



HAL
open science

**Object-based classification from high resolution satellite
image time series with Gaussian mean map kernels:
Application to grassland management practices**

Maïlys Lopes, Mathieu M. Fauvel, Stéphane Girard, David Sheeren

► **To cite this version:**

Maïlys Lopes, Mathieu M. Fauvel, Stéphane Girard, David Sheeren. Object-based classification from high resolution satellite image time series with Gaussian mean map kernels: Application to grassland management practices. 2017. hal-01424929v1

HAL Id: hal-01424929

<https://inria.hal.science/hal-01424929v1>

Preprint submitted on 3 Jan 2017 (v1), last revised 13 Jun 2017 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Object-based classification from high resolution satellite image time series with Gaussian mean map kernels: Application to grassland management practices

Mailys Lopes, *Student Member*, Mathieu Fauvel, *Senior Member*, Stéphane Girard, and David Sheeren

Abstract

This paper deals with the classification of grassland management practices using high resolution satellite image time series. Grasslands considered in this work are semi-natural elements in fragmented landscapes, *i.e.*, they are heterogeneous and small elements. The first contribution of this study is to account for grassland heterogeneity while working at the object scale by modeling its pixels distributions by a Gaussian distribution. To measure the similarity between two grasslands, a Gaussian mean map kernel is proposed as a second contribution: the α -Gaussian mean kernel. It allows to weight the influence of the covariance matrix when comparing two grasslands. This kernel is used in Support Vector Machine for the supervised classification of three management practice types in 52 grasslands from south-west France, using an intra-annual multispectral time series of Formosat-2 satellite. Results in terms of classification accuracy and processing time are compared to other pixel- and object-based approaches. The proposed modeling showed to be the best compromise between processing speed and classification accuracy. It can adapt to the classification constraints and it encompasses several similarity measures known in the literature.

Index Terms

Supervised classification, Gaussian mean map kernels, kernel methods, object scale, grasslands.

M. Lopes, M. Fauvel and D. Sheeren are with Dynafor, University of Toulouse, INRA, INPT, INP-EI Purpan, Castanet Tolosan, France (e-mail: mailys.lopes@inra.fr; mathieu.fauvel@ensat.fr; david.sheeren@ensat.fr).

S. Girard is with Team Mistis, INRIA Rhône-Alpes, LJK, Montbonnot, France (e-mail: stephane.girard@inria.fr).

This work was granted by INRA and by Défi Mastodons-CNRS.

I. INTRODUCTION

Grasslands are semi-natural elements that represent a significant source of biodiversity in farmed landscapes [1]–[4]. They provide many ecosystem services such as carbon storage, erosion regulation, food production, crop pollination, biological regulation of pests [5], that are linked to their plant and animal composition.

Agricultural management of semi-natural grasslands is essential for their biodiversity conservation. The species composition in a grassland is impacted by the farmer’s management practices [6]–[10]. However, an intensive use constitutes a threat for this biodiversity [10], [11]. Indeed, the anthropogenic events in the grasslands, like mowing, grazing, fertilizing and/or reseeding, disturb the natural cycle and the structure of the vegetation. This diversity of species needs to be preserved, as recent studies showed that service-providing species are the most sensitive to an increase in land use intensity [12], [13]. At the same time, grasslands being the main livestock feeding resource, it is important to find a tradeoff between biodiversity conservation and grassland productivity. Therefore, it is essential to identify the management practices (number of cuts, intensity of grazing) in each grassland in order to predict their effect on biodiversity and related services.

Usually, ecologists and agronomists characterize grasslands at the parcel scale through field surveys. However, these surveys require important human and material resources, the knowledge of the assessor and a sampling strategy, which make them expensive and time-consuming. They are thus limited spatially and in temporal frequency, limiting grassland characterization at the local scale.

Conversely, remote sensing offers the possibility to provide information on landscapes over large extents, because of the broad spatial coverage and regular revisit frequency of satellite sensors [14], [15]. In this context, satellite images have already appeared to be an appropriate tool to monitor grasslands over large areas with a high temporal resolution.

In the remote sensing literature, grasslands have relatively not been studied much compared to other land covers like crops or forest [16]. Most of studies focusing on grasslands have agronomic applications, such as estimating biomass productivity and growth rate [17]–[19] or derivating biophysical parameters like LAI, fPAR and chlorophyll content [20]–[24]. Studies having biodiversity conservation schemes such as assessing plant diversity and plant community composition in a grassland are usually based on ground spectral measurements or airborne acquisitions at a very high spatial resolution [25]–[31]. However, such acquisitions are time-consuming and expensive and thus do not allow continuous monitoring of grasslands over the years.

Using satellite remote sensing images, grasslands have been studied a lot at a regional scale with medium spatial resolution sensors (*i.e.*, MODIS, 250m/pixel [17], [18], [32]), where the Minimum Mapping Unit (MMU) is at least of hundreds of meters. This scale is suitable for large, extensive, homogeneous and contiguous regions like steppes [33], but not for fragmented landscapes that are usually found in Europe and in France

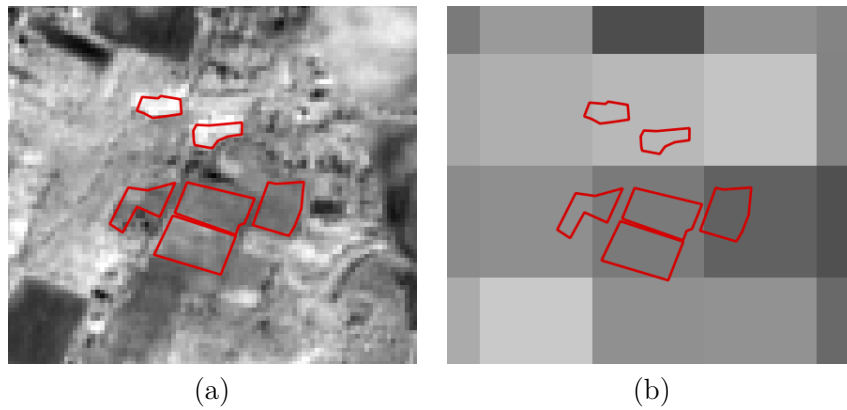


Figure 1. Digitalized grasslands (in red) from the dataset used in this study on a (a) Sentinel-2 image (10m) and a (b) MODIS image (250m).

particularly [34]. These landscapes are made of a patchwork of different land covers which have a small area [34]. In these types of landscapes, grasslands can be smaller (less than $10,000m^2$) than the pixel resolution [15] (see Fig. 1 for a graphical example). As a consequence, pixels containing grasslands are usually a mixture of other contributions, which can limit the analysis [35], [36]. As examples, Poças *et al.* [37] had to select large contiguous areas of semi-natural grasslands in a mountain region of Portugal to be able to use SPOT-VEGETATION data (1-km resolution). Halabuk *et al.* [38] also had to select only one MODIS pixel per homogeneous sample site in Slovakia to detect cutting in hay meadows. A 30-m pixel resolution is still not sufficient for grassland characterization. Lucas *et al.* [39] and Toivonen and Luoto [40] showed that it was more difficult to classify fragmented and complex elements [41], like semi-natural grasslands, than homogeneous habitats, using Landsat imagery. Price *et al.* [42] classified six grassland management types in Kansas using six Landsat images, but the accuracy of the classification was not satisfying (less than 70%). Therefore, to detect small grasslands in fragmented landscapes, high spatial resolution images are required [15], [43], [44].

For high spatial resolution images (about 10m/pixel), few intra-annual images are usually available for a given location [45]. However, Buck *et al.* [46] concluded that three Rapid-Eye images per year were not enough to detect the mowing practices in grasslands. It was confirmed by Franke *et al.* [47] who increased the classification accuracy when using more Rapid-Eye images (from three to five images) to identify grassland use intensity. Some works report results with few images per year, such as Dusseux *et al.* [48], but on artificial and/or almost monospecific grasslands (mainly composed of rye-grass and clover). They can be considered as similar to crops in terms of phenology and homogeneity and do not reflect the diversity of semi-natural grasslands. In their study for mapping grassland habitat using intra-annual Rapid-Eye imagery, Schuster *et al.* [49] concluded the more acquisition dates were used, the better the mapping quality.

Given the heterogeneity of grasslands in fragmented landscapes, their phenological cycle and the punctuality of the anthropogenic events (*e.g.*, mowing), dense high spatial resolution time series during the vegetation

cycle are necessary to identify the grassland management types [15], [49]–[51]. Until recently, satellite missions offering high revisit frequency (1-16 days) had coarse spatial resolution (*i.e.*, NOAA AVHRR - 1km, MODIS - 250/500m). Conversely, high spatial resolution missions did not provide dense time series and/or were costly (*i.e.*, QuickBird, RapidEye) . For these reasons and compared to crops, grasslands differentiation through Earth observations is still considered as a challenge [49]. However, new missions like Sentinel-2 [52], with very high revisit frequency (5 days) and high spatial resolution (10 meters in four spectral channels, 20 meters in six channels) provide new opportunities for grasslands monitoring in fragmented landscapes [51] at no cost, thanks to the ESA free data access policy. For instance, the high spatial resolution could allow to identify grassland-only pixels in the image and several pixels could belong to the same grassland plot. Hence, the analysis can be done at the object scale, not at the pixel scale, which is suitable for landscape ecologists and agronomists who usually study grasslands at the parcel scale [53]. Thus, object-oriented approaches are more likely to characterize grasslands ecologically [54], [55]. Yet, a lot of works consider pixel-based approaches without any spatial constraints [17], [40], [42], [46], [47], [49], [56].

At the object scale, grasslands are commonly represented by their mean NDVI [18]. But such representation might be too simple since it does not account for the heterogeneity in a grassland. Sometimes, distributions of pixels as individual observations are still better than the mean value to represent grasslands, as in [51]. Lucas *et al.* [39] used a rule-based method on segmented areas for habitat mapping but it did not work well on complex and heterogeneous land covers. Esch *et al.* [57] also used an object-oriented method on segmented elements then represented by their mean NDVI. These methods based on mean modeling do not capture well grasslands heterogeneity. Other representations can be found in the literature, taking standard deviation and object texture features as variables [58], but they were not applied on time series. In our knowledge, these methods do not use the high spatial and the high temporal resolutions jointly. Moreover, all these studies used vegetation indices as a variable. It has been shown that classification results are better when using more spectral information [34], [59].

A few works have been done on addressing the high spatio-spectro-temporal resolutions in the literature. Yet, it is the type of data the new satellite sensors are now offering. Such time series bring new methodological and statistical constraints given the high dimension of data (*i.e.*, number of pixels and number of spectral and temporal measurements). When dealing with more variables, it increases the number of parameters to estimate, increasing the computation time and making the computation unstable (*i.e.*, ill-conditioned covariance matrices...) [60]. Thus, the use of high dimensional data requires to adapt the methodological and statistical approaches to process a robust supervised classification. Therefore, classifying grasslands at the parcel-scale, using dense multispectral time series with high spatial resolution is still a challenge and

appropriate models are lacking.

In the present study, we introduce a model suitable for the classification of management practices in semi-natural grasslands using satellite image time series (SITS) with a high number of spectro-temporal variables, *e.g.*, Sentinel-2 data. The first contribution of the method is to model a grassland at the object-scale while accounting for the spectral variability within a grassland. We consider that the distribution of the pixel spectral reflectance in a given grassland can be modeled by a Gaussian distribution. The second contribution is to propose a measure of similarity between two Gaussian distributions that is robust to the high dimension of the data. This method is based on mean maps. The last contribution is the application of the method on grassland management practices classification, which is a non common application in that field. In our knowledge, mean maps have not yet been used on Gaussian distributions for supervised classification of SITS at the object scale.

In the next section, the materials used for this study are presented. The different types of grassland modeling are presented in section III. Then, the mean map kernel method to measure the similarity is introduced in section IV. Following that, we experiment the method and other conventional methods on the classification of a real dataset in section V. Finally, conclusions and prospects are given in section VI.

II. MATERIALS

A. Study site

The study site is located in south-west France, near the city of Toulouse (about 30 km), in a semi-rural area (center coordinates: 43°27'36"N 1°8'24"E, Fig. 2). This region is characterized by a temperate climate with oceanic and mediterranean influences. The average annual precipitation is 656 mm and the average temperature is 13°C. The north of the site, closer to the urban area of Toulouse, is flat, whereas the south-west of the site is hilly. The eastern part corresponds to the Garonne river floodplain and this location is dominated by crop production. Within this study site, livestock farming is declining in favor of annual crop production. Grasslands are mostly used for forrage or silage production. Some grasslands, located in the south-western part of the area, are pastures for cattle or sheep. The extent of the area is about 440km² and is included in the satellite image extent (Fig. 2).

B. Satellite data

The satellite image time series (SITS) consists of 17 Formosat-2 images acquired in 2013 and 15 images acquired in 2014 (Fig. 3). In this study, we worked on each year separately. Formosat-2 has four spectral bands with an 8-meter spatial resolution: B1 "Blue" (0.45 - 0.52μm), B2 "Green" (0.52 - 0.6μm), B3 "Red"

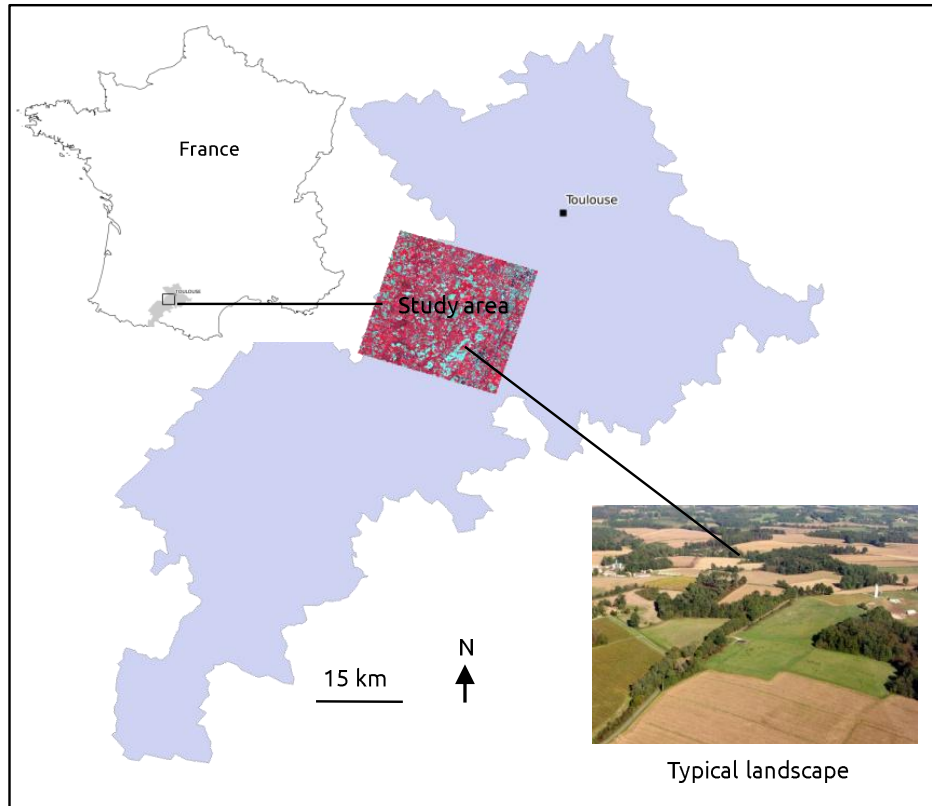


Figure 2. Study site location in south-west France. It is included in the satellite image extent.

($0.63 - 0.69\mu\text{m}$), B4 "Near Infra-Red (NIR)" ($0.76 - 0.9\mu\text{m}$). The extent of an acquisition is $24\text{km} \times 24\text{km}$. The images were all acquired with the same viewing angle. They were orthorectified, radiometrically and atmospherically corrected by the French Spatial Agency (CNES). They were provided by the Center for the Study of the Biosphere from Space (CESBIO) in reflectance with a mask of clouds and shadows issued from the MACCS (Multi-sensor Atmospheric Correction and Cloud Screening) processor [61], in the frame of the Kalideos project.

To reconstruct the time series due to missing data (clouds and their shadows), the Whittaker filter [62] was applied pixel-by-pixel on the reflectances in each spectral band. The smoother was adapted for unequally spaced intervals and accounted for missing data (see [59] for a detailed description of the method). The smoothing parameter was the same for all the pixels. It was equal to 10^5 for year 2013 and to 10^4 for 2014, after an ordinary cross-validation done on a subset of the pixels for each year. An example of smoothing on a grassland pixel is provided in Fig. 4. This pixel is hidden by a light cloud during one image acquisition (red cross). Notice that the smoothing is done at the cost of minimizing the local maxima.

The smoothed time series associated with each of the four spectral bands were then concatenated to get a unique time series per pixel.

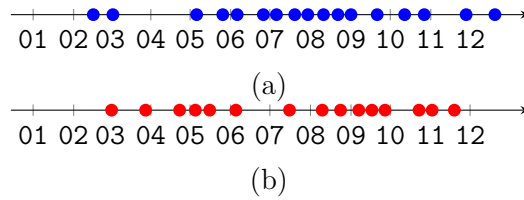


Figure 3. Satellite image time series acquisition dates in (a) 2013 and (b) 2014.

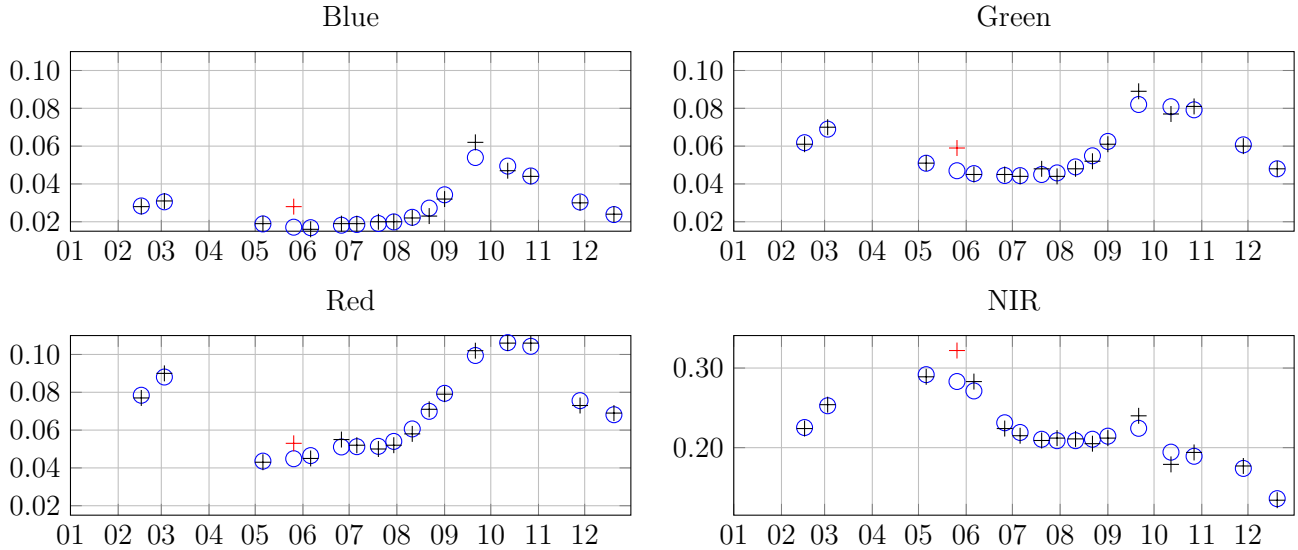


Figure 4. Example of time series reconstruction (blue dots) with Whittaker smoother for a pixel of a grassland in the four spectral bands. The black crosses correspond to the original 2013 Formosat-2 time series and the red ones correspond to missing/noisy data due to the clouds. The x-axis represents the month of year 2013 and the y-axis is the reflectance.

C. Field data

The dataset is composed of 52 grasslands that were selected during a field survey in May, 2015. The past and current management practices were determined for each parcel by interviewing the farmers or grassland owners. The practices remained stable for the years 2013 and 2014. Four management types during a vegetation cycle were identified: one mowing, two mowings, grazing and mixed (mowing then grazing). We eliminated the type "two mowings" of the dataset because of its under-representation (only three parcels). The management types were used as classes for the classification (Table I). The grasslands have been extracted from the French agricultural land use database (Registre Parcellaire Graphique) (Fig. 5) in a polygon form. A negative buffer of 8 meters was then applied to eliminate the edge effects, before rasterizing the polygons. The average grasslands surface area is about $10,000m^2$. The smallest grassland is $1,632m^2$ (which represents 25 Formosat-2 pixels) and the largest is $47,111m^2$ (735 pixels) (Fig. 6).

III. GRASSLAND MODELING

In this work, each grassland g_i is composed of a given number n_i of pixels $\mathbf{x}_{ik} \in \mathbb{R}^d$, where k is the pixel index such as $k \in \{1, \dots, n_i\}$, $i \in \{1, \dots, G\}$, G is the total number of grasslands, $N = \sum_{i=1}^G n_i$ is the total

Table I
GRASSLAND MANAGEMENT TYPES AND COMPOSITION OF THE DATASET.

Class	Nb of grasslands	Nb of pixels
Mowing	34	6265
Grazing	10	1193
Mixed	8	1170
Total	52	8628

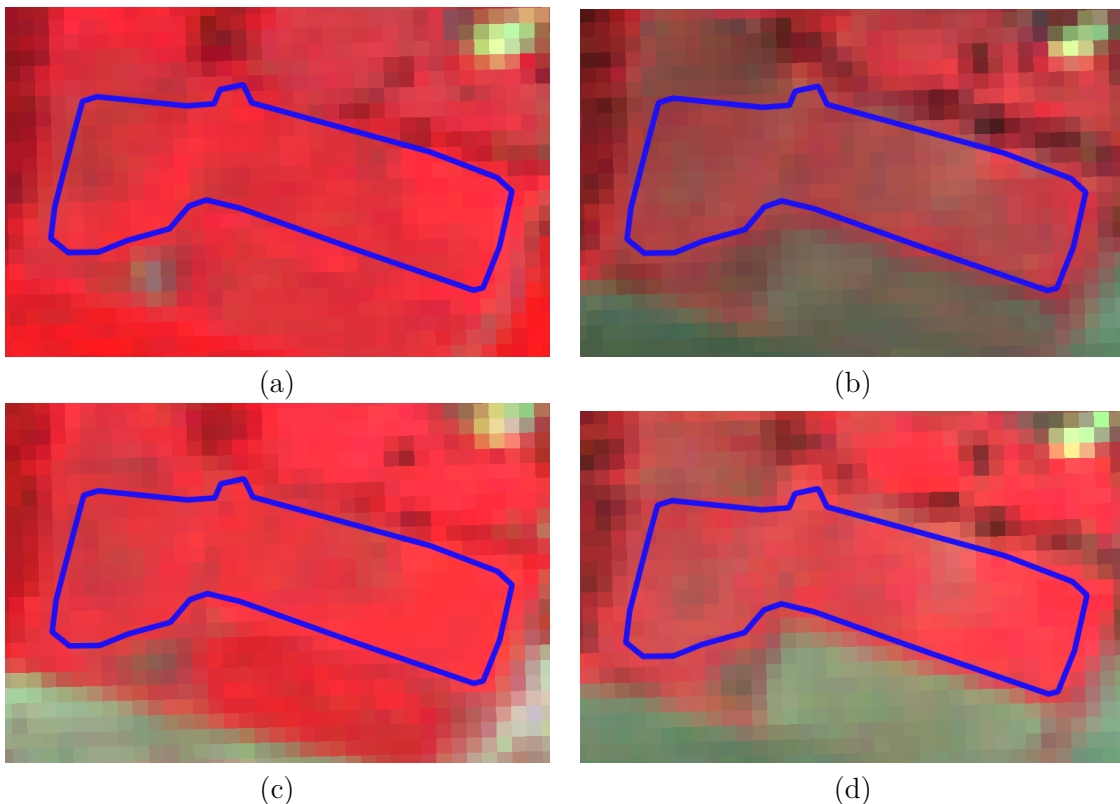


Figure 5. False color Formosat-2 images of the same grassland on two close dates (June and October) in 2013 and 2014 with the same color scale. (a) 2013-06-06, (b) 2013-10-27, (c) 2014-06-05, (d) 2014-10-23. The blue line represents the polygon limits of the grassland.

number of pixels, $d = n_B n_T$ is the number of spectro-temporal variables, n_B is the number of spectral bands and n_T is the number of temporal acquisitions. With each grassland g_i are associated a matrix \mathbf{X}_i of size $(n_i \times d)$ and a response variable $y_i \in \mathbb{R}$ which corresponds to its class label.

In the following, two types of grassland modeling are discussed, at the pixel-scale and at the object-scale. A more informative object-scale modeling is then proposed.

A. Pixel-scale

The representation of a grassland at the pixel scale has been used a lot in the remote sensing literature [17], [40], [42], [46], [47], [49], [56]. The grassland can either be represented by all its pixels or by one pixel when

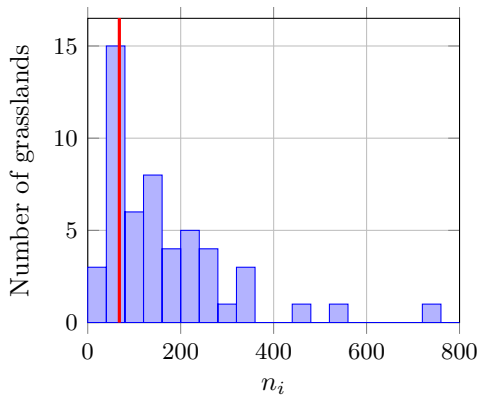


Figure 6. Histogram of grasslands size in number of pixels n_i . The red line corresponds to the number of spectro-variables $d = 68$ in 2013.

the spatial resolution of the pixel is too coarse, see for instance [37], [38]. In this representation, a sample is a pixel. Therefore, with each \mathbf{x}_{ik} is associated the response variable y_i of g_i , but \mathbf{x}_{ik} is processed independently of all other $\mathbf{x}_{ik'}$ of g_i . However, this representation usually leads to aberrant classification results (*e.g.*, salt and pepper effect) [36], which are not expected when working at the grassland scale.

B. Object-scale

At the object-scale, the mean vector $\boldsymbol{\mu}_i$ of the pixels belonging to g_i is generally used to represent g_i . It is estimated empirically by:

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

In this case, a vector $\hat{\boldsymbol{\mu}}_i \in \mathbb{R}^d$ and a response variable $y_i \in \mathbb{R}$ are associated with each grassland. This representation might be limiting for a heterogeneous object such as grasslands since the spectro-temporal variability is not encoded. To illustrate this bias, Fig. 7 shows on the left the set of pixels values in the NIR band for two grasslands. From this figure, it can be seen that if the mean vector captures the average behavior, higher variability can be captured by including the variance/covariance (right plot). The figure shows that the first and second eigenvectors of the covariance matrix capture well the general trend in the grassland and the main variations due to different phenologic behaviors in the grassland. This information cannot be recovered by considering the variance feature only, covariance must also be included.

In this study, to account for the spectro-temporal variability, we assume that the density function of pixels \mathbf{x}_i is, conditionally to grassland g_i , a Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, where $\boldsymbol{\Sigma}_i$ is the covariance matrix estimated empirically by:

$$\hat{\boldsymbol{\Sigma}}_i = \frac{1}{n_i - 1} \sum_{k=1}^{n_i} (\mathbf{x}_{ik} - \hat{\boldsymbol{\mu}}_i)(\mathbf{x}_{ik} - \hat{\boldsymbol{\mu}}_i)^\top. \quad (2)$$

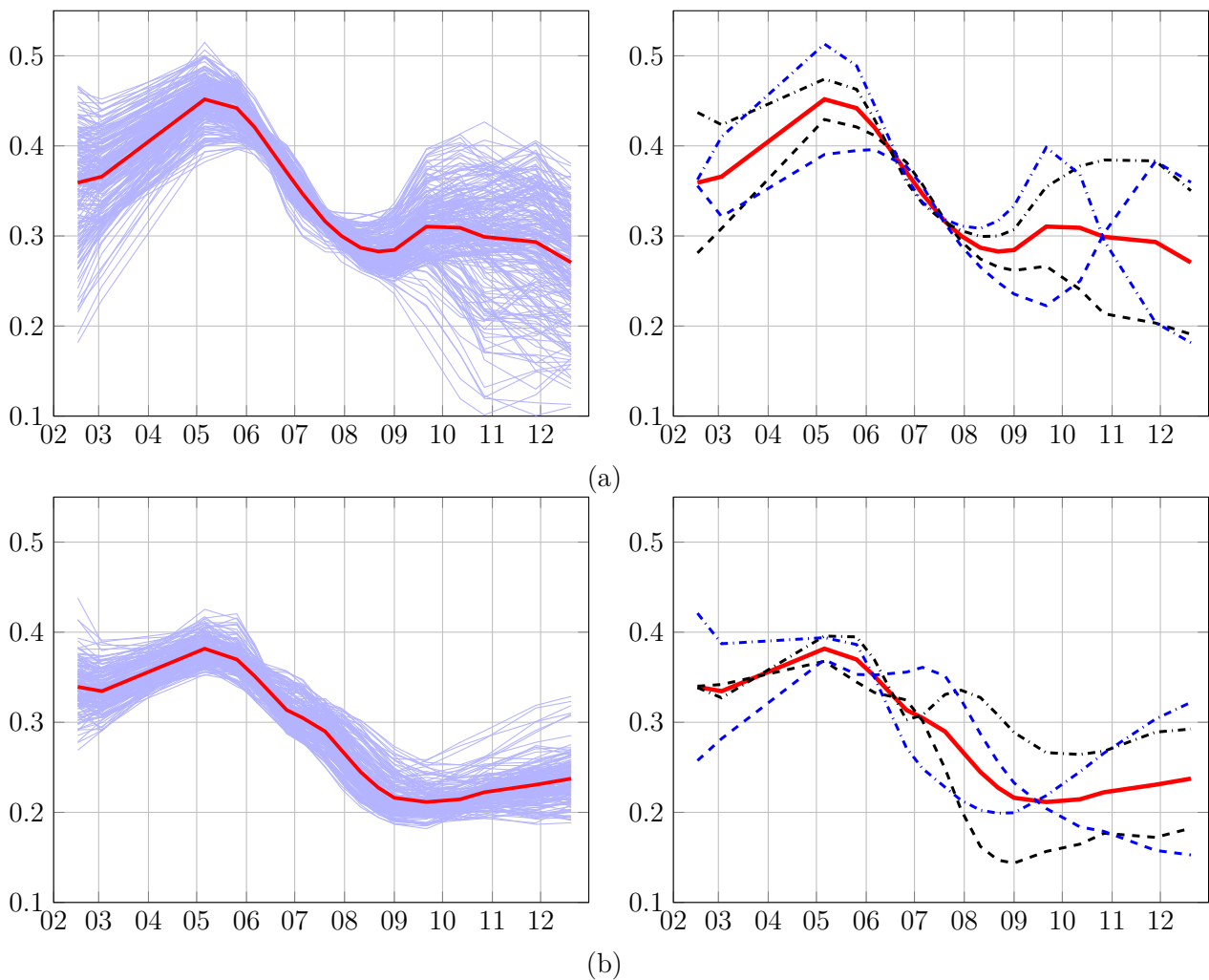


Figure 7. Examples of 2013 time series evolution in the NIR reflectance band of Formosat-2 for a grassland of class (a) "mowing" and (b) "grazing". The x-axis is the month of the year. The y-axis is the NIR reflectance value. The plot on the left shows the evolution of all the pixels in the grassland and the temporal mean of these pixels in red. The plot on the right shows the temporal mean in red, the temporal mean ± 0.2 the first eigenvector in black and the temporal mean ± 0.2 the second eigenvector in blue.

In this case, we associate with each g_i its estimated distribution $\mathcal{N}(\hat{\boldsymbol{\mu}}_i, \hat{\boldsymbol{\Sigma}}_i)$ and a response variable $y_i \in \mathbb{R}$. The Gaussian modeling encodes first and second order information on the grassland by exploiting the variance-covariance information. It is worth noting that if we constraint $\hat{\boldsymbol{\Sigma}}_i = \mathbf{I}_d$, the identity matrix of size d , for $i \in [1, \dots, G]$, the Gaussian modeling is reduced to the mean vector. In the following, $\mathcal{N}(\hat{\boldsymbol{\mu}}_i, \hat{\boldsymbol{\Sigma}}_i)$ is denoted \mathcal{N}_i .

IV. SIMILARITY MEASURE

A. Similarity measure between distributions

For classification purposes, a similarity measure between each pair of grasslands is required. With pixel-based or mean modeling approaches, conventional kernel methods such as Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel can be used since the explanatory variable is a vector. However for a

Gaussian modeling, *i.e.*, when the explanatory variable is a distribution, specific derivations are required to handle the probability distribution as an explanatory variable.

Many similarity functions generally used to compare two Gaussian distributions (*e.g.*, Kullback-Leibler divergence [63], [64], Jeffries-Matusita distance which is based on Bhattacharyya distance [65]) require the inversion of the covariance matrices and the computation of their determinants. For a conventional multivariate Gaussian model, the number of parameters to estimate for each grassland is $d(d+3)/2$. In the case where d is large, the number of parameters to estimate can be much larger than the number of samples, making ill-posed the inverse problem. This issue is faced in this study because grasslands are small elements of the landscape. They are characterized by a number of spectro-temporal variables which is about the same order than the number of pixels n_i (see Fig. 6). Therefore, most of the estimated covariance matrices are non-invertible and their determinants are null. Hence, conventional similarity measures used for moderate dimensional Gaussian distributions are not suitable for high dimensional Gaussian distributions.

B. Kernel functions between distributions

Mean map kernels are defined to operate on distributions [66]. They have been used in remote sensing for semi-supervised pixel-based learning in [67]. The authors define the similarity between two distributions p_i and p_j as the average of all pairwise kernel evaluations over the available realizations of p_i and p_j (*i.e.*, pixels that belong to grasslands g_i or g_j). It corresponds to the *empirical mean kernel* [67, eq.(8)]:

$$K^e(p_i, p_j) = \frac{1}{n_i n_j} \sum_{l,m=1}^{n_i, n_j} k(\mathbf{x}_{il}, \mathbf{x}_{jm}), \quad (3)$$

where n_i and n_j are the number of pixels associated with p_i and p_j respectively, \mathbf{x}_{il} is the l^{th} realization of p_i , \mathbf{x}_{jm} is the m^{th} realization of p_j and k is a semi-definite positive kernel function.

It is possible to include prior knowledge about the distributions by considering the *generative mean kernel* [66]:

$$K^g(p_i, p_j) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} k(\mathbf{x}, \mathbf{x}') \hat{p}_i(\mathbf{x}) \hat{p}_j(\mathbf{x}') d\mathbf{x} d\mathbf{x}'. \quad (4)$$

Note that eq. (3) acts on the realizations of p_i while eq. (4) acts on its estimation. When dealing with a large number of samples, the latter can drastically reduce the computational load with respects to the former.

In our grassland modeling, p_i and p_j are assumed to be Gaussian distributions. In that case, if k is a Gaussian kernel such as $k(\mathbf{x}, \mathbf{x}') = \exp(-\frac{\gamma}{2} \|\mathbf{x} - \mathbf{x}'\|^2)$, eq. (4) reduces to the so-called *Gaussian mean kernel* [68]:

$$K^G(\mathcal{N}_i, \mathcal{N}_j) = \frac{\exp \left\{ -0.5 (\hat{\boldsymbol{\mu}}_i - \hat{\boldsymbol{\mu}}_j)^T (\hat{\boldsymbol{\Sigma}}_i + \hat{\boldsymbol{\Sigma}}_j + \gamma^{-1} \mathbf{I}_d)^{-1} (\hat{\boldsymbol{\mu}}_i - \hat{\boldsymbol{\mu}}_j) \right\}}{|\hat{\boldsymbol{\Sigma}}_i + \hat{\boldsymbol{\Sigma}}_j + \gamma^{-1} \mathbf{I}_d|^{0.5}}, \quad (5)$$

where γ is a positive regularization parameter coming from the Gaussian kernel k and $|\cdot|$ stands for the determinant.

This kernel is not normalized, *i.e.*, $K^G(\mathcal{N}_i, \mathcal{N}_i) \neq 1$, but the normalization can be achieved easily:

$$\begin{aligned} \tilde{K}^G(\mathcal{N}_i, \mathcal{N}_j) &= \frac{K^G(\mathcal{N}_i, \mathcal{N}_j)}{K^G(\mathcal{N}_i, \mathcal{N}_i)^{0.5} K^G(\mathcal{N}_j, \mathcal{N}_j)^{0.5}} \\ &= K^G(\mathcal{N}_i, \mathcal{N}_j) |2\hat{\Sigma}_i + \gamma^{-1}\mathbf{I}_d|^{0.25} |2\hat{\Sigma}_j + \gamma^{-1}\mathbf{I}_d|^{0.25}. \end{aligned} \quad (6)$$

With respect to the Kullback-Leibler divergence (KLD) and the Jeffries-Matusita distance (JMD), the Gaussian mean kernel introduces a *ridge regularization* term $\gamma^{-1}\mathbf{I}_d$ in the computation of the inverse and of the determinant [69]. Thus, the Gaussian mean kernel is more suitable to measure the similarity in a high dimensional space than KLD and JMD. The value of γ controls the level of regularization. It is tuned during the training process as a conventional kernel parameter.

However, in the case of very small grasslands, two problems remain. The first lies in the ridge regularization: in this case, so low γ values are selected that it becomes too much regularized and it deteriorates the information. The second problem is that the estimation of the covariance matrix has a large variance when the number of samples used for the estimation is lower than the number of variables. Therefore, the covariance matrix becomes a poorly informative feature. In the following, we propose a new kernel function that allows to weight the covariance features with respect to the mean features.

C. α -Gaussian Mean Kernel

Depending on the level of heterogeneity and the size of the grassland, the covariance matrix could be more or less important for the classification process. We propose a kernel including an additional parameter which allows to weight the influence of the covariance matrix. It is done by the α -generative mean kernel:

$$K^\alpha(p_i, p_j) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} k(\mathbf{x}, \mathbf{x}') \hat{p}_i(\mathbf{x})^{(\alpha-1)} \hat{p}_j(\mathbf{x}')^{(\alpha-1)} d\mathbf{x} d\mathbf{x}'. \quad (7)$$

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

$$\begin{aligned} \tilde{K}^\alpha(\mathcal{N}_i, \mathcal{N}_j) &= \\ &= \frac{\exp \left\{ -0.5(\hat{\boldsymbol{\mu}}_i - \hat{\boldsymbol{\mu}}_j)^T \left(\alpha(\hat{\Sigma}_i + \hat{\Sigma}_j) + \gamma^{-1}\mathbf{I}_d \right)^{-1} (\hat{\boldsymbol{\mu}}_i - \hat{\boldsymbol{\mu}}_j) \right\}}{|\alpha(\hat{\Sigma}_i + \hat{\Sigma}_j) + \gamma^{-1}\mathbf{I}_d|^{0.5}} |2\alpha\hat{\Sigma}_i + \gamma^{-1}\mathbf{I}_d|^{0.25} |2\alpha\hat{\Sigma}_j + \gamma^{-1}\mathbf{I}_d|^{0.25}, \end{aligned} \quad (8)$$

where α is a positive parameter. The proof is given in the appendix. It is interesting to note that particular values of α and γ lead to known results:

- 1) $\alpha = 0$: In this case, eq. (8) reduces to the Gaussian kernel between the mean vectors. It becomes therefore equivalent to an object modeling where only the mean is considered.
- 2) $\alpha = 1$: It corresponds to the Gaussian mean kernel defined in eq. (6).
- 3) $\alpha \rightarrow +\infty$: We get a distance which works only on the covariance matrices. It is therefore equivalent to an object modeling where only the covariance is considered.
- 4) $\gamma \rightarrow +\infty$ and $\alpha = 2$: The α -Gaussian mean kernel simplifies to a RBF kernel built with the Bhattacharyya distance computed between \mathcal{N}_i and \mathcal{N}_j .

This proposed kernel includes several similarity measures known in the literature. Furthermore, new similarity measures can be defined by choosing different parameters configuration. The α -Gaussian mean kernel is therefore more flexible since it can adapt to the classification constraints:

- Whether the heterogeneity of the object is relevant or not,
- Whether the ratio between the number of pixels and the number of variables is high or low.

V. EXPERIMENTS ON CLASSIFICATION OF MANAGEMENT PRACTICES

In this section, the experimentations are detailed. We first introduce the implemented competitive methods, then the classification protocol is described and we finally present and discuss the results.

A. Competitive methods

Methods used to compare the effectiveness of the proposed method are presented below. They all rely on SVM.

- 1) *Pixel-based and mean modeling*: These conventional methods use a RBF kernel.
 - **PMV** (Pixel Majority Vote): The pixel-based method was described in section III-A. It classifies each pixel with no *a priori* information on the object which the pixel belongs to. In order to compare to other object-scaled methods, one class label is extracted per grassland by a majority vote done among the pixels belonging to the same grassland.
 - **μ** (mean): The distribution of the pixels reflectance of g_i is modeled by its mean vector μ_i (see section III-B).
- 2) *Divergence methods*: These methods are based on distance D between two Gaussian distributions. They are used in a Gaussian kernel such as $K_D(\mathcal{N}_i, \mathcal{N}_j) = \exp(-\frac{D_{ij}^2}{\sigma})$, with $\sigma > 0$:

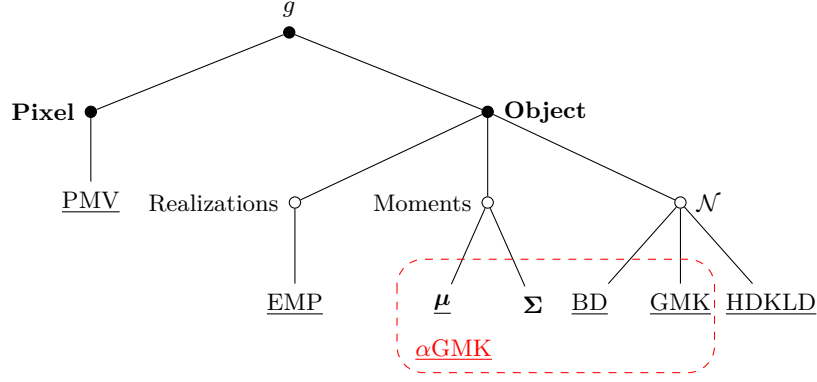


Figure 8. Contribution of the proposed method in grassland analysis for supervised classification. α GMK consists in a general modeling of the grassland at the object-scale and it encompasses several known modelings. The underlined methods are tested in this study.

Table II
METHODS USED IN THIS STUDY.

Method	PMV	EMP	μ	HDKLD	BD	GMK	α GMK
Scale	Pixel	Object	Object	Object	Object	Object	Object
Expl. variable	\mathbf{x}_{ik}	\mathbf{x}_{ik}	μ_i	\mathcal{N}_i	\mathcal{N}_i	\mathcal{N}_i	\mathcal{N}_i
Kernel	RBF	RBF	RBF	K_{HDKLD}	K_{B}	\tilde{K}^G	\tilde{K}^α
Parameters	σ, C	σ, C	σ, C	σ, C	σ, C	γ, C	γ, α, C
Nb of samples	8628	8628	52	52	52	52	52

- **HDKLD** (High Dimensional Kullback-Leibler Divergence): This method uses the Kullback-Leibler divergence for normal distributions with a regularization on covariance matrices such as described in [70].
- **BD** (Bhattacharyya Distance): This method uses the Bhattacharyya distance in the case of normal distributions :

$$B(\mathcal{N}_i, \mathcal{N}_j) = \frac{1}{8}(\hat{\mu}_i - \hat{\mu}_j)^\top \left(\frac{\hat{\Sigma}_i + \hat{\Sigma}_j}{2} \right)^{-1} (\hat{\mu}_i - \hat{\mu}_j) + \frac{1}{2} \ln \left(\frac{|\frac{\hat{\Sigma}_i + \hat{\Sigma}_j}{2}|}{|\hat{\Sigma}_i|^{0.5} |\hat{\Sigma}_j|^{0.5}} \right).$$

Small eigenvalues of the covariance matrices have been shrunk to the value 10^{-5} to make the computation tractable.

3) *Mean map kernel based methods*: These methods are based on mean maps kernels presented in section IV:

- **EMP**: This method uses the empirical mean map kernel of eq. (3).
- **GMK** (Gaussian Mean Kernel): This method is based on the normalized Gaussian mean kernel (eq. 6).
- **α GMK** (α -Gaussian Mean Kernel): This method is based on the normalized α -Gaussian mean kernel (eq. 8).

Fig. 8 illustrates the relationships between the different methods. The characteristics of each method are synthesized in Table II.

Table III
PARAMETERS TESTED FOR EACH METHOD DURING CROSS-VALIDATION.

Method	Parameters values
PMV	$\sigma \in \{2^{-17}, 2^{-16}, \dots, 2^{-10}\}$
EMP	$\sigma \in \{2^{-18}, 2^{-17}, \dots, 2^{-10}\}$
μ	$\sigma \in \{2^{-18}, 2^{-17}, \dots, 2^{-10}\}$
GMK	$\gamma \in \{2^{-17}, 2^{-18}, \dots, 2^{-10}\}$
HDKLD	$\sigma \in \{2^{15}, 2^{16}, \dots, 2^{25}\}$
BD	$\sigma \in \{2^{10}, 2^{11}, \dots, 2^{18}\}$
α GMK	$\gamma \in \{2^{-18}, 2^{-17}, \dots, 2^{-13}\}$
	$\alpha \in \{0, 10^{-3}, 10^{-2}, 10^{-1}, 0.3, 0.5, 0.7, 0.9, 1, 2, 5, 10, 15, 20, 25\}$

B. Classification protocol

We compared the efficiency in terms of classification accuracy and processing time of all the presented methods by classifying the "management practices" dataset (section II).

For each method, a SVM was used to classify the time series of year 2013 and of year 2014 independently. The dataset was split randomly into training and testing dataset (75% for training and 25% for testing), preserving the initial proportions of each class. 100 Monte Carlo runs were performed by selecting the same individuals for all the methods given a repetition.

For each repetition, the optimal parameters were tuned by cross-validation based on the best F1 score. Table III contains the grid search for all the methods. Note that a wide grid was searched for the parameter α of α GMK to further analyze the distribution of selected values. The penalty parameter C of the SVM process was chosen empirically and fixed to $C = 10$, after running several simulations. The classification accuracy for each repetition was assessed by the Kappa coefficient and the F1 score computed from the confusion matrix. The overall accuracy (OA) was computed but it is not presented here, because it does not reflect well the accuracy of the classification as it is an unbalanced dataset. In order to compare each pair of methods, a Wilcoxon rank-sum test was processed on the two distributions of the 100 Kappa coefficients.

The kernels and the SVM were implemented in Python through the Scikit library [71].

C. Results

The boxplot of Kappa and of F1 score distributions over the 100 repetitions are presented in Fig. 9. The processing times for each method for year 2013 are shown in Fig. 10.

In terms of classification accuracy, methods based on divergences (BD and HDKLD) provided the worst results. It is however worth noting that compared to BD, the high dimensional method developed in [70] works slightly better. Pixel-based methods, mean modeling method and mean generative kernel methods

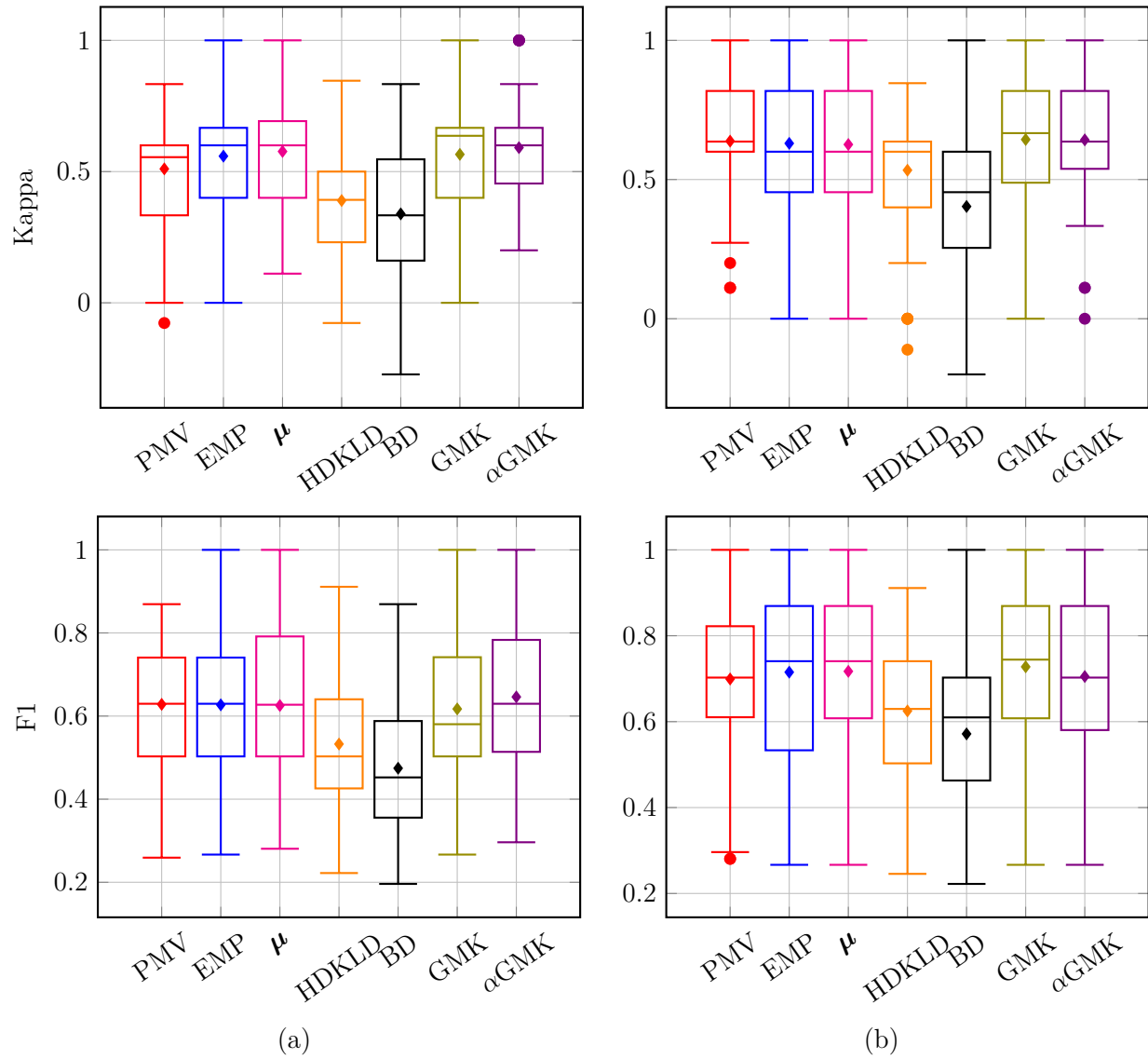


Figure 9. Boxplot of Kappa and F1 score repartitions for year (a) 2013 and (b) 2014. The line in the box stands for the median whereas the dot stands for the mean.

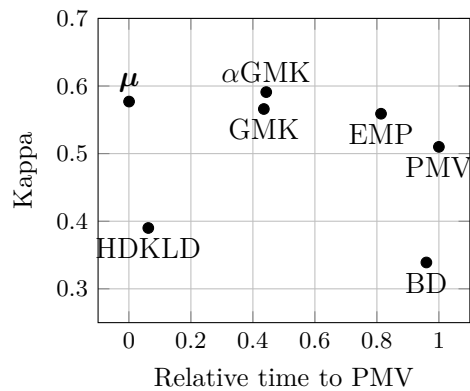


Figure 10. Relative training processing times to PMV and average Kappa of each method for the time series of 2013.

provided similar results in terms of OA, Kappa and F1 score, except for PMV which was significantly worse than the others for the year 2013. α GMK provided the highest values but it was not significantly better.

In terms of computational time (Fig. 10), the pixel-based methods required the largest processing times. BD was also very long, mainly because of the shrinkage procedure. Mean modeling was the fastest, followed closely by HDKLD. α GMK and GMK were equivalent in terms of computational times.

D. Discussion

The purpose of this work was to develop a model suitable for the classification of grasslands management practices from dense SITS and robust to the dimension of the data. The proposed method was compared to several competitive methods.

In this configuration, the divergence methods (BD and HDKLD) showed that they were not robust enough to a high dimensional space. Nevertheless, HDKLD, which applies a regularization on covariance matrices, worked better than BD which does not. Note that BD was successfully used in remote sensing for feature selection in hyperspectral data [72]. However, in that case, it was not evaluated in full dimension, unlike in our application.

Pixel-based methods are the most demanding in terms of computational time and they do not scale well with the number of pixels. Indeed, they have to process N pixels instead of G grasslands with $G \ll N$. Although they provided results close to the best results, using them on a large area might be difficult.

In addition, representing grasslands by the estimated distribution of their set of pixels decreased the complexity during the SVM process. Therefore, object-scaled methods offer a lower computational load when compared to empirical mean kernels and pixel-based methods.

Despite the mean modeling and the mean generative kernel methods (GMK and α GMK) performed similarly well, the best results were given by α GMK for the year 2013. The proposed method proved to be the most stable method in terms of variance of Kappa coefficient and F1 score over the two years.

In this context, including the covariance information helps to discriminate grasslands. However, if the dimensionality is not properly handled, it deteriorates the process (*e.g.*, BD and HDKLD). In this case, it is preferable to use the mean values only. α GMK offers the possibility to weight the influence of the covariance information in front of the mean. As a result, it provided at least as good results as the mean modeling and as GMK, since it encompasses both.

It is furthermore interesting to analyze the optimal values of α found during the cross-validation and the median of associated Kappa coefficients (Fig. 11). The highest Kappa coefficients were reached for higher values of $\hat{\alpha}$. The worst Kappa were obtained with $\hat{\alpha} < 1$. It shows the importance of the covariance information

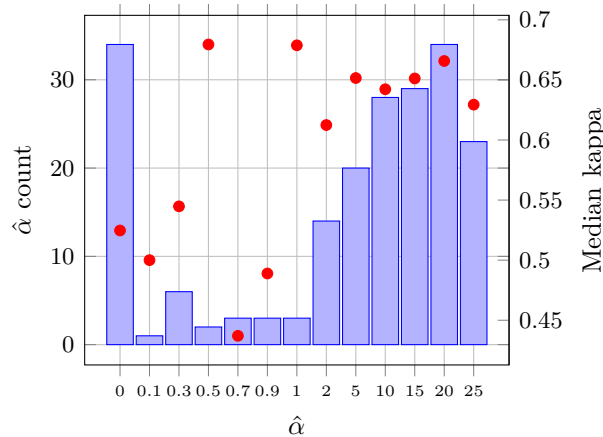


Figure 11. Bar plot of α values chosen by cross-validation and the median of associated Kappa coefficients (red dots), 2013 and 2014 taken together. NB: The values $\hat{\alpha} = 10^{-3}$ and $\hat{\alpha} = 10^{-2}$ were never selected.

in grasslands modeling: the heterogeneity in a grassland must be accounted for and it is not entirely well represented by the mean only.

Therefore, the choice of modeling grasslands pixels distribution by a Gaussian distribution makes sense in this context. It is particularly well appropriate for semi-natural grasslands, which are very heterogeneous, contrary to crops or annual "artificial" grasslands which can be assimilated to crops. Nevertheless, in the proposed kernel, this modeling was made flexible by regularizing the weight given to the covariance matrix. Indeed, grassland modeling by the mean only could work in other contexts (*i.e.*, young sown grasslands which are more homogeneous). Moreover, some grasslands are so small that the covariance matrix is too badly estimated. α GMK benefits from its high level of adaptability in front of the object configuration: no choice has to be made between a Gaussian or a mean modeling since the method encompasses both. It also includes several object-scaled methods known in the literature. However, this is at the cost of one more parameter to tune. Therefore, the classification process takes more time than GMK for instance.

Above all, although it is the first application of generative mean kernels in remote sensing classification, the α -Gaussian mean kernel proved its efficiency and stability in these experiments.

Overall, the results for time series of 2014 are better than for 2013. It might be explained by the acquisition dates in the time series. Although 2014 has less images, more clear images were acquired during Spring compared to 2013 (Fig.3). Indeed, many studies showed that the best season to discriminate grasslands is during the growing season [15], [47], [50], [51]. Spring is the period of the vegetation cycle where the management practices begin. Therefore, it is easier to differentiate the practices during this period. It might thus affect the accuracy of the classification of year 2013.

In general, the results could also be enhanced by removing some winter images which can have a negative influence on the entire annual time series [38]. However, the scope of this study was to develop a method

which is able to use a given time series, without having to process a date selection.

As described in the introduction, few studies have been carried out on the analysis of semi-natural grasslands using high spatio-spectro-temporal resolutions SITS. Usually, methods were pixel-based and applied on a few images or on a precise date selection to avoid dealing with the high dimension of data [40], [42], [47]. Schuster *et al.* [49] successfully classified grassland habitat using 21 Rapid-Eye images on a pixel basis, but there was no mention of the processing times. At the object-scale using a time series, grasslands were often represented by their mean NDVI, such as in [57], who noticed the difficulty to discriminate grasslands from crops because of mean NDVI similarities. The closest configuration might be the work of Zillman *et al.* [34], who used an object-based analysis and spectral reflectances combined with seasonal statistics of vegetation indices for mapping grasslands across Europe. The seasonal statistics were particularly relevant in the classification, because they captured well the spectral diversity of the grassland phenology. The use of these metrics could be considered for discriminating grassland management practices which impact on the phenology. The authors also concluded that the object-based analysis improves the classification compared to a pixel-based classification. However, the objects were determined by segmentation.

Finally, the relatively small Kappa coefficients obtained among the best methods in this work (around 0.64 in average) might be explained by the classes definition itself. Indeed, as many times emphasized, semi-natural grasslands are characterized by their high level of heterogeneity. Therefore, there might be a large amount of intra-class variability because of grasslands diversity. This diversity might be stronger than the effect of the management practices themselves. With this dataset, it was not possible to make groups of grasslands based on their functional types and/or species richness prior to their management practices discrimination. This idea might be reconsidered for a future application of this method. Within the same frame of ideas, other criteria for grassland discrimination should be regarded: artificial/monospecific grasslands against semi-natural/species-rich grasslands, intensively used grasslands against extensively used grasslands...

E. Prediction on the grasslands of the land use database

To show the efficiency of α GMK, we classified all the grasslands from the French agricultural land use database (RPG) covered by the Formosat-2 time series in 2014. We selected the plots of "permanent grassland" and "temporary grasslands". After applying a negative buffer of 8m and rasterizing the polygons, we removed the plots representing less than 10 Formosat-2 pixels. In the end, there were 797 grassland plots covered by the extent of Formosat-2 for a total of 252,472 pixels.

We trained the SVM on the whole field data (section II-C) using the same grid search as in the experiments. The parameters chosen after cross-validation based on F1 score were $\hat{\alpha} = 5$ and $\hat{\gamma} = 2^{-15}$. Then, the model

was used to predict the management practices of the 797 grasslands of the land use database.

The classification accuracy could not be assessed since the true labels of the grasslands are not known. However, as described in the study site, a spatial distribution of the classes could be expected. Indeed, grazed and mixed grasslands should be found in the south-west of the site, whereas more mown grasslands should be in the north.

An extract of the classification result is shown in Fig. 12. It represents the classified grasslands in their raster format. As expected, most of the grazed and mixed grasslands are located in the south-west of the image, whereas the north of the image is mostly composed of mown grasslands. Therefore, α GMK was very likely able to classify with an acceptable accuracy the grasslands management practices without any *a priori* spatial information. However, specific care should be considered, as not all the possible management practices were predicted. For instance, grasslands mown twice or grasslands not used were not in the training dataset, but it does not mean these managements do not exist in the rest of the data. The method deserves to be tested with an exhaustive grassland typology to produce more detailed grasslands maps.

In terms of processing times, the proposed method is able to classify 800 grasslands, representing more than 250,000 pixels, at the object scale from a high spatial resolution SITS within a few seconds on a conventional personal computer.

VI. CONCLUSION AND PROSPECTS

This study aimed at developing a statistical model for the classification of grasslands using satellite image time series with a high number of spectro-temporal variables. A grassland modeling at the object scale was proposed. To deal with the grasslands heterogeneity and their small size, the pixels distribution in a grassland was modeled by a Gaussian distribution. Then, to measure the similarity between two grasslands, *i.e.*, two Gaussian distributions, a kernel function based on mean maps was introduced. This kernel named α -Gaussian mean kernel, is robust to the high dimensionality of the data. The proposed method was compared to pixel-based and object-based existing classification methods for the supervised classification of grassland management practices using SVM. The α -Gaussian mean kernel provided the highest Kappa coefficients, but it was not significantly better than the best methods. However, it performed at least as well as the mean modeling method and as the Gaussian mean kernel because it encompasses both. In terms of processing times, these object-based methods were much faster than pixel-based methods.

Several contributions have been made in this work. The first lies in the grasslands pixels distributions modeling at the object-scale. A flexible kernel was proposed to encompass both Gaussian and mean modeling of grasslands, so no choice has to be made between these two modelings. It can therefore be used on very

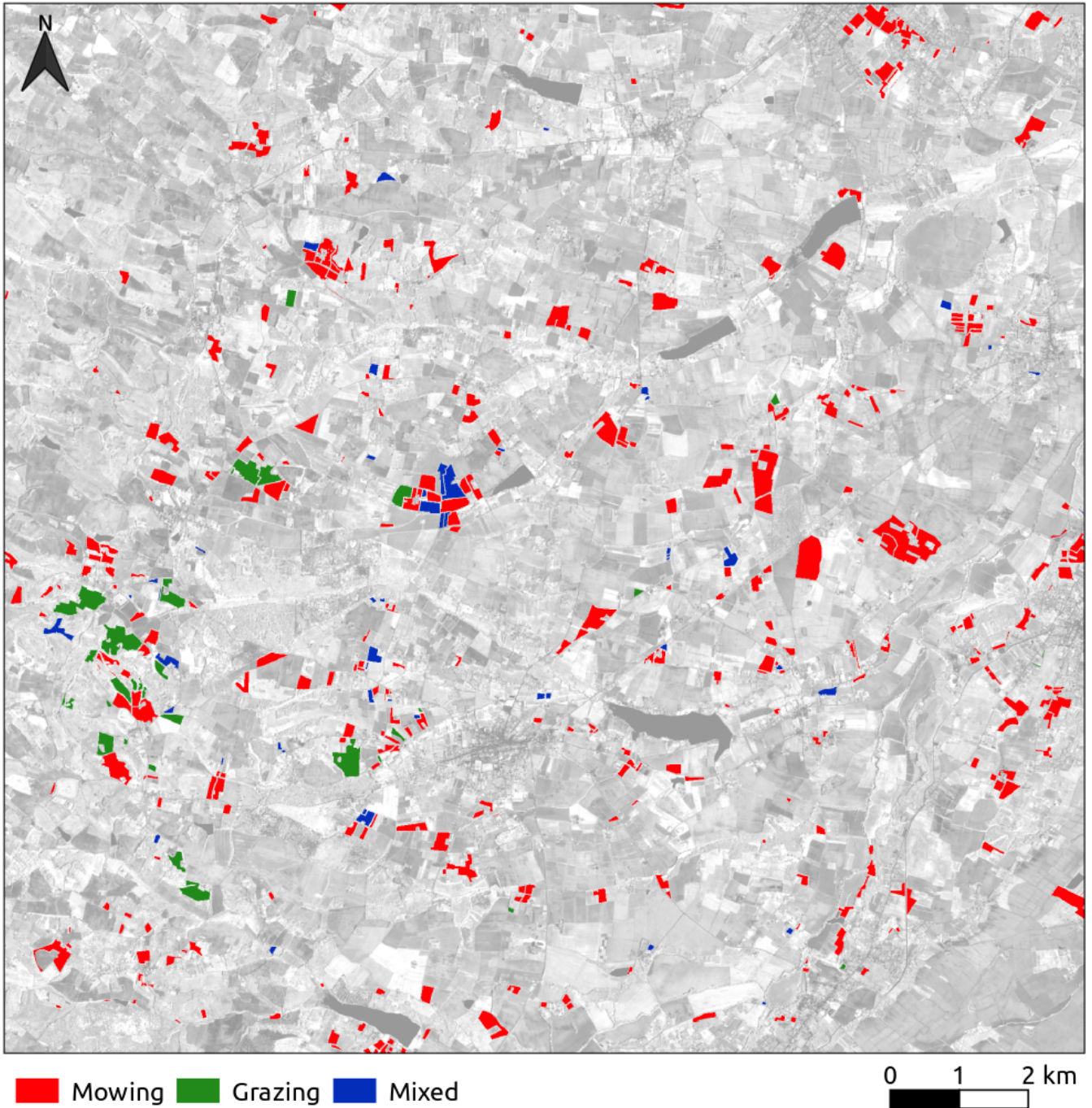


Figure 12. Classification of the grasslands from the French agricultural land use database (RPG) in 2014. The background is a May, 2014 Formosat-2 image in the NIR channel.

homogeneous objects such as artificial grasslands or on heterogeneous semi-natural grasslands. The second contribution is that this kernel is suitable for high dimensional data in a small ground sample size context. It allows to use all the multispectral data instead of a single vegetation index. Also, it can be used on a whole time series without dates selection. Indeed, this new kernel offers very low computational load. It can therefore be applied on a large dataset. With this kernel, we were able to process and to classify more than 250,000 pixels on a conventional personal computer within a few seconds. Even if it is the first application of generative mean kernels in remote sensing classification, the α -Gaussian mean kernel proved its efficiency and stability in these experiments.

The α -Gaussian mean kernel deserves to be tested on a larger dataset and with more balanced classes. Seasonal statistics could be used in the dataset to improve the representation of grassland phenology. These ideas will be considered in the future. This method was designed to deal with the dense SITS which will be provided by Sentinel-2 and to efficiently produce maps from this type of data. Other applications of the method are still possible (*e.g.*, small and heterogeneous objects such as peatlands, urban areas...).

APPENDIX

Proof of eq. (8): First, let us write the Gaussian distribution p_i to the power of α^{-1} :

$$\begin{aligned}
p_i(\mathbf{x})^{\alpha^{-1}} &= \frac{1}{(2\pi)^{d/2\alpha}} \times \frac{1}{|\boldsymbol{\Sigma}_i|^{1/2\alpha}} \times \exp \left\{ -0.5(\mathbf{x} - \boldsymbol{\mu}_i)^\top (\alpha\boldsymbol{\Sigma}_i)^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) \right\} \\
&= \frac{(2\pi)^{\frac{d}{2}(1-\frac{1}{\alpha})}}{(2\pi)^{d/2}} \times \alpha^{1/2} \times \frac{|\boldsymbol{\Sigma}_i|^{\frac{1}{2}(1-\frac{1}{\alpha})}}{|\alpha\boldsymbol{\Sigma}_i|^{1/2}} \times \exp \left\{ -0.5(\mathbf{x} - \boldsymbol{\mu}_i)^\top (\alpha\boldsymbol{\Sigma}_i)^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) \right\} \\
&= \alpha^{1/2} (2\pi)^{\frac{d}{2}(1-\frac{1}{\alpha})} |\boldsymbol{\Sigma}_i|^{\frac{1}{2}(1-\frac{1}{\alpha})} \times p(\mathbf{x}|\boldsymbol{\mu}_i, \alpha\boldsymbol{\Sigma}_i) \\
&= C(\boldsymbol{\Sigma}_i, \alpha) p(\mathbf{x}|\boldsymbol{\mu}_i, \alpha\boldsymbol{\Sigma}_i)
\end{aligned} \tag{9}$$

Then, including eq. (9) in eq. (7), we get:

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

which is eq. (5) with the covariance matrix of the Gaussian distribution scaled with α . The constants $C(\boldsymbol{\Sigma}_i, \alpha)$ and $C(\boldsymbol{\Sigma}_j, \alpha)$ are removed when normalizing the kernel and we get eq. (8). \blacksquare

ACKNOWLEDGEMENTS

The authors would like to thank CNES and CESBIO for providing Kalideos dataset.

REFERENCES

- [1] A. Eriksson, O. Eriksson, and H. Berglund, “Species abundance patterns of plants in Swedish semi-natural pastures,” *Ecography*, vol. 18, no. 3, pp. 310–317, 1995.
- [2] C. Gardi, M. Tomaselli, V. Parisi, A. Petraglia, and C. Santini, “Soil quality indicators and biodiversity in northern italian permanent grasslands,” *European Journal of Soil Biology*, vol. 38, no. 1, pp. 103 – 110, 2002.
- [3] C. Critchley, M. Burke, and D. Stevens, “Conservation of lowland semi-natural grasslands in the UK: a review of botanical monitoring results from agri-environment schemes,” *Biological Conservation*, vol. 115, no. 2, pp. 263 – 278, 2004.
- [4] C. A. Sullivan, M. S. Skeffington, M. J. Gormally, and J. A. Finn, “The ecological status of grasslands on lowland farmlands in western ireland and implications for grassland classification and nature value assessment,” *Biological Conservation*, vol. 143, no. 6, pp. 1529 – 1539, 2010.
- [5] B. P. Werling, T. L. Dickson, R. Isaacs, H. Gaines, C. Gratton, K. L. Gross, H. Liere, C. M. Malmstrom, T. D. Meehan, L. Ruan, B. A. Robertson, G. P. Robertson, T. M. Schmidt, A. C. Schrottenboer, T. K. Teal, J. K. Wilson, and D. A. Landis, “Perennial grasslands enhance biodiversity and multiple ecosystem services in bioenergy landscapes,” *Proceedings of the National Academy of Sciences*, vol. 111, no. 4, pp. 1652–1657, 2014.
- [6] M. Hansson and H. Fogelfors, “Management of a semi-natural grassland; results from a 15-year-old experiment in southern Sweden,” *Journal of Vegetation Science*, vol. 11, no. 1, pp. 31–38, 2000.
- [7] S. Kahmen, P. Poschlod, and K.-F. Schreiber, “Conservation management of calcareous grasslands. Changes in plant species composition and response of functional traits during 25 years,” *Biological Conservation*, vol. 104, no. 3, pp. 319 – 328, 2002.
- [8] D. Moog, P. Poschlod, S. Kahmen, and K.-F. Schreiber, “Comparison of species composition between different grassland management treatments after 25 years,” *Applied Vegetation Science*, vol. 5, no. 1, pp. 99–106, 2002.
- [9] H. Zechmeister, I. Schmitzberger, B. Steurer, J. Peterseil, and T. Wrba, “The influence of land-use practices and economics on plant species richness in meadows,” *Biological Conservation*, vol. 114, no. 2, pp. 165 – 177, 2003.
- [10] S. Plantureux, A. Peeters, and D. McCracken, “Biodiversity in intensive grasslands: Effect of management, improvement and challenges,” *Agronomy Research*, vol. 3, no. 2, pp. 153–164, 2005.
- [11] S. Muller, “Appropriate agricultural management practices required to ensure conservation and biodiversity of environmentally sensitive grassland sites designated under natura 2000,” *Agriculture, Ecosystems & Environment*, vol. 89, no. 3, pp. 261 – 266, 2002.
- [12] E. Allan, P. Manning, F. Alt, J. Binkenstein, S. Blaser, N. Blüthgen, S. Böhm, F. Grassein, N. Hölzel, V. H. Klaus, T. Kleinebecker, E. K. Morris, Y. Oelmann, D. Prati, S. C. Renner, M. C. Rillig, M. Schaefer, M. Schloter, B. Schmitt, I. Schöning, M. Schrupf, E. Solly, E. Sorkau, J. Steckel, I. Steffen-Dewenter, B. Stempfhuber, M. Tschapka, C. N. Weiner, W. W. Weisser, M. Werner, C. Westphal, W. Wilcke, and M. Fischer, “Land use intensification alters ecosystem multifunctionality via loss of biodiversity and changes to functional composition,” *Ecology Letters*, vol. 18, no. 8, pp. 834–843, 2015.
- [13] S. Soliveres, P. Manning, D. Prati, M. M. Gossner, F. Alt, H. Arndt, V. Baumgartner, J. Binkenstein, K. Birkhofer, S. Blaser, N. Blüthgen, S. Boch, S. Böhm, C. Börschig, F. Buscot, T. Diekötter, J. Heinze, N. Hölzel, K. Jung, V. H. Klaus, A.-M. Klein, T. Kleinebecker, S. Klemmer, J. Krauss, M. Lange, E. K. Morris, J. Müller, Y. Oelmann, J. Overmann, E. Pašalić, S. C. Renner, M. C. Rillig, H. M. Schaefer, M. Schloter, B. Schmitt, I. Schöning, M. Schrupf, J. Sikorski, S. A. Socher, E. F. Solly, I. Sonnemann, E. Sorkau, J. Steckel, I. Steffan-Dewenter, B. Stempfhuber, M. Tschapka, M. Türke, P. Venter, C. N. Weiner, W. W. Weisser, M. Werner, C. Westphal, W. Wilcke, V. Wolters, T. Wubet, S. Wurst, M. Fischer, and

- E. Allan, “Locally rare species influence grassland ecosystem multifunctionality,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 371, no. 1694, p. 20150269, 2016.
- [14] N. Pettorelli, W. F. Laurance, T. G. O’Brien, M. Wegmann, H. Nagendra, and W. Turner, “Satellite remote sensing for applied ecologists: opportunities and challenges,” *Journal of Applied Ecology*, vol. 51, no. 4, pp. 839–848, 2014.
- [15] I. Ali, F. Cawkwell, E. Dwyer, B. Barrett, and S. Green, “Satellite remote sensing of grasslands: from observation to management,” *Journal of Plant Ecology*, vol. 9, no. 6, pp. 649–671, 2016.
- [16] A. C. Newton, R. A. Hill, C. Echeverría, D. Golicher, J. M. Rey Benayas, L. Cayuela, and S. A. Hinsley, “Remote sensing and the future of landscape ecology,” *Progress in Physical Geography*, vol. 33, no. 4, pp. 528–546, 2009.
- [17] Y. Gu, B. K. Wylie, and N. B. Bliss, “Mapping grassland productivity with 250-m eMODIS NDVI and SSURGO database over the greater platte river basin, USA,” *Ecological Indicators*, vol. 24, pp. 31 – 36, 2013.
- [18] Z. Li, T. Huffman, B. McConkey, and L. Townley-Smith, “Monitoring and modeling spatial and temporal patterns of grassland dynamics using time-series MODIS NDVI with climate and stocking data,” *Remote Sensing of Environment*, vol. 138, pp. 232 – 244, 2013.
- [19] Y. Gu and B. K. Wylie, “Developing a 30-m grassland productivity estimation map for central Nebraska using 250-m MODIS and 30-m Landsat-8 observations,” *Remote Sensing of Environment*, vol. 171, pp. 291 – 298, 2015.
- [20] M. A. Friedl, J. Michaelsen, F. W. Davis, H. Walker, and D. S. Schimel, “Estimating grassland biomass and Leaf Area Index using ground and satellite data,” *International Journal of Remote Sensing*, vol. 15, no. 7, pp. 1401–1420, 1994.
- [21] B. Wylie, D. Meyer, L. Tieszen, and S. Mannel, “Satellite mapping of surface biophysical parameters at the biome scale over the north american grasslands: A case study,” *Remote Sensing of Environment*, vol. 79, no. 2–3, pp. 266 – 278, 2002. Recent Advances in Remote Sensing of Biophysical Variables.
- [22] R. Darvishzadeh, A. Skidmore, M. Schlerf, C. Atzberger, F. Corsi, and M. Cho, “LAI and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 63, no. 4, pp. 409 – 426, 2008.
- [23] Y. He, X. Guo, and J. F. Wilmshurst, “Reflectance measures of grassland biophysical structure,” *International Journal of Remote Sensing*, vol. 30, no. 10, pp. 2509–2521, 2009.
- [24] S. Asam, H. Fabritius, D. Klein, C. Conrad, and S. Dech, “Derivation of leaf area index for grassland within alpine upland using multi-temporal RapidEye data,” *International Journal of Remote Sensing*, vol. 34, no. 23, pp. 8628–8652, 2013.
- [25] S. Schmidtlein and J. Sassin, “Mapping of continuous floristic gradients in grasslands using hyperspectral imagery,” *Remote Sensing of Environment*, vol. 92, no. 1, pp. 126 – 138, 2004.
- [26] J. Ishii, S. Lu, S. Funakoshi, Y. Shimizu, K. Omasa, and I. Washitani, “Mapping potential habitats of threatened plant species in a moist tall grassland using hyperspectral imagery,” *Biodiversity and Conservation*, vol. 18, no. 9, pp. 2521–2535, 2009.
- [27] F. Fava, G. Parolo, R. Colombo, F. Gusmeroli, G. D. Marianna, A. Monteiro, and S. Bocchi, “Fine-scale assessment of hay meadow productivity and plant diversity in the european alps using field spectrometric data,” *Agriculture, Ecosystems & Environment*, vol. 137, no. 1–2, pp. 151 – 157, 2010. Special section Harvested perennial grasslands: Ecological models for farming’s perennial future.
- [28] J. Oldeland, D. Wesuls, D. Rocchini, M. Schmidt, and N. Jürgens, “Does using species abundance data improve estimates of species diversity from remotely sensed spectral heterogeneity?,” *Ecological Indicators*, vol. 10, no. 2, pp. 390 – 396, 2010.
- [29] H. Feilhauer, U. Faude, and S. Schmidtlein, “Combining isomap ordination and imaging spectroscopy to map continuous floristic gradients in a heterogeneous landscape,” *Remote Sensing of Environment*, vol. 115, no. 10, pp. 2513 – 2524, 2011.

- [30] M. C. Duniway, J. W. Karl, S. Schrader, N. Baquera, and J. E. Herrick, "Rangeland and pasture monitoring: an approach to interpretation of high-resolution imagery focused on observer calibration for repeatability," *Environmental Monitoring and Assessment*, vol. 184, no. 6, pp. 3789–3804, 2012.
- [31] S. Punalekar, A. Verhoef, I. V. Tatarenko, C. van der Tol, D. M. J. Macdonald, B. Marchant, F. Gerard, K. White, and D. Gowing, "Characterization of a highly biodiverse floodplain meadow using hyperspectral remote sensing within a plant functional trait framework," *Remote Sensing*, vol. 8, no. 2, p. 112, 2016.
- [32] T. Hilker, E. Natsagdorj, R. H. Waring, A. Lyapustin, and Y. Wang, "Satellite observed widespread decline in Mongolian grasslands largely due to overgrazing," *Global Change Biology*, vol. 20, no. 2, pp. 418–428, 2014.
- [33] R. Cao, J. Chen, M. Shen, and Y. Tang, "An improved logistic method for detecting spring vegetation phenology in grasslands from MODIS EVI time-series data," *Agricultural and Forest Meteorology*, vol. 200, pp. 9 – 20, 2015.
- [34] E. Zillmann, A. Gonzalez, E. J. M. Herrero, J. van Wolvelaer, T. Esch, M. Keil, H. Weichelt, and A. M. Garzón, "Pan-european grassland mapping using seasonal statistics from multisensor image time series," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, pp. 3461–3472, Aug 2014.
- [35] H. Nagendra, "Using remote sensing to assess biodiversity," *International Journal of Remote Sensing*, vol. 22, no. 12, pp. 2377–2400, 2001.
- [36] T. Blaschke, G. J. Hay, M. Kelly, S. Lang, P. Hofmann, E. Addink, R. Q. Feitosa, F. van der Meer, H. van der Werff, F. van Coillie, and D. Tiede, "Geographic object-based image analysis – towards a new paradigm," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 87, pp. 180 – 191, 2014.
- [37] I. Poças, M. Cunha, and L. S. Pereira, "Dynamics of mountain semi-natural grassland meadows inferred from SPOT-VEGETATION and field spectroradiometer data," *International Journal of Remote Sensing*, vol. 33, no. 14, pp. 4334–4355, 2012.
- [38] A. Halabuk, M. Mojses, M. Halabuk, and S. David, "Towards detection of cutting in hay meadows by using of NDVI and EVI time series," *Remote Sensing*, vol. 7, no. 5, p. 6107, 2015.
- [39] R. Lucas, A. Rowlands, A. Brown, S. Keyworth, and P. Bunting, "Rule-based classification of multi-temporal satellite imagery for habitat and agricultural land cover mapping," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 62, no. 3, pp. 165 – 185, 2007.
- [40] T. Toivonen and M. Luoto, "Landsat TM images in mapping of semi-natural grasslands and analysing of habitat pattern in an agricultural landscape in south-west finland," *Fennia - International Journal of Geography*, vol. 181, no. 1, pp. 49 – 67, 2003.
- [41] H. Nagendra, R. Lucas, J. P. Honrado, R. H. Jongman, C. Tarantino, M. Adamo, and P. Mairota, "Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats," *Ecological Indicators*, vol. 33, pp. 45 – 59, 2013. Biodiversity Monitoring.
- [42] K. P. Price, X. Guo, and J. M. Stiles, "Optimal Landsat TM band combinations and vegetation indices for discrimination of six grassland types in eastern kansas," *International Journal of Remote Sensing*, vol. 23, no. 23, pp. 5031–5042, 2002.
- [43] J. A. Gamon, C. B. Field, D. A. Roberts, S. L. Ustin, and R. Valentini, "Airbone imaging spectrometry functional patterns in an annual grassland during an AVIRIS overflight," *Remote Sensing of Environment*, vol. 44, no. 2, pp. 239 – 253, 1993.
- [44] C. Corbane, S. Lang, K. Pipkins, S. Alleaume, M. Deshayes, V. E. G. Millán, T. Strasser, J. V. Borre, S. Toon, and F. Michael, "Remote sensing for mapping natural habitats and their conservation status – new opportunities and challenges," *International Journal of Applied Earth Observation and Geoinformation*, vol. 37, pp. 7 – 16, 2015. Special Issue on Earth observation for habitat mapping and biodiversity monitoring.

- [45] M. A. Wulder, R. J. Hall, N. C. Coops, and S. E. Franklin, "High spatial resolution remotely sensed data for ecosystem characterization," *BioScience*, vol. 54, no. 6, pp. 511–521, 2004.
- [46] O. Buck, V. E. G. Millán, A. Klink, and K. Pakzad, "Using information layers for mapping grassland habitat distribution at local to regional scales," *International Journal of Applied Earth Observation and Geoinformation*, vol. 37, pp. 83 – 89, 2015. Special Issue on Earth observation for habitat mapping and biodiversity monitoring.
- [47] J. Franke, V. Keuck, and F. Siegert, "Assessment of grassland use intensity by remote sensing to support conservation schemes," *Journal for Nature Conservation*, vol. 20, no. 3, pp. 125 – 134, 2012.
- [48] P. Duseux, F. Vertès, T. Corpetti, S. Corgne, and L. Hubert-Moy, "Agricultural practices in grasslands detected by spatial remote sensing," *Environmental Monitoring and Assessment*, vol. 186, no. 12, pp. 8249–8265, 2014.
- [49] C. Schuster, T. Schmidt, C. Conrad, B. Kleinschmit, and M. Förster, "Grassland habitat mapping by intra-annual time series analysis – comparison of RapidEye and TerraSAR-X satellite data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 34, pp. 25 – 34, 2015.
- [50] A. Psomas, M. Kneubuhler, S. Huber, K. Itten, and N. E. Zimmermann, "Hyperspectral remote sensing for estimating aboveground biomass and for exploring species richness patterns of grassland habitats," *International Journal of Remote Sensing*, vol. 32, pp. 9007–9031, 2011.
- [51] M. J. Hill, "Vegetation index suites as indicators of vegetation state in grassland and savanna: An analysis with simulated SENTINEL-2 data for a North American transect," *Remote Sensing of Environment*, vol. 137, pp. 94 – 111, 2013.
- [52] M. Drusch, U. D. Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, B. Hoersch, C. Isola, P. Laberinti, P. Martimort, A. Meygret, F. Spoto, O. Sy, F. Marchese, and P. Bargellini, "Sentinel-2: ESA's optical high-resolution mission for GMES operational services," *Remote Sensing of Environment*, vol. 120, pp. 25 – 36, 2012. The Sentinel Missions - New Opportunities for Science.
- [53] A. S. Laliberte, E. L. Fredrickson, and A. Rango, "Combining decision trees with hierarchical object-oriented image analysis for mapping arid rangelands," *Photogrammetric engineering & Remote sensing*, vol. 73, no. 2, pp. 197–207, 2007.
- [54] J. C. Brenner, Z. Christman, and J. Rogan, "Segmentation of Landsat Thematic Mapper imagery improves buffelgrass (*pennisetum ciliare*) pasture mapping in the sonoran desert of mexico," *Applied Geography*, vol. 34, pp. 569 – 575, 2012.
- [55] S. Stenzel, F. E. Fassnacht, B. Mack, and S. Schmidtlein, "Identification of high nature value grassland with remote sensing and minimal field data," *Ecological Indicators*, vol. 74, pp. 28 – 38, 2017.
- [56] J. Evans and R. Geerken, "Classifying rangeland vegetation type and coverage using a Fourier component based similarity measure," *Remote Sensing of Environment*, vol. 105, no. 1, pp. 1 – 8, 2006.
- [57] T. Esch, A. Metz, M. Marconcini, and M. Keil, "Combined use of multi-seasonal high and medium resolution satellite imagery for parcel-related mapping of cropland and grassland," *International Journal of Applied Earth Observation and Geoinformation*, vol. 28, pp. 230 – 237, 2014.
- [58] D. C. Duro, S. E. Franklin, and M. G. Dubé, "A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery," *Remote Sensing of Environment*, vol. 118, pp. 259 – 272, 2012.
- [59] D. Sheeren, M. Fauvel, V. Josipović, M. Lopes, C. Planque, J. Willm, and J.-F. Dejoux, "Tree species classification in temperate forests using Formosat-2 satellite image time series," *Remote Sensing*, vol. 8, no. 9, p. 734, 2016.
- [60] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 652–675, 2013.
- [61] O. Hagolle, M. Huc, D. Villa Pascual, and G. Dedieu, "A multi-temporal method for cloud detection, applied to FORMOSAT-

- 2, VENUS, LANDSAT and SENTINEL-2 images,” *Remote Sensing of Environment*, vol. 114, pp. 1747–1755, Aug. 2010.
- [62] P. H. C. Eilers, “A perfect smoother,” *Analytical Chemistry*, vol. 75, no. 14, pp. 3631–3636, 2003. PMID: 14570219.
- [63] S. Kullback and R. A. Leibler, “On information and sufficiency,” *Ann. Math. Statist.*, vol. 22, pp. 79–86, 03 1951.
- [64] S. Kullback, “Letter to the editor: The Kullback-Leibler distance,” *The American Statistician*, vol. 41, no. 4, pp. 340–341, 1987.
- [65] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 3rd ed., 1999.
- [66] N. A. Mehta and A. G. Gray, “Generative and latent mean map kernels,” *CoRR*, vol. abs/1005.0188, 2010.
- [67] L. Gomez-Chova, G. Camps-Valls, L. Bruzzone, and J. Calpe-Maravilla, “Mean map kernel methods for semisupervised cloud classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, pp. 207–220, Jan 2010.
- [68] K. Muandet, K. Fukumizu, F. Dinuzzo, and B. Schölkopf, “Learning from distributions via support measure machines,” in *Advances in Neural Information Processing Systems 25* (P. Bartlett, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, eds.), pp. 10–18, 2012.
- [69] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2005.
- [70] M. Lopes, M. Fauvel, S. Girard, and D. Sheeren, “High dimensional Kullback-Leibler divergence for grassland management practices classification from high resolution satellite image time series,” in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 3342–3345, July 2016.
- [71] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [72] L. Bruzzone, F. Roli, and S. B. Serpico, “An extension of the Jeffreys-Matusita distance to multiclass cases for feature selection,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 6, pp. 1318–1321, 1995.