Monitoring Ecosystems with Remote Sensing

HDR Defense

Mathieu Fauvel

March 29, 2023

CESBIO, Université de Toulouse, CNES/CNRS/INRAe/IRD/UPS, Toulouse, FRANCE

Jury Members

Agnès Bégué	DR	UMR Tetis, CIRAD, France	Reviewer
Lorenzo Bruzzone	Professor	University of Trento, Italy	Reviewer
Florence Tupin	Professor	Telecom Paris, France	Reviewer
Frédérick Garcia	DR	MIA Toulouse, INRAe, France	Examinator
Christian Germain	Professor	Bordeaux Sciences Agro, France	Examinator
Jordi Inglada	Senior Expert	CESBIO, CNES, France	Examinator
Mathias Ortner	Senior Expert	AIRBUS DS, France	Invited

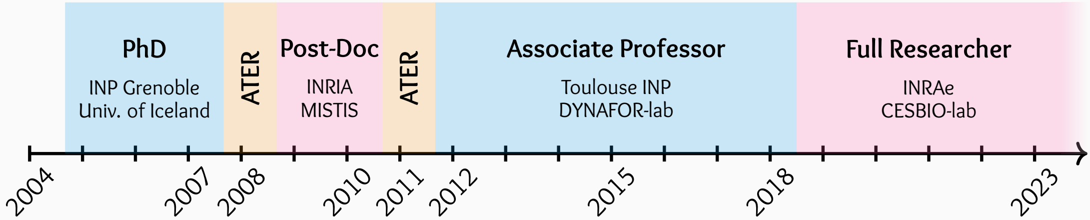
Outline

Introduction

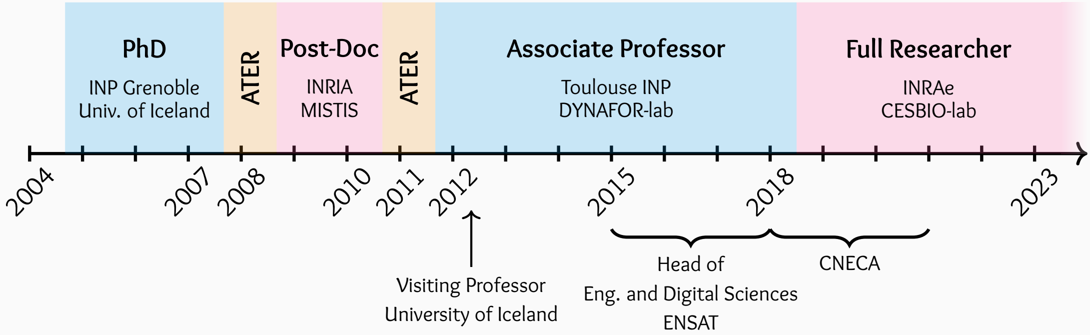
Introduction

Selected Curriculum Vitæ

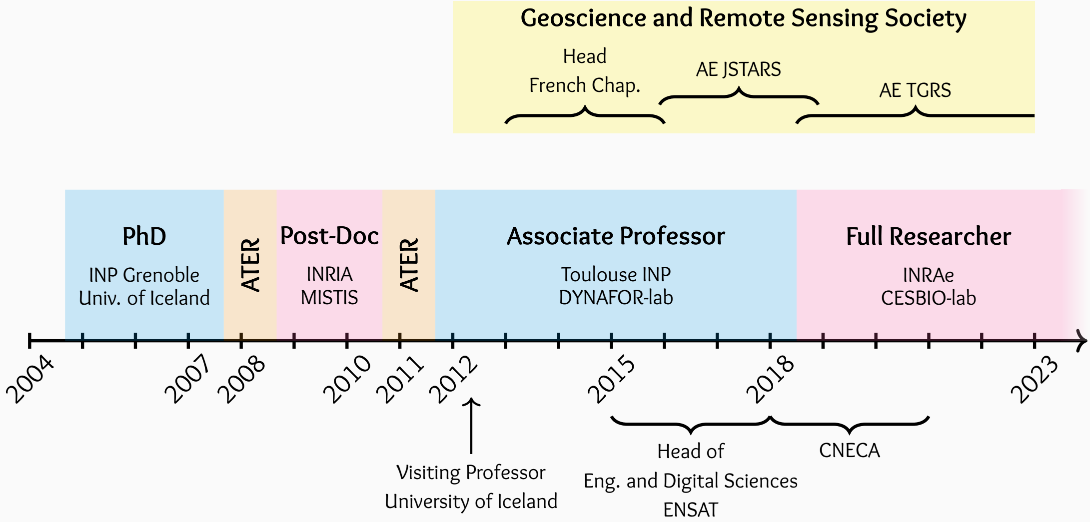
Education & Professional Experience



Education & Professional Experience

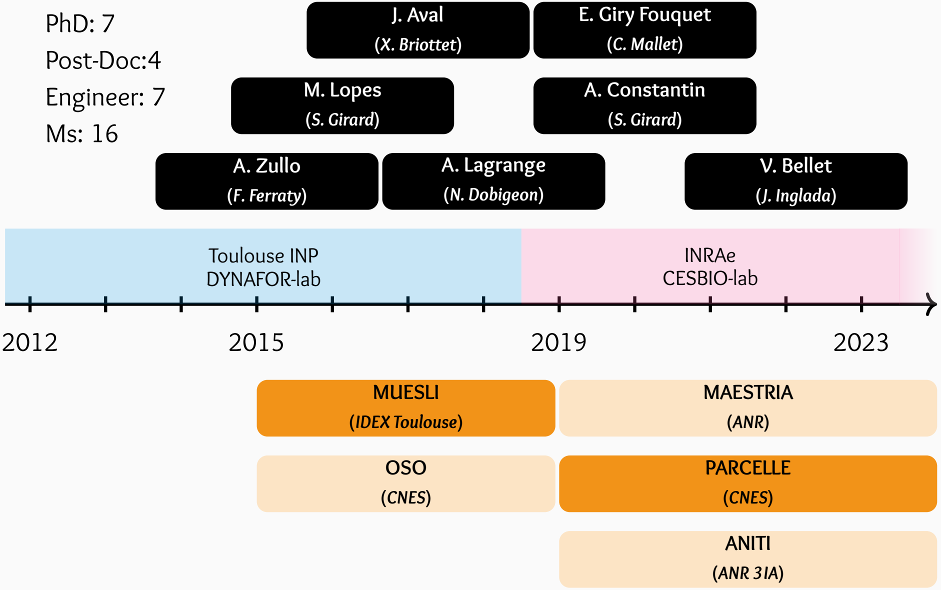


Education & Professional Experience



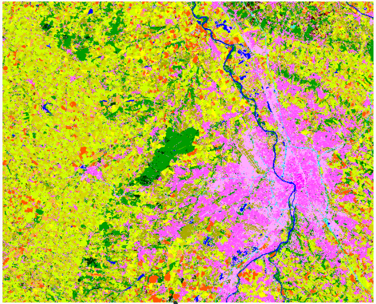
Education & Professional Experience

PhD: 7
Post-Doc:4
Engineer: 7
Ms: 16



Introduction

Scientific Context



$f(\text{pixel's features}) = \text{"Corn"}$

- Prediction: classification, regression, inversion ...
- Supervised learning, unsupervised, semi/self-supervised

Challenges in Machine Learning

- How to define predictive model for **tensorial** data ?

$$\text{Image} \times \text{Spectra} \times \text{Time Series}$$

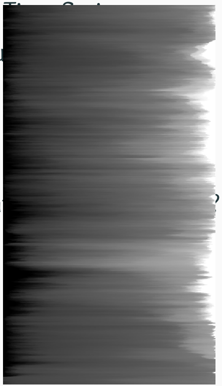
- How to learn with **high number** of **correlated** features ?
  - ⋆ More features than sample
  - ⋆ Numerical instability
  - ⋆ Overfitting
- How to include **multi-source** data in the learning/inference framework ?
- How to deal with **missing values** in features ?

Challenges in Machine Learning

- How to define predictive model for **tensorial** data ?

  ectra ×

- How to learn with **high** **ted** featu
  - More features than
  - Numerical instability
  - Overfitting
- How to include **multi-s** arning/in ?
- How to deal with **missi** s ?



Challenges in Machine Learning

- How to define predictive model for **tensorial** data ?

  Image × Spectra × Time Series

- How to learn with **high number** of **correlated** features ?
  - ⋆ More features than sample
  - ⋆ Numerical instability
  - ⋆ Overfitting
- How to include **multi-source** data in the learning/inference framework ?
- How to deal with **missing values** in features ?

Exploit the **structure** of the data w.r.t the **application**

Quick look of my contributions

Spatial & Spectral Methods - [Hyperspectral, VHR]

- Morphological filters
- Definition of kernel functions
- OBIA

High Dimensional Space - [hyperspectral, SITS]

- Pseudo-distance between spectra
- Feature selection
- Functional data analysis (*A. Zullo*)
- Representation learning (*A. Lagrange*)

Spectro & Temporal Methods - [SITS]

- Irregular & unaligned time series (*A. Constantin*)

Thematic Applications

- Urban trees (*J. Aval*)
- Hedgerows network
- Grasslands (*M. Lopes*)

Quick look of my contributions

Spatial & Spectral Methods - [Hyperspectral, VHR]

- Morphological filters
- Definition of kernel functions
- OBIA

High Dimensional Space - [hyperspectral, SITS]

- Pseudo-distance between spectra
- Feature selection
- Functional data analysis (*A. Zullo*)
- Representation learning (*A. Lagrange*)

Spectro & Temporal Methods - [SITS]

- Irregular & unaligned time series (*A. Constantin*)

Thematic Applications

- Urban trees (*J. Aval*)
- Hedgerows network
- Grasslands (*M. Lopes*)

Scientific productions

Type	#
Journal	48
Chapter	4
Proceeding	53

Quick look of my contributions

Spatial & Spectral Methods - [Hyperspectral, VHR]

- Morphological filters
- Definition of kernel functions
- OBIA

High Dimensional Space - [hyperspectral, SITS]

- Pseudo-distance between spectra
- Feature selection
- **Functional data analysis** (*A. Zullo*)
- Representation learning (*A. Lagrange*)

Spectro & Temporal Methods - [SITS]

- **Irregular & unaligned time series** (*A. Constantin*)

Thematic Applications

- Urban trees (*J. Aval*)
- Hedgerows network
- **Grasslands** (*M. Lopes*)

Scientific productions

Type	#
Journal	48
Chapter	4
Proceeding	53

CESBIO

Spatial & Spectral Methods - [Hypersp… al, VHR]

- Morphological filter…
- Defi…
- OBIA…

Fauvel, Tarabalka, et al., "Advances in Spectral-Spatial Classification of Hyperspectral Images"

High Dimensional Space - [hyperspectral, SITS]

- Pseudo-distance between spectra
- Feature selection
- **Functional data analysis** (*A. Zullo*)
- Representation learning (*A. Lagrange*)

Spectro & Temporal Methods - [SITS]

- **Irregular & unaligned time series** (*A. Constantin*)

Thematic Applications

- Urban trees (*J. Aval*)
- Hedgerows network
- **Grasslands** (*M. Lopes*)

Scientific productions

Type	#
Journal	48
Chapter	4
Proceeding	53

Selected Scientific Contributions

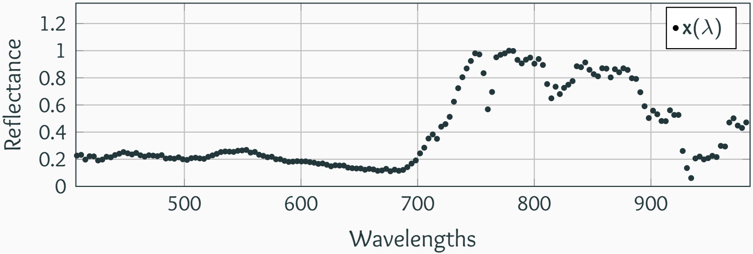
Selected Scientific Contributions

Functional Data Analysis For High Dimensional Remote Sensing Data

Functional Data Analysis

Anthony Zullo, "Analyse de données fonctionnelles en télédétection hyperspectrale : application à l'étude des paysages agri-forestiers"

Functional Data Analysis



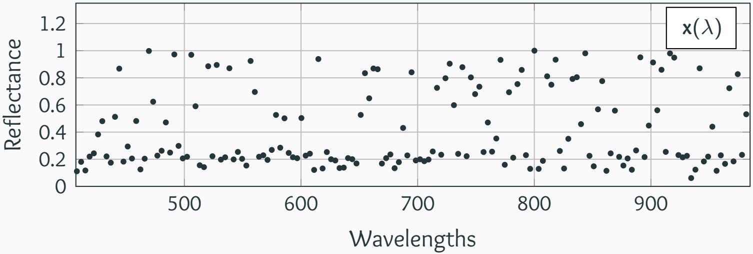
Random vector

- $\mathbf{x} = [x_{\lambda_1}, \ldots, x_{\lambda_d}] \in \mathbb{R}^d$
- Invariant to random permutation

Random curves

- $\chi = \{\chi(\lambda), \lambda \in [\lambda_1, \lambda_d]\} \in \mathcal{F}$
- Ordering relation
- Integrate curves properties (derivative, smoothness)

Functional Data Analysis



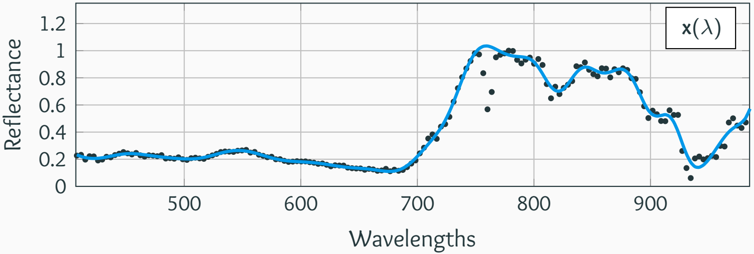
Random vector

- $\mathbf{x} = [x_{\lambda_1}, \ldots, x_{\lambda_d}] \in \mathbb{R}^d$
- Invariant to random permutation

Random curves

- $\chi = \{\chi(\lambda), \lambda \in [\lambda_1, \lambda_d]\} \in \mathcal{F}$
- Ordering relation
- Integrate curves properties (derivative, smoothness)

Functional Data Analysis



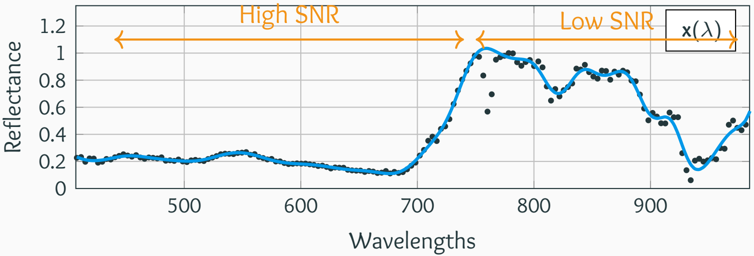
Random vector

- $\mathbf{x} = [x_{\lambda_1}, \ldots, x_{\lambda_d}] \in \mathbb{R}^d$
- Invariant to random permutation

Random curves

- $\chi = \{\chi(\lambda), \lambda \in [\lambda_1, \lambda_d]\} \in \mathcal{F}$
- Ordering relation
- Integrate curves properties (derivative, smoothness)

Functional Data Analysis

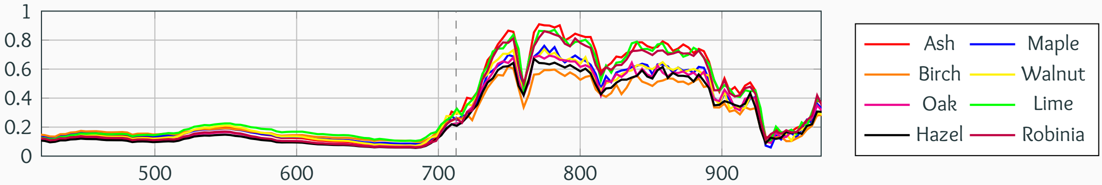


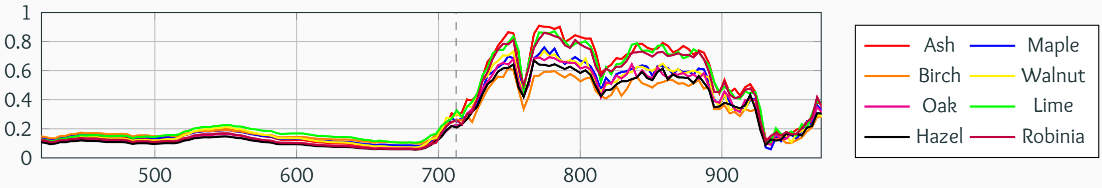
Random vector

- $\mathbf{x} = [x_{\lambda_1}, \ldots, x_{\lambda_d}] \in \mathbb{R}^d$
- Invariant to random permutation

Random curves

- $\chi = \{\chi(\lambda), \lambda \in [\lambda_1, \lambda_d]\} \in \mathcal{F}$
- Ordering relation
- Integrate curves properties (derivative, smoothness)


1
0.8
0.6
0.4
0.2
0
500
600
700
800
900
Ash
Maple
Birch
Walnut
Oak
Lime
Hazel
Robinia



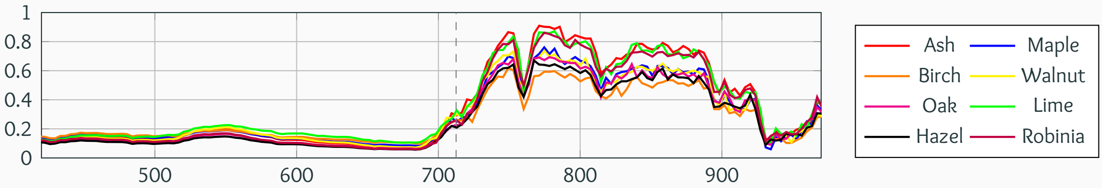
Heteroscedeastic noise

$$\mathbf{x}_i(\lambda) \xrightarrow{g} \tilde{\mathbf{x}}_i(\lambda) \xrightarrow{f} \hat{Y}$$

$$\tilde{\mathbf{x}}_i(\lambda) = \sum_{j=1}^{d} a_{\lambda_j}(\lambda)\mathbf{x}_i(\lambda_j)$$

Ferraty, A. Zullo, and Fauvel, "Nonparametric regression on contaminated functional predictor with application to hyperspectral data"

Classification of Hyperspectral Data



Heteroscedeastic noise

$$\mathbf{x}_i(\lambda) \xrightarrow{g} \tilde{\mathbf{x}}_i(\lambda) \xrightarrow{f} \hat{Y}$$

$$\tilde{\mathbf{x}}_i(\lambda) = \sum_{j=1}^{d} a_{\lambda_j}(\lambda)\mathbf{x}_i(\lambda_j)$$

Ferraty, A. Zullo, and Fauvel, "Nonparametric regression on contaminated functional predictor with application to hyperspectral data"

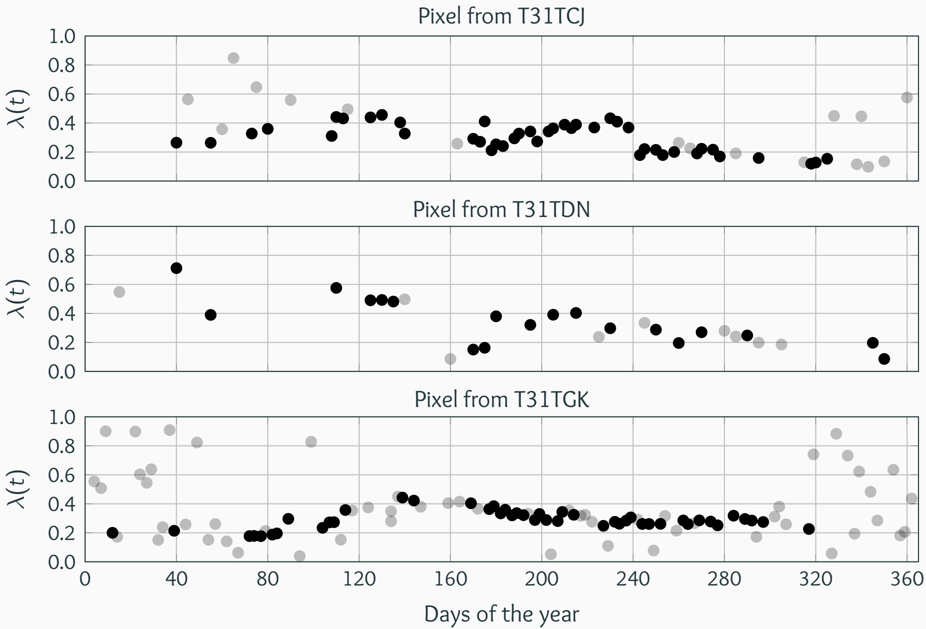
Classification results

	Error rate	std
NPFE	10.5	1.0
GMM	14.9	1.3
SVM	13.8	1.4
RF	18.5	1.2

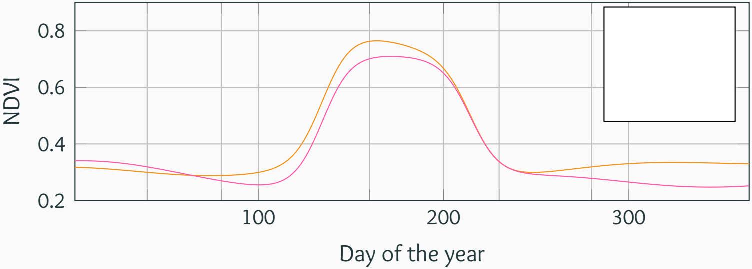
Selected Scientific Contributions

Gaussian Process For Irregular And Unaligned Time Series

Irregular and Unaligned SITS



Gaussian Process

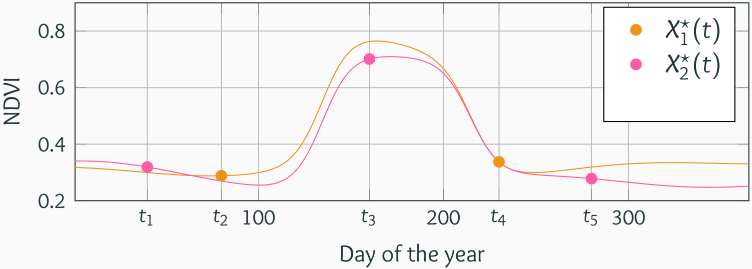


Definition

A Gaussian process (GP) is a stochastic process such that any finite-dimensional marginal follows a multivariate Gaussian distribution.

$$X \sim \mathcal{GP}(m, K), \; X(t) \in \mathbb{R}$$

Alexandre Constantin, "Analyse de séries temporelles massives d'images satellitaires : Applications à la cartographie des écosystèmes"

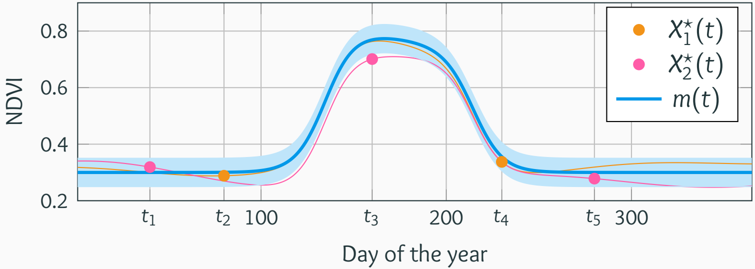


Definition

A Gaussian process (GP) is a stochastic process such that any finite-dimensional marginal follows a multivariate Gaussian distribution.

$$X \sim \mathcal{GP}(m, K),\ X(t) \in \mathbb{R}$$

$$\begin{bmatrix} X_1^\star(t_1) \\ X_1^\star(t_3) \\ X_1^\star(t_5) \end{bmatrix} \sim \mathcal{N}_3 \left(\begin{bmatrix} m(t_1) \\ m(t_3) \\ m(t_5) \end{bmatrix}, \begin{bmatrix} K(t_1, t_1) & \dots & K(t_1, t_5) \\ \vdots & \ddots & \vdots \\ K(t_5, t_1) & \dots & K(t_5, t_5) \end{bmatrix} \right) \begin{bmatrix} X_2^\star(t_2) \\ X_2^\star(t_4) \end{bmatrix} \sim \mathcal{N}_2 \left(\begin{bmatrix} m(t_2) \\ m(t_4) \end{bmatrix}, \begin{bmatrix} K(t_2, t_2) & K(t_2, t_4) \\ K(t_2, t_4) & K(t_4, t_4) \end{bmatrix} \right)$$



Definition

A Gaussian process (GP) is a stochastic process such that any finite-dimensional marginal follows a multivariate Gaussian distribution.

$$X \sim \mathcal{GP}(m, K),\ X(t) \in \mathbb{R}$$

$$\begin{bmatrix} X_1^\star(t_1) \\ X_1^\star(t_3) \\ X_1^\star(t_5) \end{bmatrix} \sim \mathcal{N}_3\left(\begin{bmatrix} m(t_1) \\ m(t_3) \\ m(t_5) \end{bmatrix}, \begin{bmatrix} K(t_1, t_1) & \dots & K(t_1, t_5) \\ \vdots & \ddots & \vdots \\ K(t_5, t_1) & \dots & K(t_5, t_5) \end{bmatrix} \right) \begin{bmatrix} X_2^\star(t_2) \\ X_2^\star(t_4) \end{bmatrix} \sim \mathcal{N}_2\left(\begin{bmatrix} m(t_2) \\ m(t_4) \end{bmatrix}, \begin{bmatrix} K(t_2, t_2) & K(t_2, t_4) \\ K(t_2, t_4) & K(t_4, t_4) \end{bmatrix} \right)$$

Multivariate Gaussian Processes (MGP) for SITS classification

Definition

A Multivariate Gaussian Processes (MGP), conditionally to $Z = c$, is defined as

$$\mathbf{X}(t) = \mathbf{A}_c\mathbf{W}_c(t) + \mathbf{m}_c(t) \text{ with } \mathcal{W}_{cb} | \; Z = c \sim \mathcal{GP}(0, K_c)$$

A. Constantin, Fauvel, and Girard, "Mixture of multivariate Gaussian processes for classification of irregularly sampled satellite image time-series"

Multivariate Gaussian Processes (MGP) for SITS classification

Definition

A Multivariate Gaussian Processes (MGP), conditionally to $Z = c$, is defined as

$$\mathbf{X}(t) = \mathbf{A}_c \mathbf{W}_c(t) + \mathbf{m}_c(t) \text{ with } \mathcal{W}_{cb} | Z = c \sim \mathcal{GP}(0, K_c)$$

$$\begin{pmatrix} \mathbf{x}^\star_{1,1} & \mathbf{x}^\star_{1,2} & \cdots & \mathbf{x}^\star_{1,l} & \cdots & \mathbf{x}^\star_{1,q} \\ \mathbf{x}^\star_{2,1} & \mathbf{x}^\star_{2,2} & \cdots & \mathbf{x}^\star_{2,l} & \cdots & \mathbf{x}^\star_{2,q} \\ \vdots & \vdots & & \vdots & & \vdots \\ \mathbf{x}^\star_{k,1} & \mathbf{x}^\star_{k,2} & \cdots & \mathbf{x}^\star_{k,l} & \cdots & \mathbf{x}^\star_{k,q} \\ \vdots & \vdots & & \vdots & & \vdots \\ \mathbf{x}^\star_{p,1} & \mathbf{x}^\star_{p,2} & \cdots & \mathbf{x}^\star_{p,l} & \cdots & \mathbf{x}^\star_{p,q} \end{pmatrix}$$

$(\mathbf{X}^\star)_l \sim \mathcal{N}_p(\,.\,, \mathbf{\Lambda})$

$(\mathbf{X}^\star)_k^\top \sim \mathcal{N}_q(\,.\,, \mathbf{\Sigma})$

$$\boldsymbol{\lambda}(t) = \begin{bmatrix} \lambda_1(t) \\ \vdots \\ \lambda_p(t) \end{bmatrix}$$

Definition

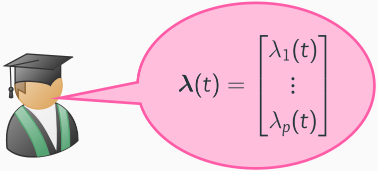
A Multivariate Gaussian Processes (MGP), conditionally to $Z = c$, is defined as

$$\mathbf{X}(t) = \mathbf{A}_c \mathbf{W}_c(t) + \mathbf{m}_c(t) \text{ with } \mathcal{W}_{cb} \mid Z = c \sim \mathcal{GP}(0, K_c)$$

$$\begin{pmatrix} \mathbf{X}^\star_{1,1} & \mathbf{X}^\star_{1,2} & \cdots & \mathbf{X}^\star_{1,l} & \cdots & \mathbf{X}^\star_{1,q} \\ \mathbf{X}^\star_{2,1} & \mathbf{X}^\star_{2,2} & \cdots & \mathbf{X}^\star_{2,l} & \cdots & \mathbf{X}^\star_{2,q} \\ \vdots & \vdots & & \vdots & & \vdots \\ \mathbf{X}^\star_{k,1} & \mathbf{X}^\star_{k,2} & \cdots & \mathbf{X}^\star_{k,l} & \cdots & \mathbf{X}^\star_{k,q} \\ \vdots & \vdots & & \vdots & & \vdots \\ \mathbf{X}^\star_{p,1} & \mathbf{X}^\star_{p,2} & \cdots & \mathbf{X}^\star_{p,l} & \cdots & \mathbf{X}^\star_{p,q} \end{pmatrix}$$

$(\mathbf{X}^\star)_l \sim \mathcal{N}_p(. , \mathbf{\Lambda})$

$(\mathbf{X}^\star)_k^\top \sim \mathcal{N}_q(. , \mathbf{\Sigma})$



- Negative log likelihood for class c

$$p \sum_{i|Z^i=c} \log |\mathbf{\Sigma}_i(\boldsymbol{\theta}_c)| + q_i \log |\mathbf{A}_c \mathbf{A}_c^\top| + \operatorname{tr}\left[\sum_{i|Z_i=c} (\mathbf{X}_i^\star - \boldsymbol{\alpha}_c \mathbf{B}_i)\{\mathbf{\Sigma}_i(\boldsymbol{\theta}_c)\}^{-1}(\mathbf{X}_i^\star - \boldsymbol{\alpha}_c \mathbf{B}_i)^\top \{\mathbf{A}_c \mathbf{A}_c^\top\}^{-1}\right] + \kappa$$

- Classification of sample $\mathbf{X}_u^\star$ observed at timestamps T_u **and** imputation of time $\tilde{t}_u$

S2 Land Cover Classification

- 190,000 learning pixels for 3 tiles
- 14 land cover classes
- Non-parametric models RF
- Parametric model QDA

	Average F1	std F1
M2GP	63.1	1.15
QDA	64.2	1.36
RF	**74.2**	1.78

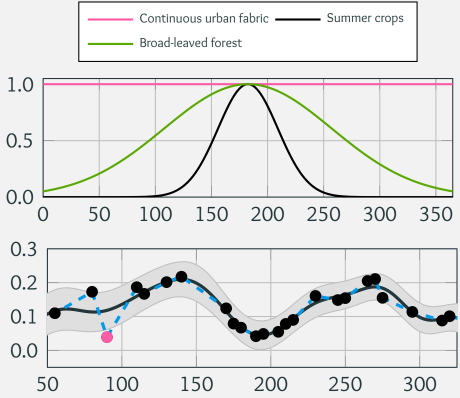
Classification and imputation of irregular and unaligned SITS

S2 Land Cover Classification

- 190,000 learning pixels for 3 tiles
- 14 land cover classes
- Non-parametric models RF
- Parametric model QDA

	Average F1	std F1
M2GP	63.1	1.15
QDA	64.2	1.36
RF	**74.2**	1.78

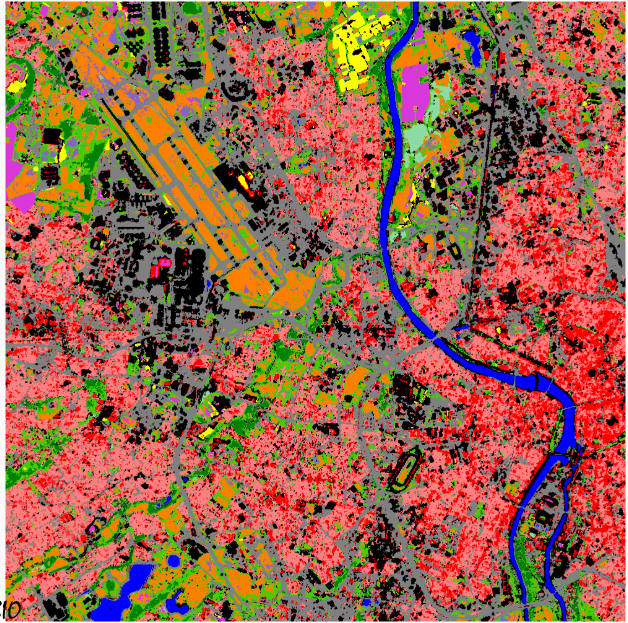
By-product



Classification and imputation of irregular and unaligned SITS

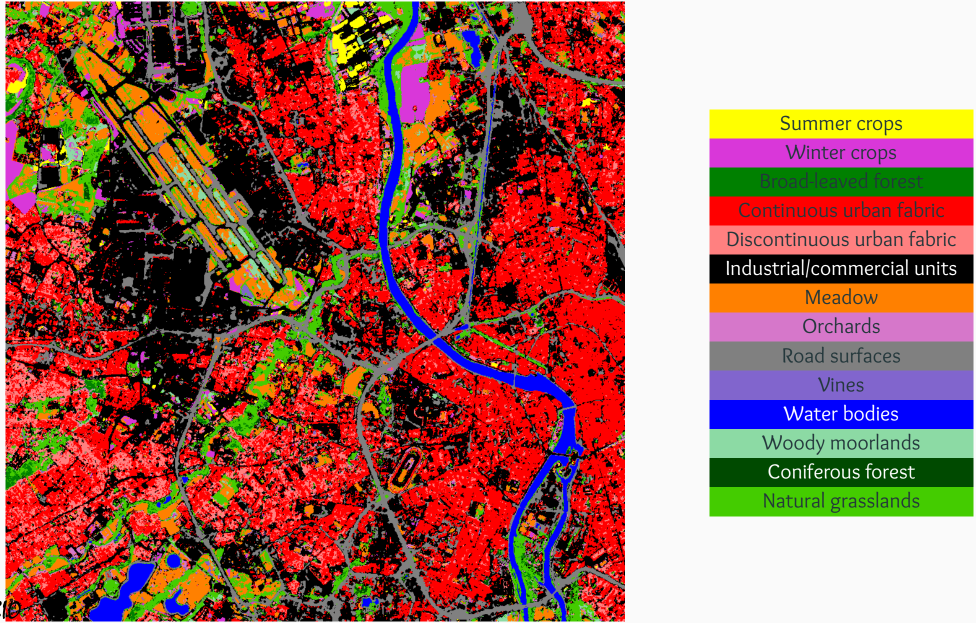
Summer crops
Winter crops
Broad-leaved forest
Continuous urban fabric
Discontinuous urban fabric
Industrial/commercial units
Meadow
Orchards
Road surfaces
Vines
Water bodies
Woody moorlands
Coniferous forest
Natural grasslands

Classification and imputation of irregular and unaligned SITS



- Summer crops
- Winter crops
- Broad-leaved forest
- Continuous urban fabric
- Discontinuous urban fabric
- Industrial/commercial units
- Meadow
- Orchards
- Road surfaces
- Vines
- Water bodies
- Woody moorlands
- Coniferous forest
- Natural grasslands

Classification and imputation of irregular and unaligned SITS



Selected Scientific Contributions

Grassland plant taxonomic diversity estimation

State of the art (as of 2017)

- Plant diversity:
  - ✗ MSR (e.g., MODIS)
  - ✓ VHSR (e.g., Rapid Eye)
  - ✓ Aerial hyperspectral data
- Limited in time and space !

Monitoring (European) Grasslands Diversity From Space

State of the art (as of 2017)

- Plant diversity:
  - ✗ MSR (e.g., MODIS)
  - ✓ VHSR (e.g., Rapid Eye)
  - ✓ Aerial hyperspectral data
- Limited in time and space !
- Can Sentinel 1&2 help to monitor diversity ?

Plant diversity at scale

- Maïlys Lopes, "Ecological monitoring of semi-natural grasslands : statistical analysis of dense satellite image time series with high spatial resolution"

✓ Old/Young grassland classification

- Go beyond:
  - ⋆ Pixel wise prediction
  - ⋆ Joint use S1/S2

Monitoring (European) Grasslands Diversity From Space

State of the art (as of 2017)

- Plant diversity:
  - ✗ MSR (e.g., MODIS)
  - ✓ VHSR (e.g., Rapid Eye)
  - ✓ Aerial hyperspectral data
- Limited in time and space !
- Can Sentinel 1&2 help to monitor diversity ?

Plant diversity at scale

- Maïlys Lopes, "Ecological monitoring of semi-natural grasslands : statistical analysis of dense satellite image time series with high spatial resolution"
- ✓ Old/Young grassland classification
- Go beyond:
  - ⋆ Pixel wise prediction
  - ⋆ Joint use S1/S2

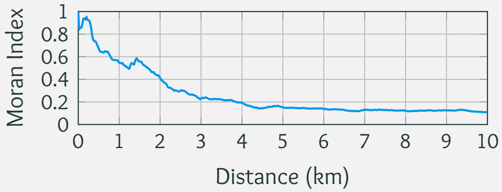
Objectives

- Predict diversity index: *Shannon* and *Simpson*
- Using all S1 & S2 from 2017-08 to 2018-12

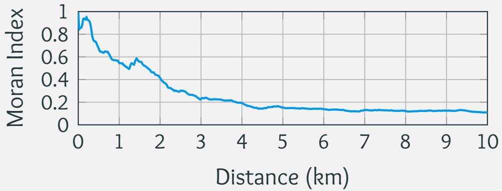
Data	S2	NDVI	NDVI-S2	PCA-S2	R-IR	S1	S1-S2	PCA-S1-S2
Size	188	47	235	49	94	279	467	117

Fauvel, M. Lopes, et al., "Prediction of plant diversity in grasslands using Sentinel-1 and -2 satellite image time series"

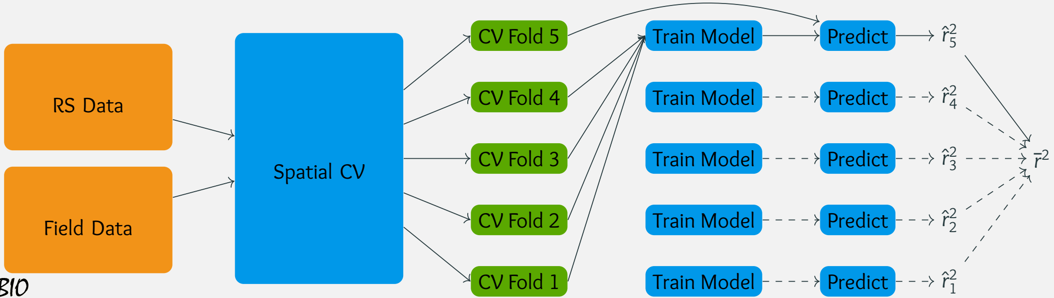
Spatial auto-correlation of the diversity index



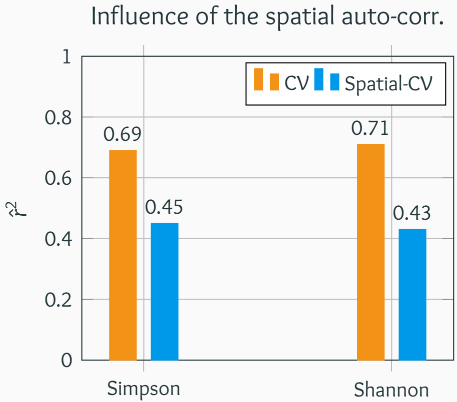
Spatial auto-correlation of the diversity index

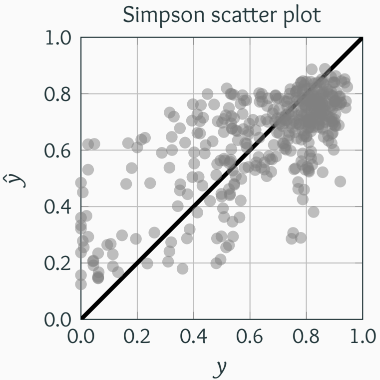


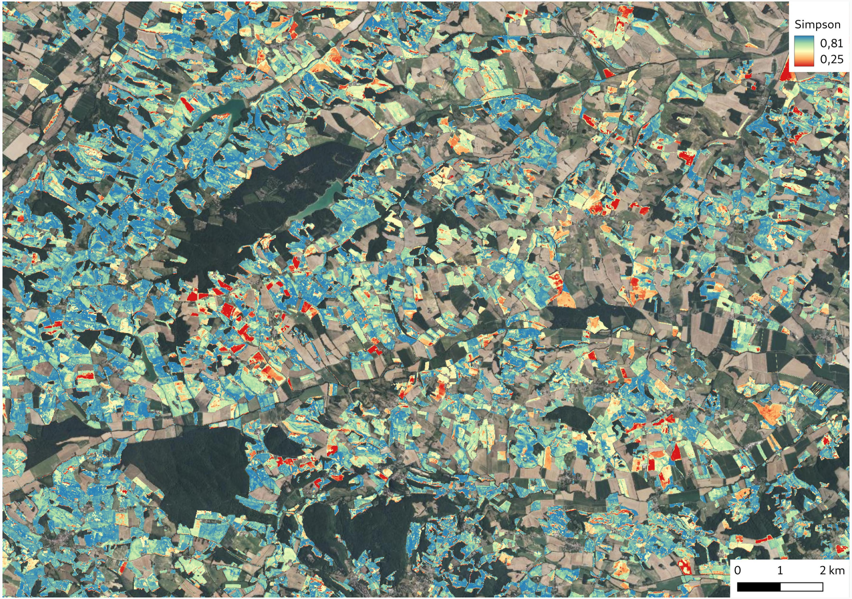
Estimation of the accuracy



Results obtained with R-IR bands only, using RF






Simpson
0,81
0,25
0
1
2 km

Perspectives

Monitoring State and Trend of Ecosystem

Objectives

- State: composition and configuration
- Trend: vegetation dynamic index
- Various vegetation type: crops, grasslands, forests …
- New practices: intercroping …
- Reduce dependence on ground truth

Constraint

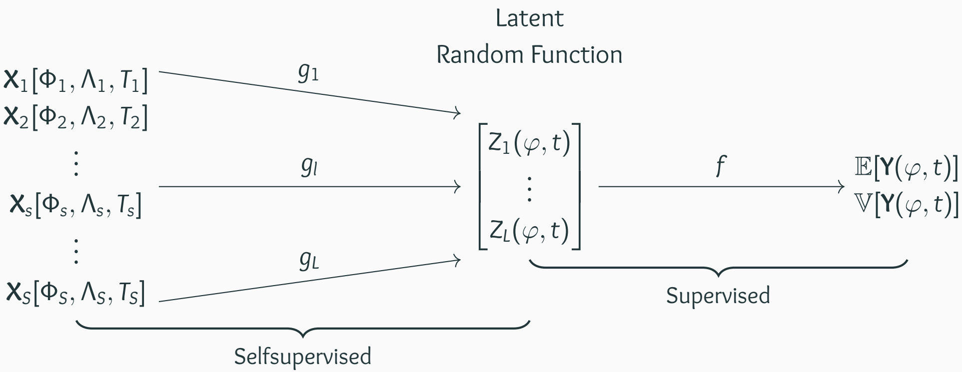
- Training
  - ⋆ Semi-/Weakly-/Self-supervised
  - ⋆ Scale w.r.t the number of sample
  - ⋆ Irregular sampled multi-resolution data
- Inference
  - ⋆ Function over space and time
  - ⋆ With *confidence*

Vegetation Dynamic

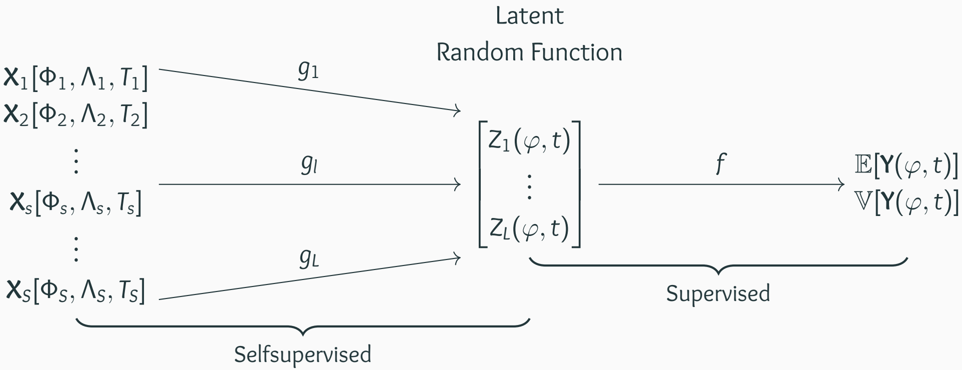
- Applications

  - ⋆ Plant/Animal diversity
  - ⋆ Carbon storage
  - ⋆ Adaptation to climate change
- Current issues
  - ⋆ Coarse resolution
  - ⋆ Unadapted model (e.g, double logistic)
  - ⋆ Based on NDVI/LAI - derived spectral index
  - ⋆ Limited use of S1 w.r.t S2
- CES Theia - Variables for biodiversity

Framework

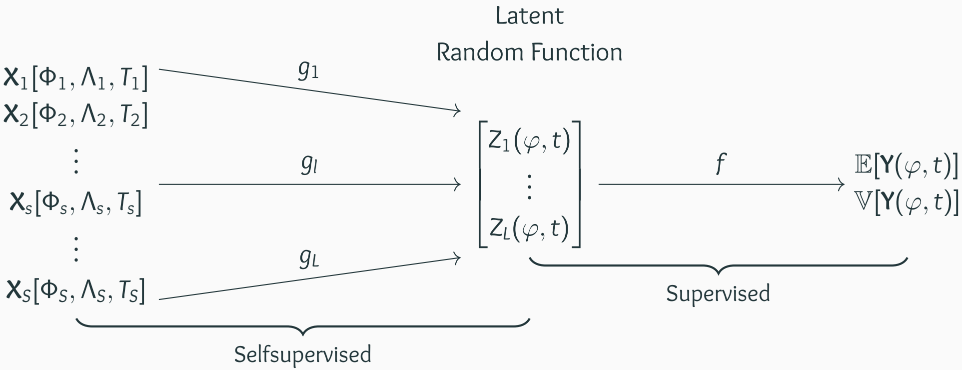

Latent
Random Function
$\mathbf{X}_1[\Phi_1, \Lambda_1, T_1]$
$\mathbf{X}_2[\Phi_2, \Lambda_2, T_2]$
$\mathbf{X}_s[\Phi_s, \Lambda_s, T_s]$
$\mathbf{X}_S[\Phi_S, \Lambda_S, T_S]$
g_1
g_l
g_L
$Z_1(\varphi, t)$
$Z_L(\varphi, t)$
f
$\mathbb{E}[\mathbf{Y}(\varphi, t)]$
$\mathbb{V}[\mathbf{Y}(\varphi, t)]$
Supervised
Selfsupervised

Framework



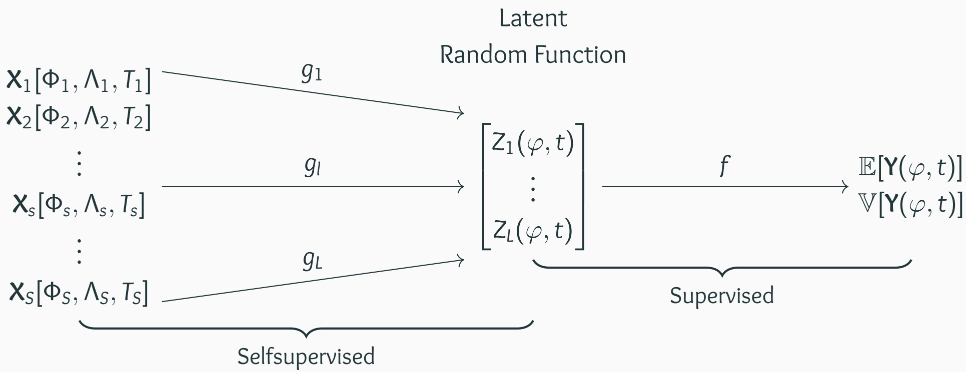
- Encoding functions g_l: “deep” structured networks

Framework



- Encoding functions g_l: “deep” structured networks
- Link function f: analytic formulation (or easy approximation) for posterior

Framework



- Encoding functions g_l: "deep" structured networks
- Link function f: analytic formulation (or easy approximation) for posterior
- **Latent functional space**
  - ⋆ Learn a fix discretization grid on virtual design points
  - ⋆ Decouple time and space: $Z_l(\psi, t) = \tilde{Z}_l(\psi) + \bar{Z}_l(t)$
  - ⋆ Learn parametric functional basis, e.g., *neural operator*

References

[1] A. Constantin, M. Fauvel, and S. Girard. "Mixture of multivariate Gaussian processes for classification of irregularly sampled satellite image time-series." In: *Statistics and Computing* 32.5 (Sept. 2022), p. 79.

[2] Alexandre Constantin. "Analyse de séries temporelles massives d'images satellitaires : Applications à la cartographie des écosystèmes." 2021GRALM069. PhD thesis. 2021. URL: http://www.theses.fr/2021GRALM069/document.

[3] M. Fauvel, M. Lopes, et al. "Prediction of plant diversity in grasslands using Sentinel-1 and -2 satellite image time series." In: *Remote Sensing of Environment* 237 (Feb. 2020).

[4] M. Fauvel, Y. Tarabalka, et al. "Advances in Spectral-Spatial Classification of Hyperspectral Images." In: *Proceedings of the IEEE* 101.3 (2013), pp. 652–675. DOI: 10.1109/JPROC.2012.2197589.

[5] F. Ferraty, A. Zullo, and M. Fauvel. “Nonparametric regression on contaminated functional predictor with application to hyperspectral data.” In: *Econometrics and Statistics* 9 (2019), pp. 95–107.

[6] Maïlys Lopes. “Ecological monitoring of semi-natural grasslands : statistical analysis of dense satellite image time series with high spatial resolution.” 2017INPT0095. PhD thesis. 2017. URL: http://www.theses.fr/2017INPT0095/document.

[7] Anthony Zullo. “Analyse de données fonctionnelles en télédétection hyperspectrale : application à l’étude des paysages agri-forestiers.” 2016TOU30135. PhD thesis. 2016. URL: http://www.theses.fr/2016TOU30135/document.

Thank you for your attention