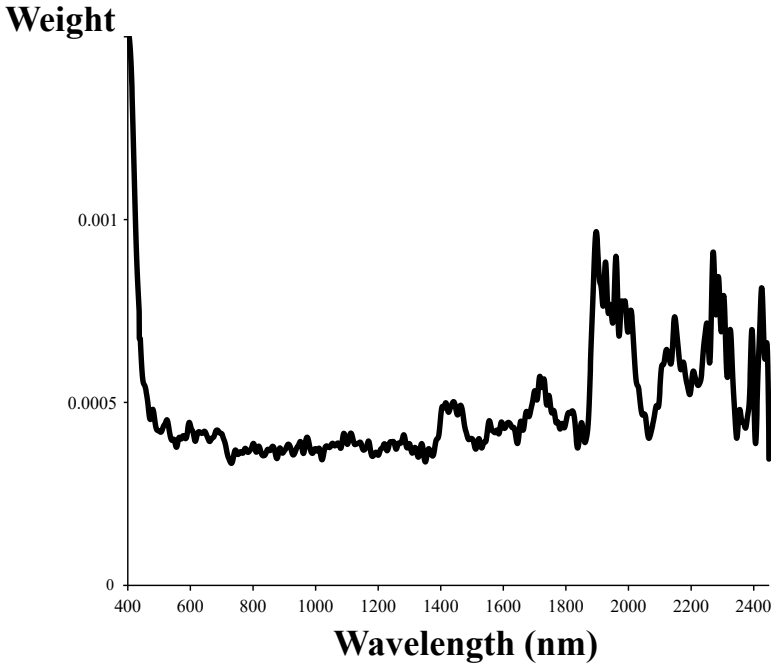

**[Yacoub et Bennani, IJNS, 2000]** M. Yacoub, Y. Bennani, “Features selection and architecture optimization in connectionist systems”, *International Journal of Neural Systems*, **10**, pp. 379-95, 2000.

# Band selection algorithm

## Step 1: Band weight computation [Yacoub et Bennani, IJNS, 2000]



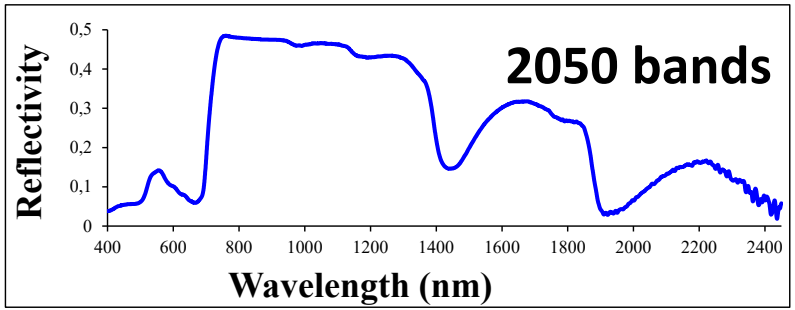



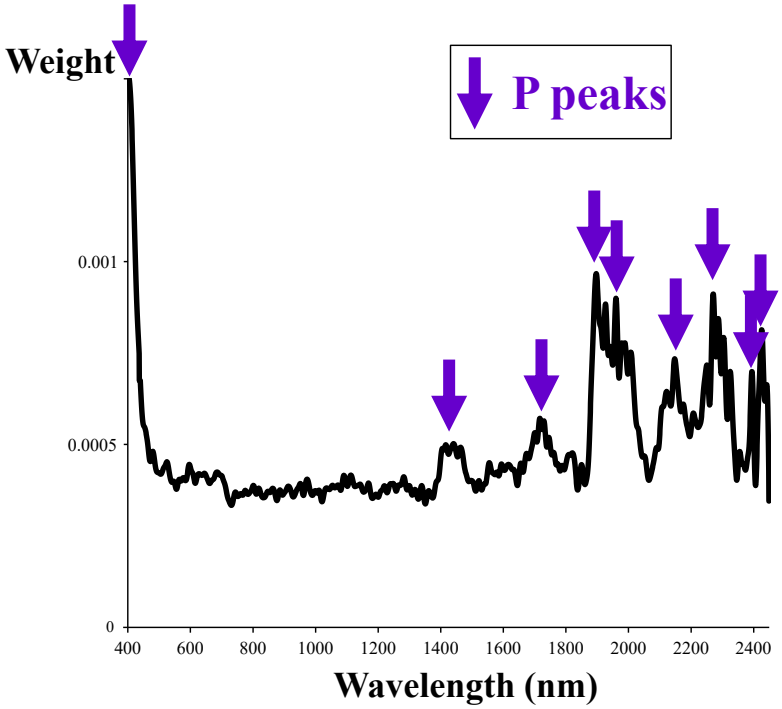



# Band selection algorithm

## Step 1: Band weight computation [Yacoub et Bennani, IJNS, 2000]

## Step 2: P peaks selection

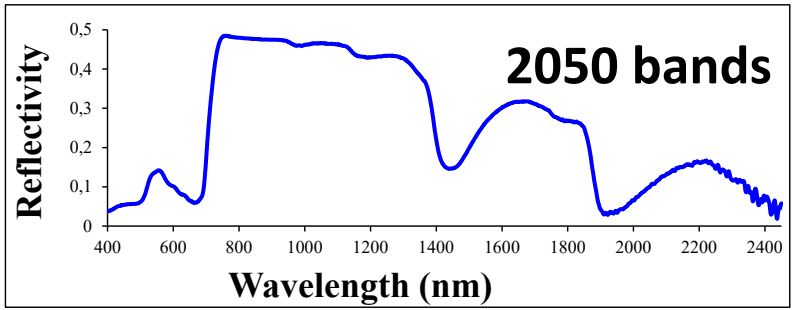


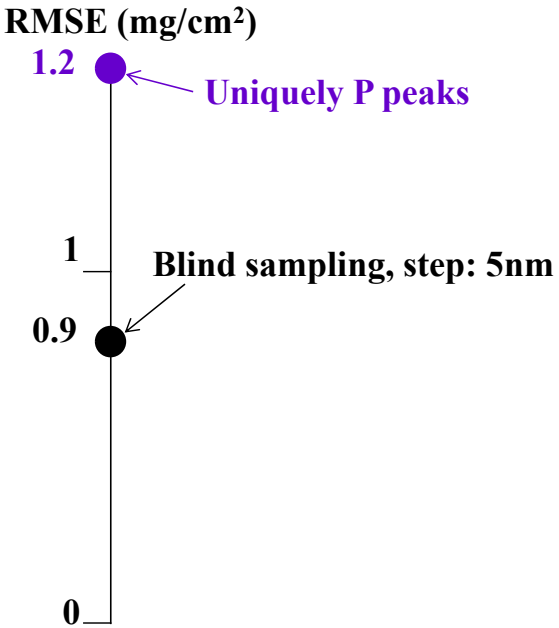



# Band selection algorithm

## Step 1: Band weight computation [Yacoub et Bennani, IJNS, 2000]

## Step 2: P peaks selection



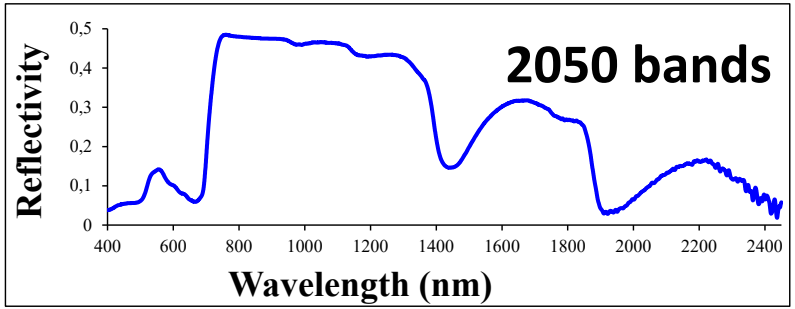
# Band selection algorithm

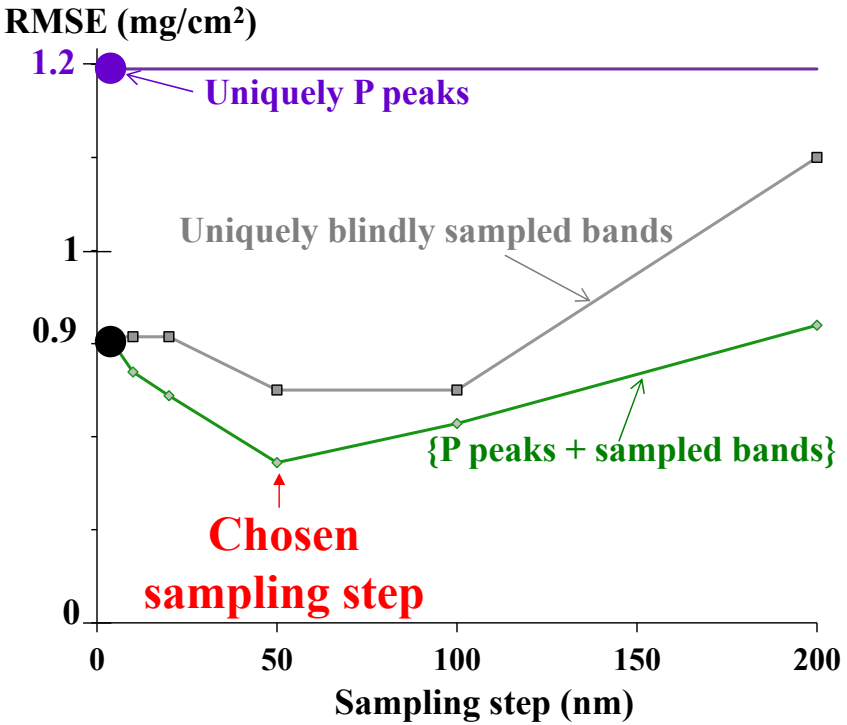
Step 1: Band weight computation [Yacoub et Bennani, IJNS, 2000]

Step 2: P peaks selection

**Step 3: Sampling step choice**

**Congregation of P peaks and sampled bands**
**~50 bands**



# Band selection algorithm

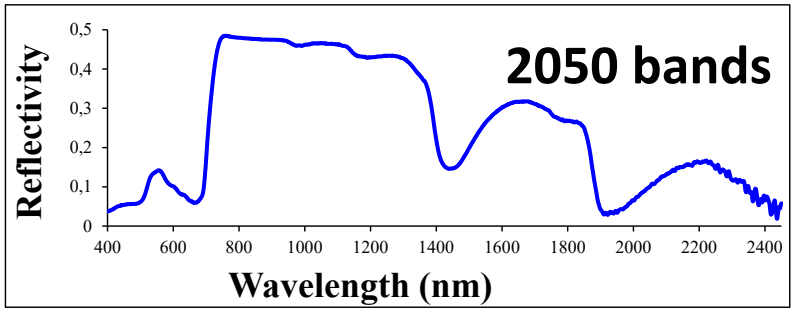
Step 1: Band weight computation [Yacoub et Bennani, IJNS, 2000]

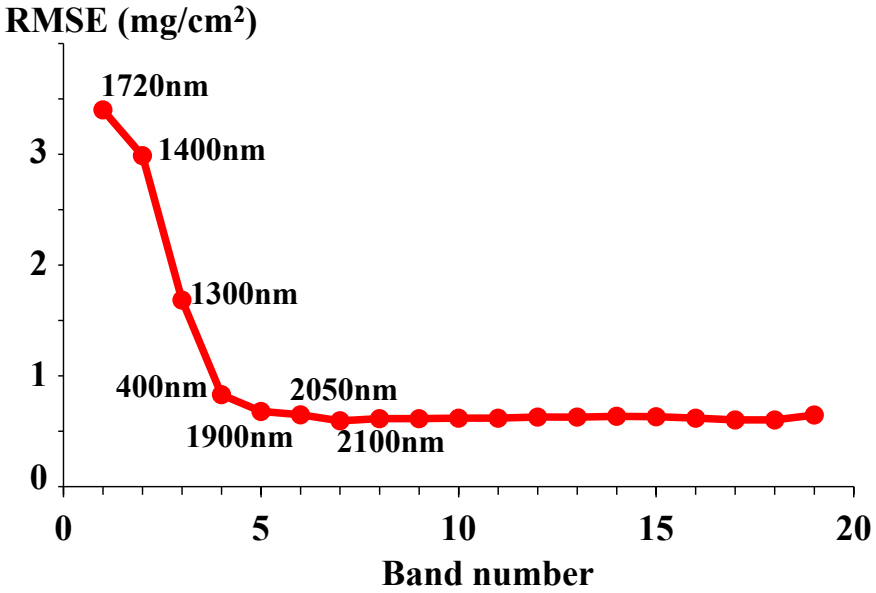
Step 2: P peaks selection

Step 3: Sampling step choice

→ Congregation of P peaks and sampled bands ~50 bands

**Step 4: Backward elimination loop**



# Band selection algorithm

Step 1: Band weight computation [Yacoub et Bennani, IJNS, 2000]

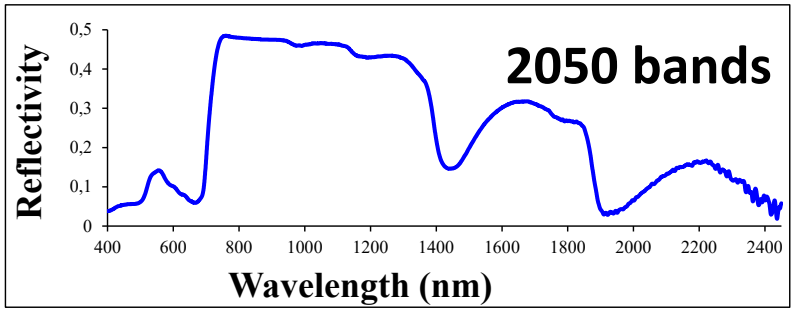
Step 2: P peaks selection

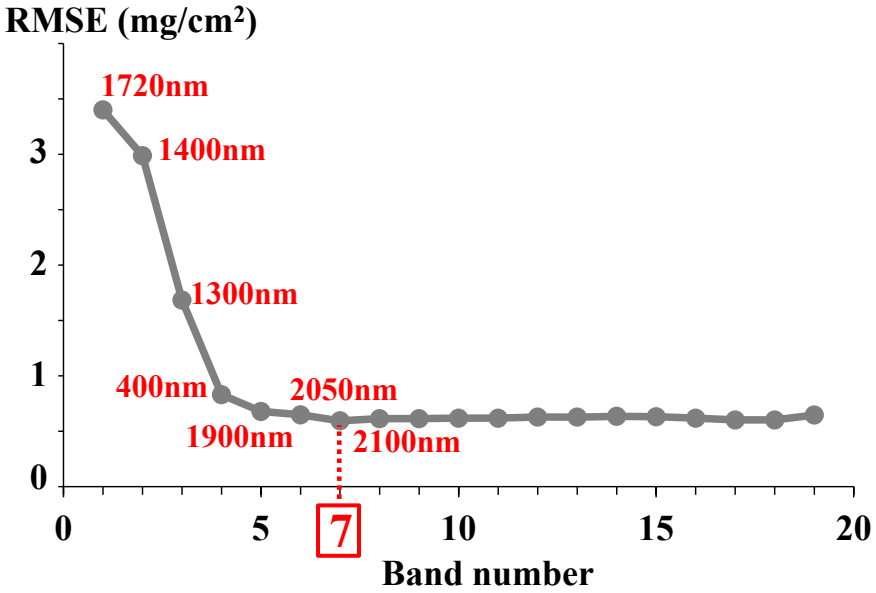
Step 3: Sampling step choice

→ Congregation of P peaks and sampled bands
~50 bands

Step 4: Backward elimination loop

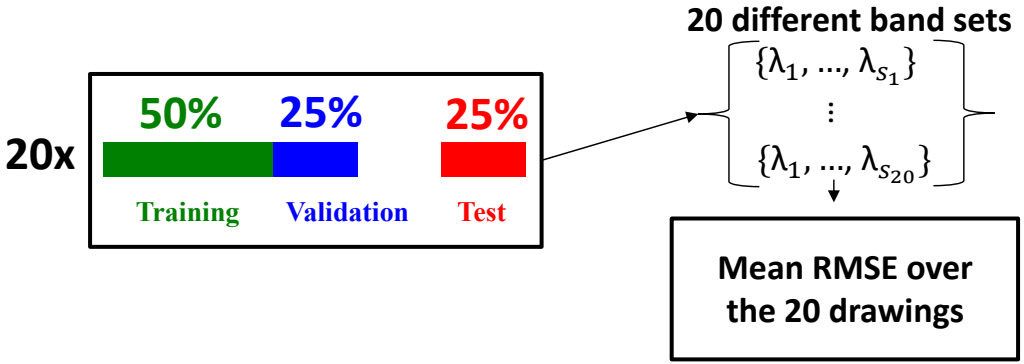
**Step 5: Final band number choice**



# Band selection algorithm testing

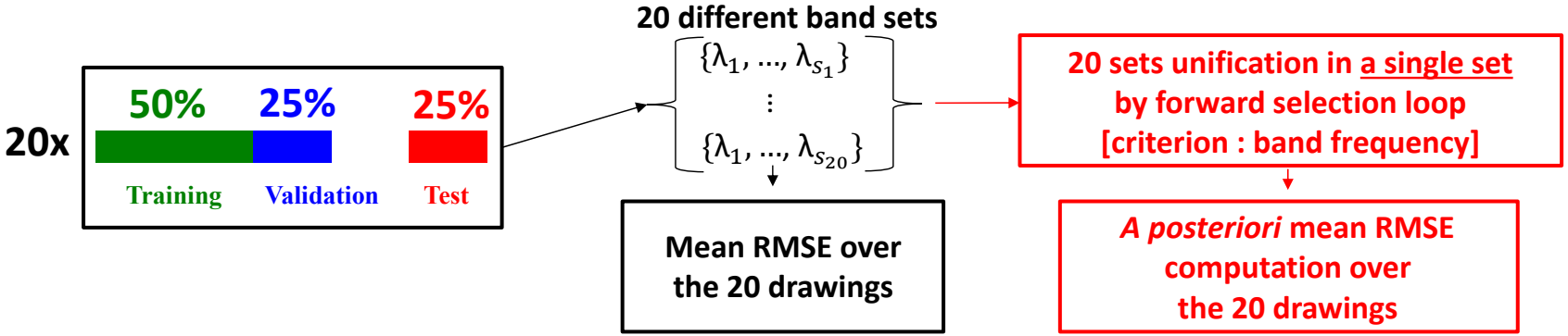
- **Testing protocol:**
  - 20 random drawings in training, validation and test subsets (50%, 25%, 25%)
  - Band selection using uniquely the training and validation subsets (and not the test subset)
  - 20 different ultimate band subsets (1 per drawing)
  - Result : Mean RMSE over the 20 drawings



# Band selection algorithm testing

- **Testing protocol:**
  - 20 random drawings in training, validation and test subsets (50%, 25%, 25%)
  - Band selection using uniquely the training and validation subsets (and not the test subset)
  - 20 different ultimate band subsets (1 per drawing)
  - Result : Mean RMSE over the 20 drawings

- **A posteriori single band set finding**
  - *A posteriori* unique band set determination (band frequency-based forward selection loop)
  - *A posteriori* mean RMSE computation over the 20 drawings (obtained with the unique band set)



# Results: Prediction Accuracy

Mean performances over the 20 drawings

## Best vegetation index

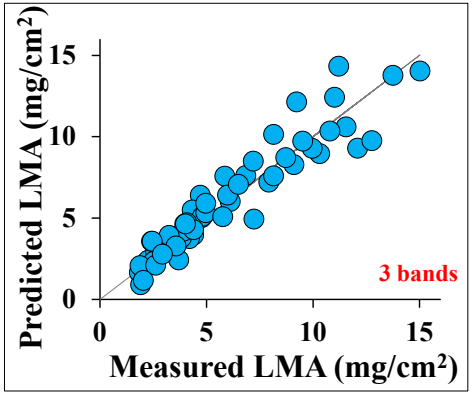
$ND=[R_{1368}-R_{1722}]/[R_{1368}+R_{1722}]$ and $R_{2300}$

**Number of bands: 3**

**Mean Calibration RMSE: 18.6%**

**Mean Test RMSE: 22%**

Graphical comparison for a specific drawing



# Results: Prediction Accuracy

Mean performances over the 20 drawings

## Best vegetation index

$ND=[R_{1368}-R_{1722}]/[R_{1368}+R_{1722}]$ and $R_{2300}$

Number of bands: 3

Mean Calibration RMSE: 18.6%

Mean Test RMSE: 22%

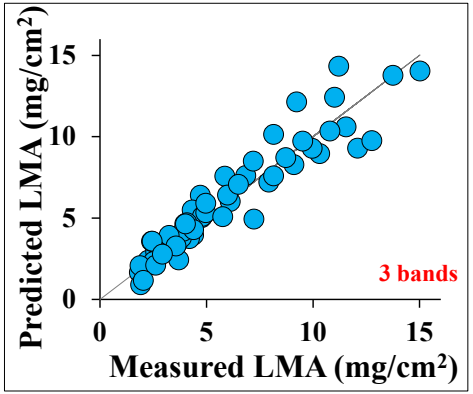
## Neural networks a priori band set

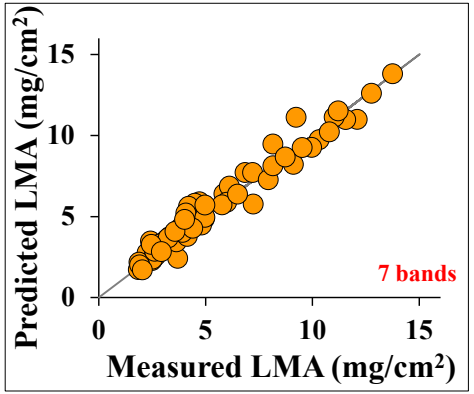
Number of bands: 3-9

Mean Calibration RMSE: 14.8%

Mean Test RMSE: 17.5%

Graphical comparison for a specific drawing



# Results: Prediction Accuracy

Mean performances over the 20 drawings

## Best vegetation index

$ND=[R_{1368}-R_{1722}]/[R_{1368}+R_{1722}]$ and $R_{2300}$

Number of bands: 3
Mean Calibration RMSE: 18.6%
Mean Test RMSE: 22%

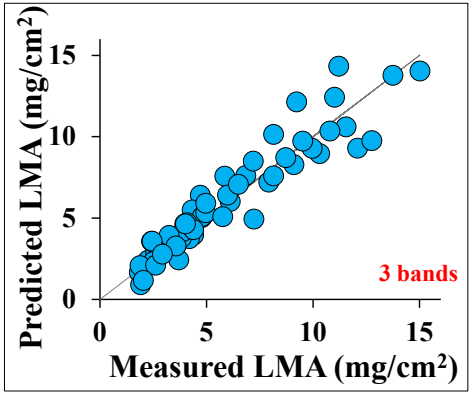
## Neural networks a priori band set

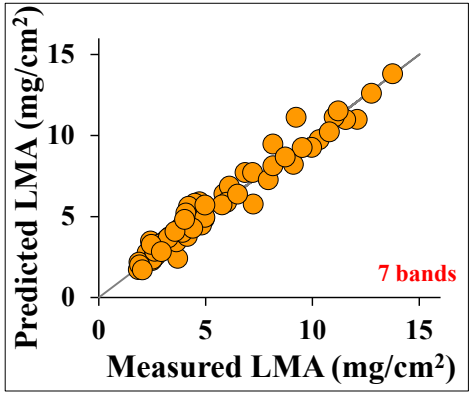
Number of bands: 3-9
Mean Calibration RMSE: 14.8%
Mean Test RMSE: 17.5%

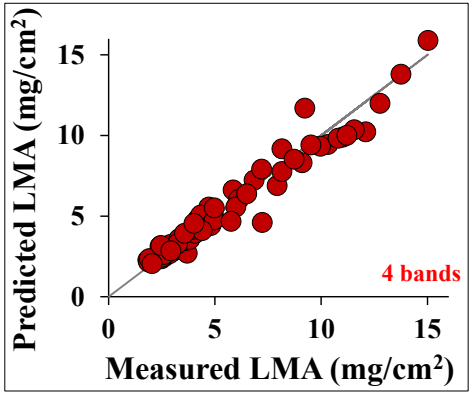
## Neural networks a posteriori band set

Number of bands: 4
Mean Calibration RMSE: 15.2%
Mean Test RMSE: 18.1%

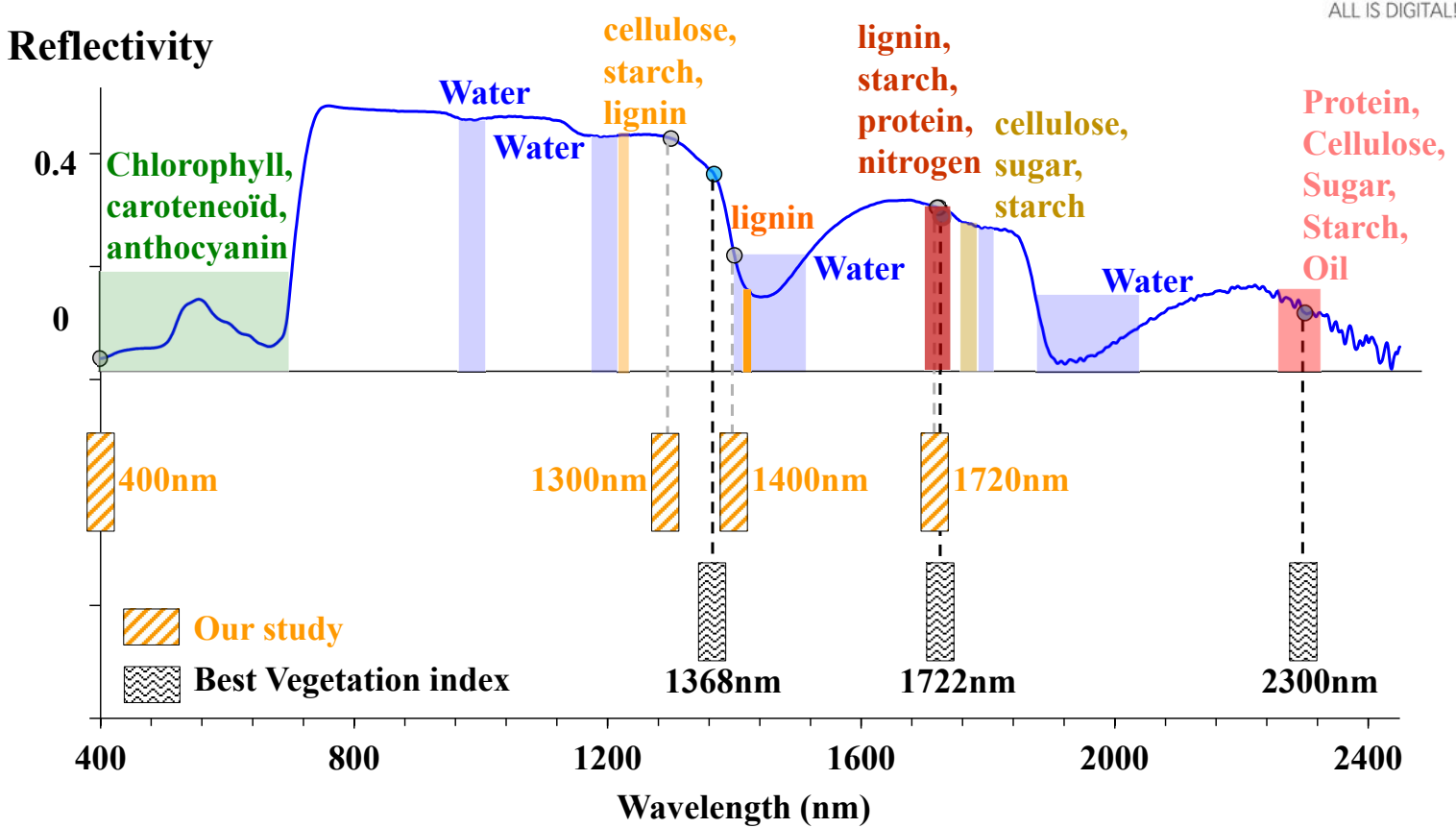
Graphical comparison for a specific drawing



# Results: Selected Band Analysis




*Depiction of, only :
- the main water peaks
- The leaf constituent absorption peaks adjacent to the selected bands

**Band extracted from [S. Jacquemoud and S. Ustin, *Leaf Optical Properties*, Cambridge University Press, Cambridge, 2019].

# Conclusions and prospects

## Conclusions

**Tested on the publicly available *Angers* database, the neural network band selection scheme:**

- **allows to reduce the relative RMSE of LMA prediction by 4.5 points as compared to vegetation indices**
- **while requiring a low number of bands (3 to 9)**
- **and leads to the a posteriori selection of 4 bands for LMA prediction [400nm, 1300nm, 1400nm, 1720nm]**

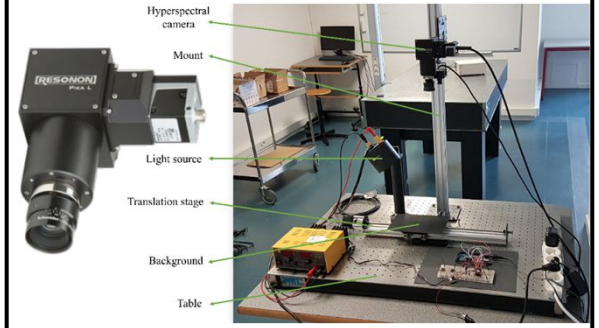
## Prospects

**- Further improvement of LMA prediction accuracy (via the creation of a new experimental database)**

**- Improvement of the prediction accuracy of other leaf parameters (Chlorophyll, Carotenoïd, Anthocyanin, …)**

**- Analytical formulae proposal for replacing neural networks**

**- Application to the detection of Tomato Plant Water Stress**




**ISEN's hyperspectral bench**

# Thank you for your attention !