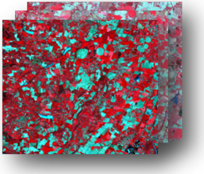
Object-based classification of grassland from high resolution satellite image time series with Gaussian mean map kernels

Presented by Stéphane Girard[1]
In collaboration with Mailys Lopes[1] and Mathieu Fauvel[2]

[1]Team Mistis, INRIA Grenoble, France,
[2]Dynafor, INRA, University of Toulouse, France

TIES-GRASPA Conference - 24-26 July 2017 - Bergamo, Italy



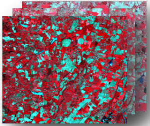
INRA SCIENCE & IMPACT

Inria INVENTORS FOR THE DIGITAL WORLD

Study objectives

Agroecological application

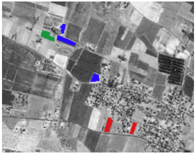
Discrimination of "old" permanent and "young" temporary grasslands



Data

SITS* with high spatial (≈ 10m) resolution and temporal (2-3 images per month) resolution

*satellite image time series



Method

Supervised classification of spatial objects

Context: grassland classification using dense satellite image time series

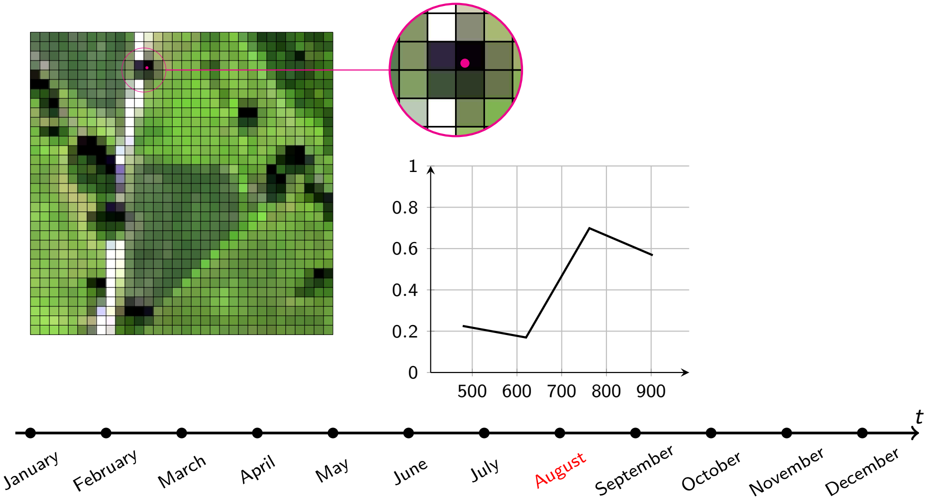
Gaussian mean kernel

Experimental results

Conclusion

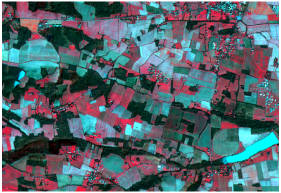
Remote sensing imagery

A digital remote sensing image corresponds to a spatial, spectral and temporal sampling of a landscape.



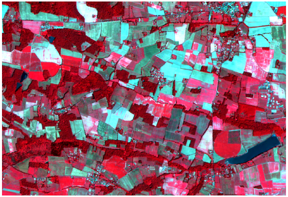
Satellite image time series

Formosat-2 (False color composites, Green, Red, NIR)



February

May



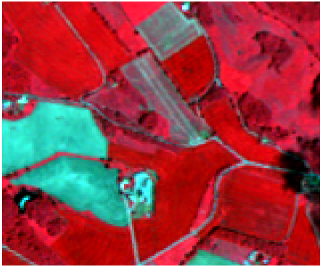
August

December

Normalized Difference Vegetation Index (NDVI)

NDVI: vegetation index that reflects the photosynthetic activity of the vegetation.

$$NDVI = \frac{NIR - Red}{NIR + Red}, \quad -1 \leq NDVI \leq 1$$



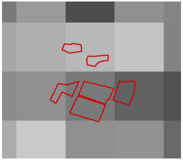
False Color (G, R, NIR)

$\Rightarrow$

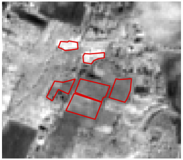
NDVI

Semi-natural grasslands in Europe:

- Relatively **small** ($\approx$ 100m $\times$ 100m) $\Rightarrow$ need **high spatial resolution images**
- **Heterogeneous** in species composition $\Rightarrow$ need **multispectral images**
- Have different **temporal behaviors** (phenology) $\Rightarrow$ need **high temporal** resolution images



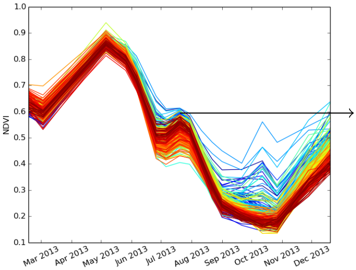
MODIS (250m)



Sentinel-2 (10m)

We propose to use **dense multispectral time series with high spatial resolution** to classify grasslands.

Grassland's pixels spectro-temporal profile:



$$\mathbf{x}_{ik} = \begin{pmatrix} x_{ik}(t_1) \\ \vdots \\ x_{ik}(t_d) \end{pmatrix} \in \mathbb{R}^d$$

with

- g_i: grassland with index i,
- n_i: number of pixels in g_i
- k: pixel index, $k \in \{1, \ldots, n_i\}$
- d: number of spectro-temporal variables
- $x_{ik}(t_l)$: spectral value of pixel k at time l

Grassland representation:

- $\mathbf{X}_i = \left[\mathbf{x}_{i1} | \ldots | \mathbf{x}_{in_i}\right]$ is a matrix of size $(n_i \times d)$ that contains all the pixels inside g_i.
- Learn f such as $y_i = f(\mathbf{X}_i)$, where y_i is the predicted label.

Thematic contributions

- **Grassland** classification (semi-natural elements)
- **Sentinel-2** contribution (new generation satellites, dense time series)

Methodological contributions

- **Model grassland's pixels distribution**
- Process grassland **supervised classification at the grassland scale**
- **Robust** to
  - the **dimension of data** (n_i pixels, d spectro-temporal variables with $n_i \approx d$),
  - the **total number of grasslands pixels** which might be large.

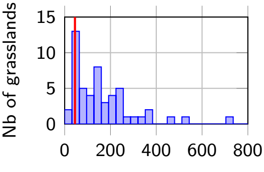


Figure: Histogram of grasslands size in number of pixels n_i. The red line corresponds to the number of variables $d = 45$.

Several ways of modeling grasslands in the remote sensing literature:

- **Pixel level**, where the pixels are the samples: the response variable y_i of g_i is associated with each pixel $\mathbf{x}_{ik}$, but each $\mathbf{x}_{ik}$ is processed independently of all others $\mathbf{x}_{ik'}$ of g_i.
- **Object level**
  - **Mean vector** $\boldsymbol{\mu}_i$ of g_i is used to represent g_i:

$$\hat{\boldsymbol{\mu}}_i = \frac{1}{n_i}\sum_{k=1}^{n_i}\mathbf{x}_{ik}.$$

Type	Pros	Cons
Pixel by pixel	Account for the heterogeneity in the grassland	Large computational cost with SVM
Mean	Reduced processing time	Limited representation, does not account for heterogeneity

We chose to model the grassland's pixels distribution by a Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ where:

$$\hat{\boldsymbol{\Sigma}}_i = \frac{1}{n_i - 1} \sum_{k=1}^{n_i} (\mathbf{x}_{ik} - \hat{\boldsymbol{\mu}}_i)(\mathbf{x}_{ik} - \hat{\boldsymbol{\mu}}_i)^\top.$$

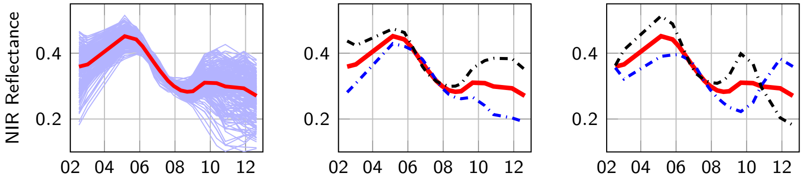


Figure: Left: temporal profile of all the pixels in the grassland and their temporal mean in red. Middle: temporal mean in red, $+0.2\times$ the 1st eigenvector in blue and $-0.2\times$ the 1st eigenvector in black. Right: temporal mean in red, $+0.2\times$ the 2nd eigenvector in blue and $-0.2\times$ the 2nd eigenvector in black.

- Pixel based and mean modelings: conventional RBF kernel
- Distributions
  - Kullback-Leibler divergence
  - Bhattacharyya distance

Conventional similarity measures used for moderate dimensional Gaussian distributions are not suitable for high dimensional Gaussian distributions.

- **Empirical mean kernel**:

$$K^e(p_i, p_j) = \frac{1}{n_i n_j} \sum_{l,m=1}^{n_i,n_j} k(\mathbf{x}_{il}, \mathbf{x}_{jm}),$$

where p_i and p_j are distributions. $\mathbf{x}_{il}$ is the l^{th} realization of p_i, and k is a semi-definite positive kernel function.

- **Generative mean kernel**:

$$K^g(p_i, p_j) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} k(\mathbf{x}, \mathbf{x}')\hat{p}_i(\mathbf{x})\hat{p}_j(\mathbf{x}')d\mathbf{x}d\mathbf{x}'.$$

- When p_i and p_j are Gaussian distributions and k is a Gaussian kernel, this becomes the **Gaussian mean kernel**:

$$\tilde{K}^G(\mathcal{N}_i, \mathcal{N}_j) = \frac{\exp\left\{-0.5(\hat{\boldsymbol{\mu}}_i - \hat{\boldsymbol{\mu}}_j)^T \left(\hat{\boldsymbol{\Sigma}}_i + \hat{\boldsymbol{\Sigma}}_j + \gamma^{-1}\mathbf{I}_d\right)^{-1} (\hat{\boldsymbol{\mu}}_i - \hat{\boldsymbol{\mu}}_j)\right\}}{|\hat{\boldsymbol{\Sigma}}_i + \hat{\boldsymbol{\Sigma}}_j + \gamma^{-1}\mathbf{I}_d|^{0.5}|2\hat{\boldsymbol{\Sigma}}_i + \gamma^{-1}\mathbf{I}_d|^{0.25}|2\hat{\boldsymbol{\Sigma}}_j + \gamma^{-1}\mathbf{I}_d|^{0.25}},$$

where γ is a positive regularization parameter coming from the Gaussian kernel k.

Proposition:
α-**generative mean kernel**:

$$K^{\alpha}(p_i, p_j) \quad = \quad \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} k(\mathbf{x}, \mathbf{x}') \hat{p}_i(\mathbf{x})^{(\alpha^{-1})} \hat{p}_j(\mathbf{x}')^{(\alpha^{-1})} d\mathbf{x} d\mathbf{x}'.$$

When p_i and p_j are Gaussian distributions, k is a Gaussian kernel and the normalization is applied, the expression gives rise to the α-**Gaussian mean kernel**:

$$\tilde{K}^{\alpha}(\mathcal{N}_i, \mathcal{N}_j) = \frac{\exp\left\{-0.5(\hat{\boldsymbol{\mu}}_i - \hat{\boldsymbol{\mu}}_j)^T \left(\alpha(\hat{\boldsymbol{\Sigma}}_i + \hat{\boldsymbol{\Sigma}}_j) + \gamma^{-1}\mathbf{I}_d\right)^{-1} (\hat{\boldsymbol{\mu}}_i - \hat{\boldsymbol{\mu}}_j)\right\}}{|\alpha(\hat{\boldsymbol{\Sigma}}_i + \hat{\boldsymbol{\Sigma}}_j) + \gamma^{-1}\mathbf{I}_d|^{0.5} |2\alpha\hat{\boldsymbol{\Sigma}}_i + \gamma^{-1}\mathbf{I}_d|^{0.25} |2\alpha\hat{\boldsymbol{\Sigma}}_j + \gamma^{-1}\mathbf{I}_d|^{0.25}}.$$

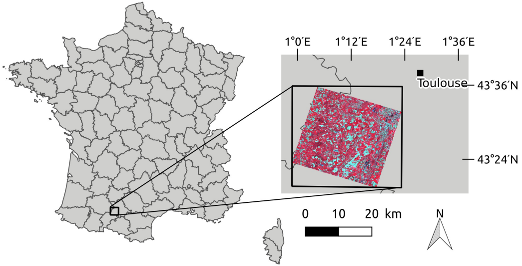
Context: grassland classification using dense satellite image time series

Gaussian mean kernel

Experimental results

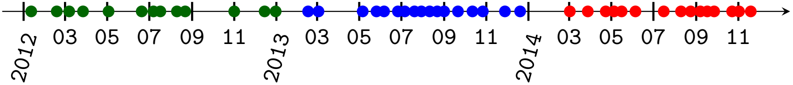
Conclusion

Study area



Satellite data

Formosat-2 (8m) inter-annual time series of **NDVI** from 2012 to 2014 (**45 dates**).



Data to classify

- Old grasslands: 14 years old and more
- Young grasslands: less than 5 years old

Class	Nb of grasslands	Nb of pixels
Old	59	31,166
Young	416	129,348
Total	475	160,514

Methods based on RBF kernel:

- **PMV** (Pixel Majority Vote): It classifies each pixel with no *a priori* information on the object which the grassland belongs to. Then, a majority vote is performed.
- $\boldsymbol{\mu}$ (mean): The distribution of the pixels reflectance of g_i is modeled by its mean vector $\boldsymbol{\mu}_i$.
- **BD** (Bhattacharyya Distance): This method uses the Bhattacharyya distance in the case of Gaussian distributions.

Method based on mean map kernels:

- **EMK** (Empirical Mean Kernel)
- **GMK** (Gaussian Mean Kernel)
- α**GMK** (α-Gaussian Mean Kernel).

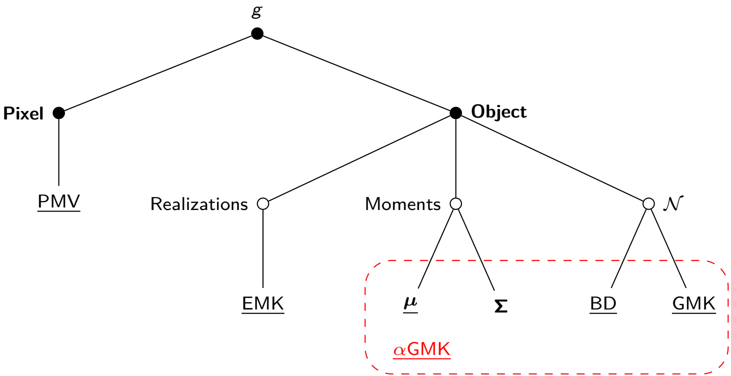
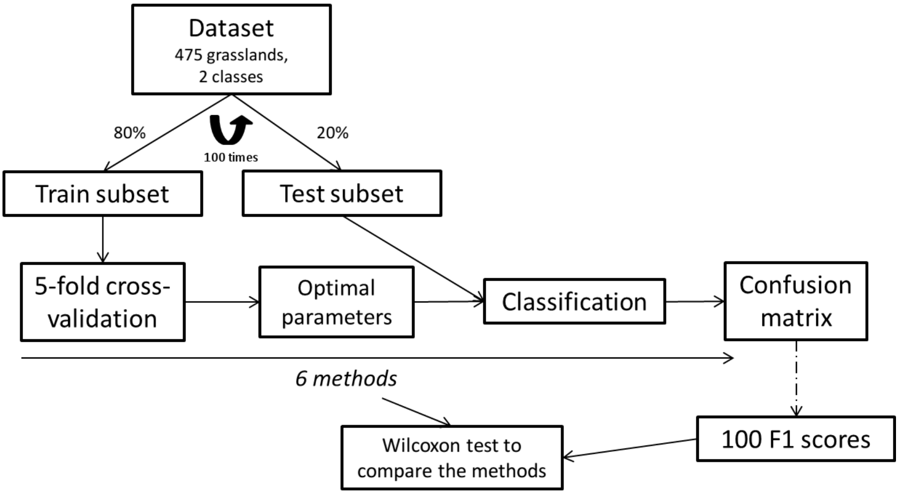


Figure: Contribution of the proposed method in grassland analysis for supervised classification. αGMK consists in a general modeling of the grassland at the object level and it encompasses several known modelings. The underlined methods are tested in this study.


Dataset
475 grasslands,
2 classes
80%
100 times
20%
Train subset
Test subset
5-fold cross-validation
Optimal parameters
Classification
Confusion matrix
6 methods
Wilcoxon test to compare the methods
100 F1 scores

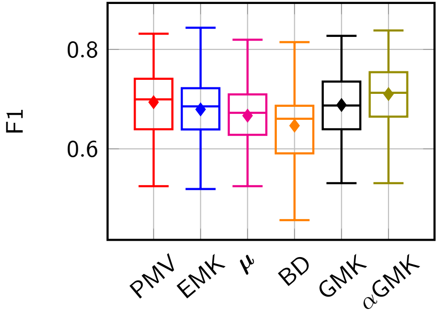


Table: Absolute value of Wilcoxon rank-sum test statistics on F1 score. ** indicates the results are significantly different, *i.e.*, p-value < 0.05.

Method	**PMV**	$\boldsymbol{\mu}$	**BD**	**EMK**	**GMK**	α**GMK**
PMV	-	3.52**	4.83**	1.93	0.98	1.32
$\boldsymbol{\mu}$		-	1.76	1.55	2.28**	4.80**
BD			-	3.23**	3.95**	6.09**
EMK				-	0.94	3.35**
GMK					-	2.42**
α**GMK**						-

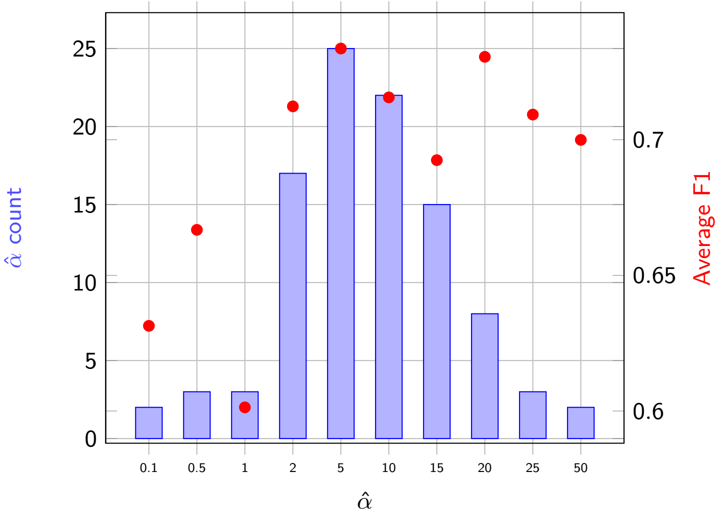


Figure: Bar plot of $\hat{\alpha}$ values chosen by cross-validation and the average of associated F1 scores (red dots) using αGMK. NB: The value $\hat{\alpha} = 0$ was never selected.

- **Flexible kernel** that encompasses both Gaussian and mean modelings.
- Kernel suitable for **high dimensional data** (low computational load).
- Good compromise between **processing speed and accuracy**.
- First application of generative mean kernels in remote sensing.
- Suitable for the **classification of small and heterogeneous objects** such as grasslands, but it could be used for other land cover (urban areas, peatlands..).

Thank you for your attention

Appendix

Table: Characteristics of the methods used in this study.

Method	PMV	EMK	$\boldsymbol{\mu}$	BD	GMK	αGMK
Level	Pixel	Object	Object	Object	Object	Object
Expl. variable	$\mathbf{x}_{ik}$	$\mathbf{x}_{ik}$	$\boldsymbol{\mu}_i$	$\mathcal{N}_i$	$\mathcal{N}_i$	$\mathcal{N}_i$
Kernel	RBF	RBF	RBF	K_{B}	$\tilde{K}^G$	$\tilde{K}^\alpha$
Parameters	σ, C	σ, C	σ, C	σ, C	γ, C	γ, α, C
Nb of samples	$1/10 \cdot 162,500$	$1/10 \cdot 162,500$	475	475	475	475