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
Study objectives

**Agroecological application**
 Discrimination of "old" permanent and "young" temporary grasslands

**Data**
 SITS* with high spatial ($\approx 10$m) 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.

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

We propose to use **dense multispectral time series with high spatial resolution** to classify grasslands.
Grassland’s pixels spectro-temporal profile:

\[ x_{ik} = \begin{pmatrix} x_{ik}(t_1) \\ \vdots \\ x_{ik}(t_d) \end{pmatrix} \in \mathbb{R}^d \]

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:
- \( \mathbf{X}_i = \begin{bmatrix} x_{i1} & \cdots & x_{in_i} \end{bmatrix} \) 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.

![Histogram of grasslands size in number of pixels $n_i$. The red line corresponds to the number of variables $d = 45$.](image)

**Figure**: Histogram of grasslands size in number of pixels $n_i$. The red line corresponds to the number of variables $d = 45$. 
<table>
<thead>
<tr>
<th>Context: grassland classification using dense satellite image time series</th>
<th>Gaussian mean kernel</th>
<th>Experimental results</th>
<th>Conclusion</th>
</tr>
</thead>
</table>

**Gaussian mean kernel**

**Experimental results**

**Conclusion**
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 $x_{ik}$, but each $x_{ik}$ is processed independently of all others $x_{ik'}$ of $g_i$.

- **Object level**
  - Mean vector $\mu_i$ of $g_i$ is used to represent $g_i$:
    \[
    \hat{\mu}_i = \frac{1}{n_i} \sum_{k=1}^{n_i} x_{ik}.
    \]

<table>
<thead>
<tr>
<th>Type</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pixel by pixel</strong></td>
<td>Account for the heterogeneity in the grassland</td>
<td>Large computational cost with SVM</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>Reduced processing time</td>
<td>Limited representation, does not account for heterogeneity</td>
</tr>
</tbody>
</table>
We chose to model the grassland’s pixels distribution by a Gaussian distribution $\mathcal{N}(\mu_i, \Sigma_i)$ where:

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

**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.
Similarity measures between distributions

- 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:

\[ K^e(p_i, p_j) = \frac{1}{n_i n_j} \sum_{l,m=1}^{n_i,n_j} k(x_{il}, x_{jm}), \]

where \( p_i \) and \( p_j \) are distributions. \( x_{il} \) is the \( l^{th} \) realization of \( p_i \), and \( k \) is a semi-definite positive kernel function.

Generative mean kernel:

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

When \( p_i \) and \( p_j \) are Gaussian distributions and \( k \) is a Gaussian kernel, this becomes the Gaussian mean kernel:

\[ \tilde{K}^G(N_i, N_j) = \exp \left\{ -0.5(\hat{\mu}_i - \hat{\mu}_j)^T \left( \hat{\Sigma}_i + \hat{\Sigma}_j + \gamma^{-1} I_d \right)^{-1} (\hat{\mu}_i - \hat{\mu}_j) \right\} \]

\[ \frac{1}{|\hat{\Sigma}_i + \hat{\Sigma}_j + \gamma^{-1} I_d|^{0.5} |2\hat{\Sigma}_i + \gamma^{-1} I_d|^{0.25} |2\hat{\Sigma}_j + \gamma^{-1} I_d|^{0.25}}, \]

where \( \gamma \) is a positive regularization parameter coming from the Gaussian kernel \( k \).
Proposition:

$\alpha$-generative mean kernel:

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

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 $\alpha$-Gaussian mean kernel:

$$\tilde{K}^\alpha(N_i, N_j) = \exp \left\{ -0.5(\hat{\mu}_i - \hat{\mu}_j)^T \left( \alpha(\hat{\Sigma}_i + \hat{\Sigma}_j) + \gamma^{-1} I_d \right)^{-1} (\hat{\mu}_i - \hat{\mu}_j) \right\}$$

$$\frac{|\alpha(\hat{\Sigma}_i + \hat{\Sigma}_j) + \gamma^{-1} I_d|^{0.5} |2\alpha\hat{\Sigma}_i + \gamma^{-1} I_d|^{0.25} |2\alpha\hat{\Sigma}_j + \gamma^{-1} I_d|^{0.25}}{\alpha(\hat{\Sigma}_i + \hat{\Sigma}_j) + \gamma^{-1} I_d}. $$
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

<table>
<thead>
<tr>
<th>Class</th>
<th>Nb of grasslands</th>
<th>Nb of pixels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Old</td>
<td>59</td>
<td>31,166</td>
</tr>
<tr>
<td>Young</td>
<td>416</td>
<td>129,348</td>
</tr>
<tr>
<td>Total</td>
<td>475</td>
<td>160,514</td>
</tr>
</tbody>
</table>
Methods based on RBF kernel:

- **PMV** (Pixel Majority Vote): It classifies each pixel with no \textit{a priori} information on the object which the grassland belongs to. Then, a majority vote is performed.

- **\( \mu \)** (mean): The distribution of the pixels reflectance of \(g_i\) is modeled by its mean vector \(\mu_i\).

- **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)

- **\( \alpha \text{GMK} \)** (\( \alpha \)-Gaussian Mean Kernel).
Figure: Contribution of the proposed method in grassland analysis for supervised classification. $\alpha$GMK 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.
Classification protocol

Dataset
475 grasslands, 2 classes

Train subset
Test subset

80% 20%
100 times

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.
Figure: Bar plot of \( \hat{\alpha} \) values chosen by cross-validation and the average of associated F1 scores (red dots) using \( \alpha \)GMK. NB: The value \( \hat{\alpha} = 0 \) was never selected.
<table>
<thead>
<tr>
<th>Context: grassland classification using dense satellite image time series</th>
<th>Gaussian mean kernel</th>
<th>Experimental results</th>
<th>Conclusion</th>
</tr>
</thead>
</table>

Context: grassland classification using dense satellite image time series

Gaussian mean kernel

Experimental results

Conclusion
- **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
### Table: Characteristics of the methods used in this study.

<table>
<thead>
<tr>
<th>Method</th>
<th>PMV</th>
<th>EMK</th>
<th>$\mu$</th>
<th>BD</th>
<th>GMK</th>
<th>$\alpha$GMK</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level</strong></td>
<td>Pixel</td>
<td>Object</td>
<td>Object</td>
<td>Object</td>
<td>Object</td>
<td>Object</td>
</tr>
<tr>
<td><strong>Expl. variable</strong></td>
<td>$x_{ik}$</td>
<td>$x_{ik}$</td>
<td>$\mu_i$</td>
<td>$N_i$</td>
<td>$N_i$</td>
<td>$N_i$</td>
</tr>
<tr>
<td><strong>Kernel</strong></td>
<td>RBF</td>
<td>RBF</td>
<td>RBF</td>
<td>$K_B$</td>
<td>$\tilde{K}^G$</td>
<td>$\tilde{K}^\alpha$</td>
</tr>
<tr>
<td><strong>Parameters</strong></td>
<td>$\sigma, C$</td>
<td>$\sigma, C$</td>
<td>$\sigma, C$</td>
<td>$\gamma, C$</td>
<td>$\gamma, \alpha, C$</td>
<td></td>
</tr>
<tr>
<td><strong>Nb of samples</strong></td>
<td>$1/10 \cdot 162,500$</td>
<td>$1/10 \cdot 162,500$</td>
<td>475</td>
<td>475</td>
<td>475</td>
<td>475</td>
</tr>
</tbody>
</table>