



**HAL**  
open science

# Object-based classification of grasslands from high resolution satellite image time series using Gaussian mean map kernels

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 of grasslands from high resolution satellite image time series using Gaussian mean map kernels. *Remote Sensing*, 2017, 9 (7), pp.Article 688. 10.3390/rs9070688 . hal-01424929v3

**HAL Id: hal-01424929**

**<https://inria.hal.science/hal-01424929v3>**

Submitted on 13 Jun 2017

**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.



Distributed under a Creative Commons Attribution 4.0 International License

Article

# Object-based classification of grasslands from high resolution satellite image time series using Gaussian mean map kernels

Mailys Lopes<sup>1\*</sup>, Mathieu Fauvel<sup>1</sup>, Stéphane Girard<sup>2</sup> and David Sheeren<sup>1</sup>

<sup>1</sup> Dynafor, University of Toulouse, INRA, INPT, INPT-EI PURPAN, 31326 Castanet Tolosan, France; mailys.lopes@inra.fr, mathieu.fauvel@ensat.fr, david.sheeren@ensat.fr.

<sup>2</sup> Team Mistis, INRIA Rhône-Alpes, LJK, 38334 Montbonnot, France; stephane.girard@inria.fr.

\* Correspondence: mailys.lopes@inra.fr; Tel.: +33-534-32-39-27

Academic Editor: name

Version June 12, 2017 submitted to Remote Sens.

**Abstract:** This paper deals with the classification of grasslands 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 level by modeling its pixels distributions by a Gaussian distribution. To measure the similarity between two grasslands, a new kernel is proposed as a second contribution: the  $\alpha$ -Gaussian mean kernel. It allows to weight the influence of the covariance matrix when comparing two Gaussian distributions. This kernel is introduced in Support Vector Machine for the supervised classification of grasslands from south-west France. A dense intra-annual multispectral time series of Formosat-2 satellite is used for the classification of grasslands management practices, while an inter-annual NDVI time series of Formosat-2 is used for old and young grasslands discrimination. Results are compared to other existing pixel- and object-based approaches in terms of classification accuracy and processing time. The proposed method shows to be a good compromise between processing speed and classification accuracy. It can adapt to the classification constraints and it encompasses several similarity measures known in the literature. It is appropriate for the classification of small and heterogeneous objects such as grasslands.

**Keywords:** Supervised classification; SVM; Gaussian mean map kernels; kernel methods; object analysis; grasslands.

## 1. 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], which are linked to their plant and animal composition.

Different factors impact on grassland biodiversity conservation. Among them, the age of a grassland (*i.e.*, the time since last ploughing/sowing) is directly related to its plant and animal composition. Old "permanent" grasslands, often called semi-natural grasslands, hold a richer biodiversity than temporary grasslands [2,6–8]. Indeed, they had time to establish and stabilize their vegetation cover, contrarily to temporary grasslands which are part of a crop rotation. Additionally, agricultural management of grasslands (*i.e.*, mowing, grazing, fertilizing, reseeding...) influences their structure and composition [9–12]. Management is essential for their biodiversity conservation because it prevents from the woody establishment. Conversely, an intensive use constitutes a threat for this biodiversity [12,13]. Therefore, it is important to know the age of a grassland and to identify the

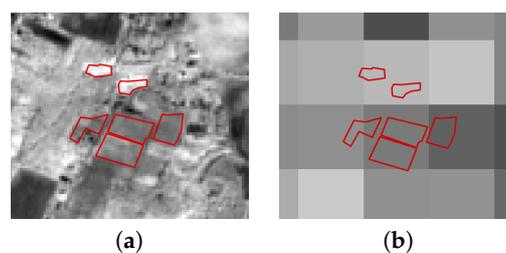
32 management practices in order to monitor their effect on biodiversity and related services. But these  
33 factors are defined at different temporal scales: over the years for the age of a grassland, and during a  
34 vegetation season (*i.e.*, a year) for the management practice.

35 Usually, ecologists and agronomists characterize grasslands at the parcel scale through field  
36 surveys. However, these surveys require important human and material resources, the knowledge of  
37 the assessor and a sampling strategy, which make them expensive and time-consuming [14]. They are  
38 thus limited in spatial extent and in temporal frequency, limiting grassland characterization to a local  
39 scale and over a short period of time.

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

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

53 Using satellite remote sensing images, grasslands have been studied a lot at a regional scale with  
54 medium spatial resolution sensors (*i.e.*, MODIS, 250m/pixel [17,18,32]), where the Minimum Mapping  
55 Unit (MMU) is at least of hundreds of meters. This scale is suitable for large, extensive, homogeneous  
56 and contiguous regions like steppes [33], but not for fragmented landscapes which are usually found  
57 in Europe and in France particularly [34,35]. These fragmented landscapes are made of a patchwork  
58 of different land covers which have a small area [35]. In these types of landscapes, grasslands can  
59 be smaller (less than 10,000m<sup>2</sup>) than the pixel resolution [36] (see Figure 1 for a graphical example).  
60 As a consequence, pixels containing grasslands are usually a mixture of other contributions, which  
61 can limit the analysis [37,38]. As examples, Poças *et al.* [39] had to select large contiguous areas of  
62 semi-natural grasslands in a mountain region of Portugal to be able to use SPOT-VEGETATION data  
63 (1-km resolution). Halabuk *et al.* [40] also had to select only one MODIS pixel per homogeneous  
64 sample site in Slovakia to detect cutting in hay meadows. A 30-m pixel resolution is still not sufficient  
65 for grassland characterization. Indeed, Lucas *et al.* [41] and Toivonen and Luoto [42] showed that it  
66 was more difficult to classify fragmented and complex elements [43], like semi-natural grasslands, than  
67 homogeneous habitats, using Landsat imagery. Price *et al.* [44] classified six grassland management  
68 types in Kansas using six Landsat images, but the accuracy of the classification was not satisfying (less  
69 than 70%). Therefore, to detect small grasslands in fragmented landscapes, high spatial resolution  
70 images are required [36,45,46].



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

71 For high spatial resolution images (about 10m/pixel), few intra-annual images are usually  
72 available for a given location [47]. However, Buck *et al.* [48] concluded that three RapidEye images per  
73 year were not enough to detect the mowing practices in grasslands. It was confirmed by Franke *et al.*  
74 [49] who classified grassland use intensity into four categories: Semi-natural Grassland, Extensively  
75 Used Grassland, Intensively Used Grassland and Tilled Grassland. They increased the classification  
76 accuracy when increasing the number of RapidEye images from three to five scenes. Additionally,  
77 Schmidt *et al.* [50] concluded that about seven to ten images, depending on the vegetation index used,  
78 are a good tradeoff between the amount of satellite data and classification accuracy of grassland use  
79 intensity. Some works report results with few images per year, such as Dusseux *et al.* [51], but they  
80 worked on LAI. In their study for mapping grassland habitat using intra-annual RapidEye imagery,  
81 Schuster *et al.* [52] concluded the more acquisition dates used, the better the mapping quality.

82 Given the heterogeneity of grasslands in fragmented landscapes, their phenological cycle and the  
83 punctuality of the anthropogenic events (*e.g.*, mowing), dense high spatial resolution intra-annual time  
84 series are necessary to identify the grassland management types [36,52–54]. Moreover, to discriminate  
85 semi-natural grasslands from temporary grasslands, inter-annual time series are necessary. Until  
86 recently, satellite missions offering high revisit frequency (1-16 days) had coarse spatial resolution  
87 (*i.e.*, NOAA AVHRR - 1km, MODIS - 250/500m). Conversely, high spatial resolution missions did  
88 not provide dense time series and/or were costly (*i.e.*, QuickBird, RapidEye). For these reasons  
89 and compared to crops, grasslands differentiation through Earth observations is still considered as a  
90 challenge [52]. However, new missions like Sentinel-2 [55], with very high revisit frequency (5 days)  
91 and high spatial resolution (10 meters in four spectral channels, 20 meters in six channels) provide  
92 new opportunities for grasslands monitoring over the years in fragmented landscapes [54] at no cost,  
93 thanks to the ESA free data access policy. For instance, the high spatial resolution is assumed to make  
94 possible the identification of grassland-only pixels in the image and several pixels can belong to the  
95 same grassland plot. Hence, the analysis can be done at the object level, not at the pixel level, which is  
96 suitable for landscape ecologists and agronomists who usually study grasslands at the parcel scale [56].  
97 Thus, object-oriented approaches are more likely to characterize grasslands ecologically [57,58]. Yet, a  
98 lot of works consider pixel-based approaches without any spatial constraints [17,42,44,48,49,52,59].

99 At the object level, grasslands are commonly represented by their mean NDVI [18]. But such  
100 representation might be too simple since it does not account for the heterogeneity in a grassland.  
101 Sometimes, distributions of pixels as individual observations are still better than the mean value  
102 to represent grasslands, as in [54]. Lucas *et al.* [41] used a rule-based method on segmented areas  
103 for habitat mapping but it did not work well on complex and heterogeneous land covers. Esch *et al.*  
104 [60] also used an object-oriented method on segmented elements then represented by their mean  
105 NDVI. These methods based on mean modeling do not capture well grasslands heterogeneity. Other  
106 representations can be found in the literature, taking standard deviation and object texture features as  
107 variables [61], but they were not applied to time series. In our knowledge, these methods do not use  
108 the high spatial and the high temporal resolutions jointly. Moreover, all these studies used vegetation  
109 indices as a variable, although it has been shown that classification results are better when using more  
110 spectral information [35,62].

111 To deal with the high spatio-spectro-temporal resolutions new satellite sensors are now offering,  
112 dimension reduction is usually performed through the use of a vegetation index such as NDVI [50,52,63,  
113 64], PCA [65] or spectro-temporal metrics [35,66]. But a large amount of spectro-temporal information  
114 is lost with these solutions. Franke *et al.* [49] developed an indicator of the spectral variability of a pixel  
115 over the time series, the Mean Absolute Spectral Dynamics, but its efficiency was assessed using a  
116 decision tree algorithm. Decision trees are usually not recommended because they tend to over-fit the  
117 data [67]. Therefore, the high spatio-spectro-temporal resolutions have not really been addressed in  
118 the literature of remote sensing classification. Indeed, such time series bring new methodological and  
119 statistical constraints given the high dimension of data (*i.e.*, number of pixels and number of spectral  
120 and temporal measurements). Dealing with more variables increases the number of parameters to

121 estimate, increasing the computation time and making the computation unstable (*i.e.*, ill-conditioned  
 122 covariance matrices...) [68,69]. Hence, conventional models are not appropriate if one wants to use  
 123 all the spectro-temporal information of time series with high spatial and temporal resolutions. Thus,  
 124 classifying grasslands with this type of data is still considered as a challenge [52].

125 In the present study, we introduce a model suitable for the classification of grasslands using  
 126 satellite image time series (SITS) with a high number of spectro-temporal variables (*e.g.*, Sentinel-2  
 127 data). Two temporal scales are considered in this work: (i) an inter-annual time series of three years to  
 128 discriminate **old** grasslands from **young** grasslands and (ii) an intra-annual time series to identify the  
 129 management practices. Note that in this work, the objects are not found from segmentation [38] but  
 130 from existing dataset in a polygon form.

131 The first contribution of this study is to model a grassland at the object **level** while accounting  
 132 for the spectral variability within a grassland. We consider that the distribution of the pixel spectral  
 133 reflectance in a given grassland can be modeled by a Gaussian distribution. The second contribution  
 134 is to propose a measure of similarity between two Gaussian distributions that is robust to the high  
 135 dimension of the data. This method is based on the use of covariance through mean maps. The  
 136 last contribution is the application of the method to **old and young** grasslands discrimination and of  
 137 management practices classification, which are non common applications in remote sensing. Moreover,  
 138 in our knowledge, mean maps have not yet been used on Gaussian distributions for supervised  
 139 classification of SITS at the object **level**.

140 In the next section, the materials used for the experimental part of this study are presented.  
 141 Then the methods, including the different types of grassland modeling and the measures of similarity  
 142 between distributions are introduced in section 3. Following that, we experiment the **proposed**  
 143 **methods** on the classification of a real dataset in section 4. Finally, conclusions and prospects are given  
 144 in section 5.

## 145 2. Materials

### 146 2.1. Study site

147 The study site is located in south-west France, near the city of Toulouse (about 30km), in a  
 148 semi-rural area (center coordinates:  $43^{\circ}27'36''\text{N } 1^{\circ}8'24''\text{E}$ , Figure 2). This region is characterized by a  
 149 temperate climate with oceanic and Mediterranean influences. The average annual precipitation is  
 150 656mm and the average temperature is  $13^{\circ}\text{C}$ . The north of the site, closer to the urban area of Toulouse,  
 151 is flat, whereas the south-west of the site is hilly. The eastern part corresponds to the Garonne river  
 152 floodplain and this location is dominated by crop production. Within this study site, livestock farming  
 153 is declining in favor of annual crop production. Grasslands are mostly used for forage or silage  
 154 production. Some grasslands, located in the south-western part of the area, are pastures for cattle or  
 155 sheep. The extent of the area is included in the satellite image extent (Figure 2) and is about  $24 \times 24\text{km}^2$ .

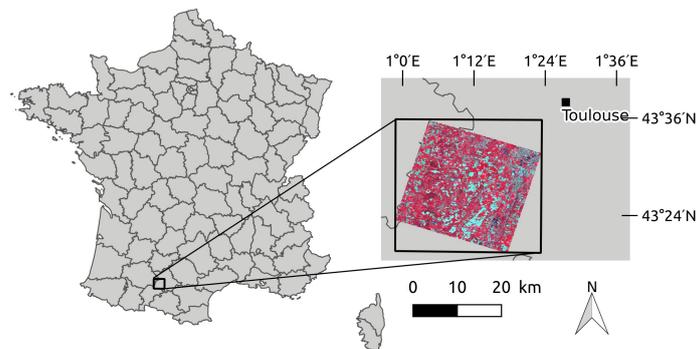


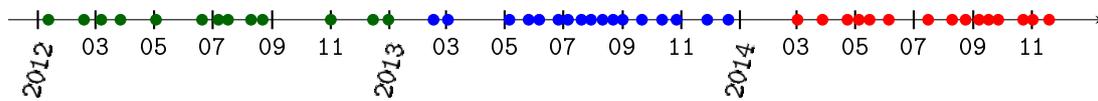
Figure 2. Study site location in south-west France. It is included in the satellite image extent.

## 156 2.2. Satellite data

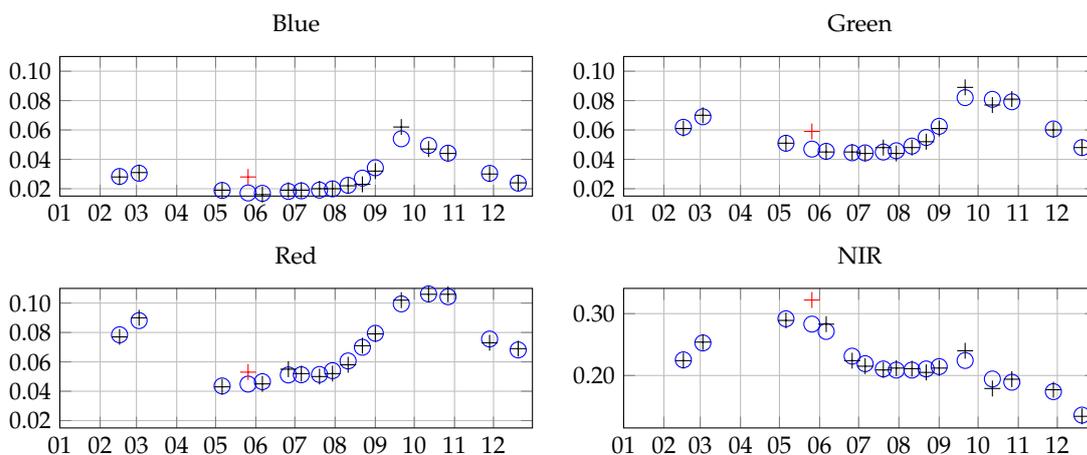
157 Time series of Formosat-2 were used in this experiment. Formosat-2 has four spectral bands with  
 158 an 8-meter spatial resolution: B1 "Blue" (0.45 - 0.52 $\mu\text{m}$ ), B2 "Green" (0.52 - 0.6 $\mu\text{m}$ ), B3 "Red" (0.63 -  
 159 0.69 $\mu\text{m}$ ), B4 "Near Infra-Red (NIR)" (0.76 - 0.9 $\mu\text{m}$ ). The extent of an acquisition is 24km  $\times$  24km. The  
 160 images were all acquired with the same viewing angle. They were orthorectified, radiometrically and  
 161 atmospherically corrected by the French Spatial Agency (CNES). They were provided by the Center  
 162 for the Study of the Biosphere from Space (CESBIO) in reflectance with a mask of clouds and shadows  
 163 issued from the MACCS (Multi-sensor Atmospheric Correction and Cloud Screening) processor [70],  
 164 in the frame of the Kalideos project.

165 For the inter-annual analysis, we used all the acquisitions of the consecutive years 2012 (13  
 166 observations), 2013 (17 observations) and 2014 (15 observations) (Figure 3 and Figure S1 in the  
 167 supplementary materials). The acquisitions of year 2013 and of year 2014 were used separately for the  
 168 classification of management practices.

169 To reconstruct the time series due to missing data (clouds and their shadows), the Whittaker  
 170 filter [71] was applied pixel-by-pixel on the reflectances in each spectral band for each year  
 171 independently. **The Whittaker filter is a non-parametric filter which has a smoothing parameter  
 172 that controls the roughness of the reconstructed curve. It has been successfully applied to smooth  
 173 NDVI time series in the literature [72–75].** The smoother was adapted for unequally spaced intervals  
 174 and accounted for missing data (see [62] for a detailed description of the method). The smoothing  
 175 parameter was the same for all the pixels. It was equal to  $10^5$  for year 2013 and to  $10^4$  for 2012 and  
 176 2014, after an ordinary cross-validation done on a subset of the pixels for each year. An example of  
 177 smoothing on a grassland pixel is provided in Figure 4. This pixel is hidden by a light cloud during  
 178 one image acquisition (red cross). Notice that the smoothing is done at the cost of under-estimating  
 179 the local maxima of the temporal profile.



**Figure 3.** Formosat-2 acquisition dates in 2012 (green dots), 2013 (blue dots) and 2014 (red dots) used in this experiment.



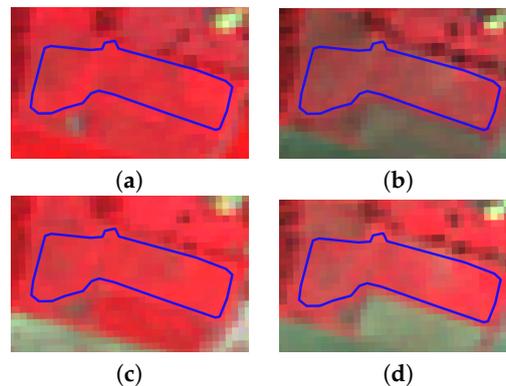
**Figure 4.** Example of time series reconstruction (blue dots) with Whittaker smoother for a pixel of a grassland in the four spectral bands. The black crosses correspond to the original 2013 Formosat-2 time series and the red ones correspond to missing/noisy data due to the clouds. The x-axis represents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

180 For the intra-annual time series, we used all the spectral information. Therefore, the smoothed  
 181 time series associated with each of the four spectral bands were concatenated to get a unique time  
 182 series per pixel. For the inter-annual time series, as using all the spectral bands would result in a too  
 183 large number of variables to process, we worked on the NDVI, computed from the red and NIR bands.

### 184 2.3. Reference data

#### 185 2.3.1. Old and young grasslands

186 In this study, "old" grasslands are 14 years old or more, whereas "young" grasslands are less  
 187 than five years old. The French agricultural land use database (Registre Parcellaire Graphique) was  
 188 used to extract the grasslands depending on their age. It registers on an annual basis the cultivated  
 189 areas declared by the farmers in a GIS. Grasslands are declared as "permanent" or "temporary".  
 190 Permanent grasslands are at least five years old, whereas temporary grasslands are less than five  
 191 years old (Commission Regulation EU No 796/2004). For every plot declared as a grassland in 2014,  
 192 its age was computed from the previous years declarations. We kept only the grasslands which  
 193 were at least 14 years old in 2014 ("old"), and the grasslands which were less than 5 years old in  
 194 2014 ("young"). A negative buffer of 8 meters was then applied to all the polygons to eliminate  
 195 the edge effects (Figure 5). Then they were rasterized using the GDAL command *gdal\_rasterize*  
 196 ([http://www.gdal.org/gdal\\_rasterize.html](http://www.gdal.org/gdal_rasterize.html)) to obtain the pixels inside each grassland. Only the  
 197 grasslands having an area of at least 1,000m<sup>2</sup> were kept to ensure a minimum number of 16 pixels to  
 198 represent each grassland. In the end, there were 59 old grasslands (at least 14 years old) and 416 young  
 199 grasslands (Table 1), for an average area of about 26,600m<sup>2</sup>.



**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.

**Table 1.** Composition of the old and young grasslands dataset.

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

#### 200 2.3.2. Management practices

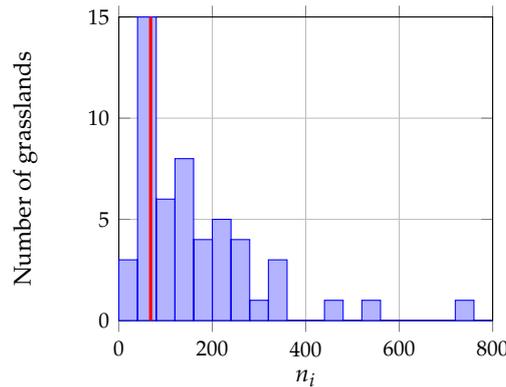
201 The information of the agricultural practices performed in the crops are not featured in the  
 202 land use database. Therefore, this dataset comes exclusively from field data. As mentioned in the

203 introduction, ground data is difficult to obtain in ecology since field work is fastidious. A field survey  
 204 was conducted in May 2015 to determine the past and current management practices of 52 grasslands  
 205 by interviewing the farmers or grasslands' owners. The practices remained stable for the years 2013  
 206 and 2014. Four management types during a vegetation cycle were identified: one mowing, two  
 207 mowings, grazing and mixed (mowing then grazing). We eliminated the type "two mowings" of the  
 208 dataset because of its under-representation (only three grasslands).

209 The management types were used as classes for the classification (Table 2). The grasslands were  
 210 digitalized manually after field work. A negative buffer of 8 meters was then applied to eliminate the  
 211 edge effects, before rasterizing the polygons. The average grasslands surface area is about 10,000m<sup>2</sup>.  
 212 The smallest grassland is 1,632m<sup>2</sup> (which represents 25 Formosat-2 pixels) and the largest is 47,111m<sup>2</sup>  
 213 (735 pixels) (Figure 6).

**Table 2.** Grassland management types and composition of the dataset.

Class	Nb of grasslands	Nb of pixels
Mowing	34	6,265
Grazing	10	1,193
Mixed	8	1,170
Total	52	8,628



**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.

### 214 3. Methods

#### 215 3.1. Grassland modeling

216 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  
 217 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  
 218 the total number of pixels,  $d = n_B n_T$  is the number of spectro-temporal variables,  $n_B$  is the number of  
 219 spectral bands and  $n_T$  is the number of temporal acquisitions. In the experimental part, when working  
 220 on the intra-annual time series of 2013 using the four spectral bands,  $d = 4 \times 17 = 68$ . In 2014,  $d =$   
 221  $4 \times 15 = 60$ . When working on the inter-annual times series using NDVI,  $d = 1 \times (13 + 17 + 15) = 45$ .  
 222 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  
 223 corresponds to its class label.

224 In the following, two types of grassland modeling are discussed, at the pixel **level** and at the  
 225 object **level**. A more informative object **level** modeling is then proposed. **Then, similarity measures are**  
 226 **discussed.**

### 227 3.1.1. Pixel level

228 The representation of a grassland at the pixel level has been used a lot in the remote sensing  
 229 literature [17,42,44,48,49,52,59]. The grassland can either be represented by all its pixels or by one pixel  
 230 when the spatial resolution of the pixel is too coarse, see for instance [39,40]. In this representation,  
 231 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  
 232 processed independently of all other  $\mathbf{x}_{ik'}$  of  $g_i$ . However, this representation usually leads to aberrant  
 233 classification results (*e.g.*, salt and pepper effect) [38], which are not expected when working at the  
 234 grassland level.

### 235 3.1.2. Object level

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

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

238 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.  
 239 This representation might be limiting for heterogeneous objects such as grasslands since the  
 240 spectro-temporal variability is not encoded. To illustrate this bias, Figure 7 shows on the left the  
 241 set of pixels values in the NIR band for two grasslands (a and b). From this figure, it can be seen that if  
 242 the mean vector captures the average behavior, higher variability can be captured by including the  
 243 variance/covariance (middle and right plots). The figure shows that the first and second eigenvectors  
 244 of the covariance matrix capture well the general trend in the grassland and the main variations due to  
 245 different phenologic behaviors in the grassland. This information cannot be recovered by considering  
 246 the variance feature only: covariance must also be included.

247 In this study, to account for the spectro-temporal variability, we assume that the distribution of  
 248 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  
 249 matrix estimated empirically by:

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

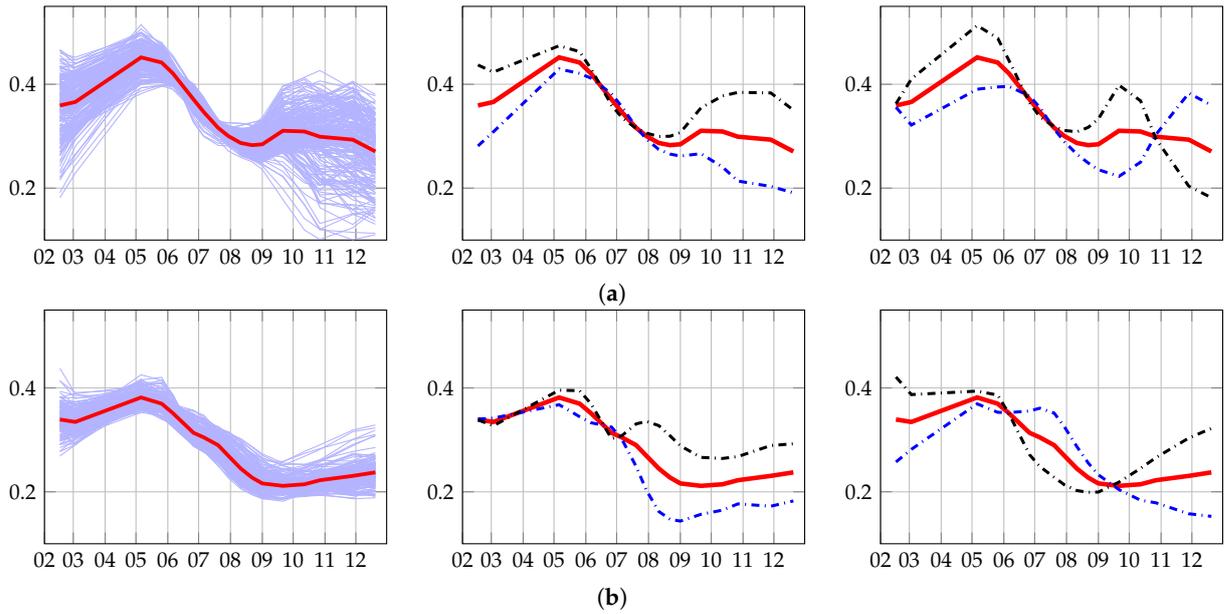
250 In this case, we associate with each  $g_i$  its estimated distribution  $\mathcal{N}(\hat{\boldsymbol{\mu}}_i, \hat{\boldsymbol{\Sigma}}_i)$  and a response variable  
 251  $y_i \in \mathbb{R}$ . The Gaussian modeling encodes first and second order information on the grassland by  
 252 exploiting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

## 255 3.2. Similarity measure

### 256 3.2.1. Similarity measure between distributions

257 For classification purposes, a similarity measure between each pair of grasslands is required.  
 258 With pixel-based or mean modeling approaches, conventional kernel methods such as Support Vector  
 259 Machine (SVM) with a Radial Basis Function (RBF) kernel can be used since the explanatory variable  
 260 is a vector. However for a Gaussian modeling, *i.e.*, when the explanatory variable is a distribution,  
 261 specific derivations are required to handle the probability distribution as an explanatory variable.

262 Many similarity functions generally used to compare two Gaussian distributions (*e.g.*,  
 263 Kullback-Leibler divergence [76], Jeffries-Matusita distance which is based on Bhattacharyya  
 264 distance [77]) require the inversion of the covariance matrices and the computation of their  
 265 determinants. For a conventional multivariate Gaussian model, the number of parameters to estimate



**Figure 7.** Examples of 2013 time series evolution in the NIR reflectance band of Formosat-2 for a grassland of management practice (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 in the middle shows the temporal mean in red, the temporal mean  $+0.2 \times$  the first eigenvector in blue and the temporal mean  $-0.2 \times$  the first eigenvector in black. The plot on the right shows the temporal mean in red, the temporal mean  $+0.2 \times$  the second eigenvector in blue and the temporal mean  $-0.2 \times$  the second eigenvector in black.

266 for each grassland is  $d(d+3)/2$  ( $d$  parameters for the mean vector and  $d(d+1)/2$  parameters for the  
 267 symmetric covariance matrix). In the case where  $d$  is large, the number of parameters to estimate can  
 268 be much larger than the number of samples, making the inverse problem ill-posed. This issue is faced  
 269 in this study because grasslands are small elements of the landscape. They are characterized by a  
 270 number of spectro-temporal variables which is about of the same order as the number of pixels  $n_i$  (see  
 271 Figure 6). Therefore, most of the estimated covariance matrices are singular and their determinants are  
 272 null. Hence, conventional similarity measures used for moderate dimensional Gaussian distributions  
 273 are not suitable for high dimensional Gaussian distributions. In the following, we propose to use  
 274 mean map kernels and we introduce a derivation of mean map kernels to weight the influence of the  
 275 covariance matrix.

### 276 3.2.2. Mean map kernels between distributions

277 Mean map kernels are similarity measures which operate on distributions [78]. They have  
 278 been used in remote sensing for semi-supervised pixel-based learning in [79]. In their work, the  
 279 authors define the similarity between two distributions  $p_i$  and  $p_j$  as the average of all pairwise kernel  
 280 evaluations over the available realizations of  $p_i$  and  $p_j$  (*i.e.*, pixels that belong to grasslands  $g_i$  or  $g_j$ ). It  
 281 corresponds to the *empirical mean kernel* [79, eq.(8)]:

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

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

284 It is possible to include prior knowledge on the distributions by considering the *generative mean*  
 285 *kernel* [78]:

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

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

289 In our grassland modeling,  $p_i$  and  $p_j$  are assumed to be Gaussian distributions. In that case, if  $k$  is  
 290 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*  
 291 *kernel* [80]:

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

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

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

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

295 With respect to the Kullback-Leibler divergence (KLD) and the Jeffries-Matusita distance (JMD),  
 296 the Gaussian mean kernel introduces a *ridge regularization* term  $\gamma^{-1} \mathbf{I}_d$  in the computation of the inverse  
 297 and of the determinant [81]. Thus, the Gaussian mean kernel is more suitable to measure the similarity  
 298 in a high dimensional space than KLD and JMD. The value of  $\gamma$  tunes the level of regularization. It is  
 299 tuned during the training process as a conventional kernel parameter.

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

### 307 3.2.3. $\alpha$ -Gaussian Mean Kernel

308 Depending on the level of heterogeneity and the size of the grassland, the covariance matrix could  
 309 be more or less important for the classification process. We propose a kernel including an additional  
 310 positive parameter  $\alpha$  which allows to weight the influence of the covariance matrix, the  $\alpha$ -*generative*  
 311 *mean kernel*:

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

312 When  $p_i$  and  $p_j$  are Gaussian distributions,  $k$  is a Gaussian kernel and the normalization is applied,  
 313 the expression gives rise to the  $\alpha$ -*Gaussian mean kernel*:

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

314 The proof is given in the appendix. It is interesting to note that particular values of  $\alpha$  and  $\gamma$  lead to  
315 known results:

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

323 This proposed kernel thus includes several similarity measures known in the literature.  
324 Furthermore, new similarity measures can be defined by choosing different parameters configuration.  
325 The  $\alpha$ -Gaussian mean kernel ( $\alpha$ GMK) is therefore more flexible since it can adapt to the classification  
326 constraints:

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

#### 329 4. Experiments on grasslands classification

330 In this section, the experiments for grassland classification are detailed. We first introduce the  
331 seven competitive methods, then the classification protocol is described and we finally present and  
332 discuss the results.

##### 333 4.1. Competitive methods

334 Several existing pixel-based and object-based classification methods using SVM are presented  
335 below. They are compared to assess the effectiveness of the proposed object-based method which relies  
336 on the weighted use of the covariance matrix,  $\alpha$ GMK, for the classification of grasslands.

##### 337 4.1.1. Pixel-based and mean modeling

338 These conventional methods use a RBF kernel.

- 339 • **PMV** (Pixel Majority Vote): The pixel-based method was described in section 3.1.1. It classifies  
340 each pixel with no *a priori* information on the object which the pixel belongs to. In order to  
341 compare to other object **level** methods, one class label is extracted per grassland by a majority  
342 vote done among the pixels belonging to the same grassland.
- 343 •  $\boldsymbol{\mu}$  (mean): The distribution of the pixels reflectance of  $g_i$  is modeled by its mean vector  $\boldsymbol{\mu}_i$  (see  
344 section 3.1.2).

##### 345 4.1.2. Divergence methods

346 These methods are based on a distance  $D$  between two Gaussian distributions. They are used in a  
347 Gaussian kernel such as  $K_D(\mathcal{N}_i, \mathcal{N}_j) = \exp(-\frac{D_D^2}{\sigma})$ , with  $\sigma > 0$ :

- 348 • **HDKLD** (High Dimensional Kullback-Leibler Divergence): This method uses the Kullback-Leibler  
349 divergence for Gaussian distributions with a regularization on covariance matrices such as  
350 described in [82].

- 351 • **BD** (Bhattacharyya Distance): This method uses the Bhattacharyya distance in the case of Gaussian  
 352 distributions:

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

353 Small eigenvalues of the covariance matrices are shrunk to the value  $10^{-5}$  to make the  
 354 computation tractable [83].

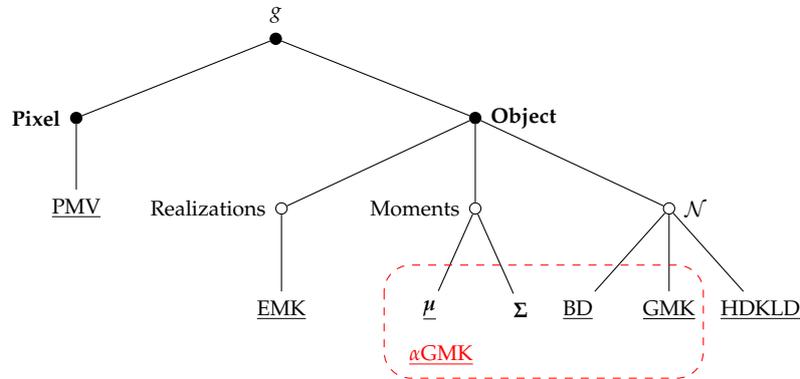
#### 355 4.1.3. Mean map kernel based methods

356 These methods are based on mean map kernels presented in section 3.2:

- 357 • **EMK** (Empirical Mean Kernel): This method uses the empirical mean map kernel of eq. (3) and it  
 358 is pixel-based.  
 359 • **GMK** (Gaussian Mean Kernel): This method is based on the normalized Gaussian mean kernel  
 360 (eq. 6).  
 361 •  **$\alpha$ GMK** ( $\alpha$ -Gaussian Mean Kernel): This method is based on the proposed normalized  $\alpha$ -Gaussian  
 362 mean kernel (eq. 8).

363 Figure 8 illustrates the relationships between the different methods. The characteristics of each  
 364 method are synthesized in Table 3.

365 For memory issues during the SVM process, the number of pixels processed for the **old and young**  
 366 grasslands classification was divided by 10 for the two methods based on pixels (PMV and EMK).  
 367 Only 1 pixel out of 10 was kept per grassland.



**Figure 8.** 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. **PMV, EMK and  $\mu$  are not based on Gaussian modeling while the others are.**

**Table 3.** Characteristics of the methods used in this study.

Method	PMV	EMK	$\mu$	HDKLD	BD	GMK	$\alpha$ GMK
<b>Level</b>	Pixel	Object	Object	Object	Object	Object	Object
<b>Expl. variable</b>	$\mathbf{x}_{ik}$	$\mathbf{x}_{ik}$	$\boldsymbol{\mu}_i$	$\mathcal{N}_i$	$\mathcal{N}_i$	$\mathcal{N}_i$	$\mathcal{N}_i$
<b>Kernel</b>	RBF	RBF	RBF	$K_{\text{HDKLD}}$	$K_B$	$\tilde{K}^G$	$\tilde{K}^\alpha$
<b>Parameters</b>	$\sigma, C$	$\sigma, C$	$\sigma, C$	$\sigma, C$	$\sigma, C$	$\gamma, C$	$\gamma, \alpha, C$
<b>Nb of samples</b>	16,250 / 8,628	16,250 / 8,628	475 / 52	475 / 52	475 / 52	475 / 52	475 / 52

## 368 4.2. Classification protocol

369 We compared the efficiency in terms of classification accuracy and processing time of all the  
370 presented methods by classifying the two grassland datasets on inter-annual and intra-annual time  
371 series (section 2).

372 For each method, a Monte Carlo procedure was performed on 100 runs. For each run, the dataset  
373 was split randomly into training and testing datasets (75% for training and 25% for testing), preserving  
374 the initial proportions of each class. The same grasslands were selected for a given Monte Carlo  
375 repetition regardless of the method.

376 During each repetition, the optimal parameters were tuned by cross-validation based on the best  
377 F1 score. Table 4 contains the parameters grid search for all the methods. Note that a wide grid was  
378 searched for the parameter  $\alpha$  of  $\alpha$ GMK to further analyze the distribution of selected values. The  
379 penalty parameter  $C$  of the SVM process was chosen empirically and fixed to  $C = 10$ , after running  
380 several simulations. The classification accuracy for each repetition was assessed by the F1 score  
381 computed from the confusion matrix. The overall accuracy (OA) was computed but it is not presented  
382 here, because it does not reflect well the accuracy of the classification since unbalanced datasets were  
383 used.

384 In order to compare each pair of methods, a Wilcoxon rank-sum test was processed on the pair of  
385 distributions of the 100 F1 scores. **This nonparametric test is designed for two independent samples  
386 that are not assumed to be normally distributed [84]. It tests if the two samples are drawn from  
387 populations having the same distribution.**

388 The kernels and the SVM were implemented in Python through the Scikit library [85].

**Table 4.** Parameters tested for each method during cross-validation.

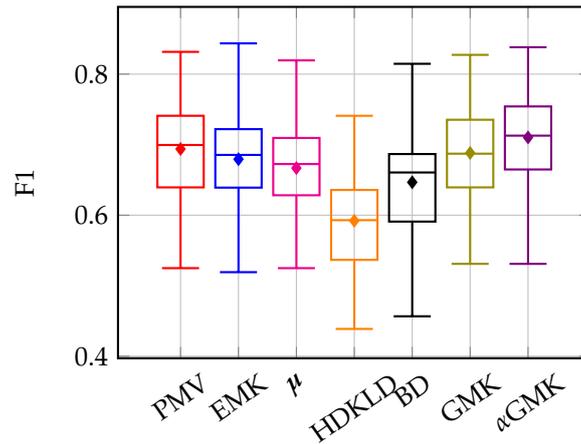
Method	Parameters values	
	Inter-annual analysis	Intra-annual analysis
PMV	$\sigma \in \{2^0, 2^1, \dots, 2^{10}\}$	$\sigma \in \{2^{-17}, 2^{-16}, \dots, 2^{-10}\}$
EMK	$\sigma \in \{2^0, 2^1, \dots, 2^{10}\}$	$\sigma \in \{2^{-18}, 2^{-17}, \dots, 2^{-10}\}$
$\mu$	$\sigma \in \{2^0, 2^1, \dots, 2^{10}\}$	$\sigma \in \{2^{-18}, 2^{-17}, \dots, 2^{-10}\}$
HDKLD	$\sigma \in \{2^{10}, 2^{11}, \dots, 2^{20}\}$	$\sigma \in \{2^{15}, 2^{16}, \dots, 2^{25}\}$
BD	$\sigma \in \{2^0, 2^1, \dots, 2^{10}\}$	$\sigma \in \{2^{10}, 2^{11}, \dots, 2^{18}\}$
GMK	$\gamma \in \{2^0, 2^1, \dots, 2^{10}\}$	$\gamma \in \{2^{-17}, 2^{-18}, \dots, 2^{-10}\}$
$\alpha$ GMK	$\gamma \in \{2^0, 2^1, \dots, 2^{10}\}$	$\gamma \in \{2^{-18}, 2^{-17}, \dots, 2^{-13}\}$
	$\alpha \in \{0, 0.1, 0.5, 1, 2, 5, 10, 15, 20, 25, 50\}$	$\alpha \in \{0, 10^{-3}, 10^{-2}, 10^{-1}, 0.3, 0.5, 0.7, 0.9, 1, 2, 5, 10, 15, 20, 25\}$

## 389 4.3. Results

### 390 4.3.1. Old and young grasslands: inter-annual time series

391 Figure 9 sums up the **old and young** grasslands classification results for each method over the  
392 100 repetitions as a boxplot of F1 scores. **The Kappa coefficients can be found in Figure S2 in the  
393 supplementary materials. Since the cross-validation was not based on the Kappa coefficient, the results  
394 are discussed in terms of F1 scores.** The method reaching the best scores is  $\alpha$ GMK with a F1 average of  
395 0.71 followed by PMV and GMK with an average of 0.69.

396 Table 5 contains the Wilcoxon rank-sum test statistics between each pair of methods. **It tests  
397 the null hypothesis that the two sets of observations are drawn from the same distribution. The  
398 null hypothesis is rejected if the test statistics is greater than 1.96 with a confidence level of 5%  
399 (p-value < 0.05). In this case, it accepts the alternative hypothesis that values in one population are  
400 more likely to be larger than the values from the other.** The two best methods,  $\alpha$ GMK and PMV are not  
401 significantly different. But  $\alpha$ GMK is significantly better than all the other methods, whereas PMV is  
402 not significantly different than the mean map methods (EMK and GMK). The worst method is HDKLD  
403 with a mean F1 of 0.59.



**Figure 9.** Boxplot of F1 score repartitions for the classification of the **old and young** grasslands. The line in the box stands for the median whereas the dot stands for the mean.

**Table 5.** Absolute value of Wilcoxon rank-sum test statistics on F1 score for the **old and young** grasslands classification. \*\* indicates the results are significantly different, *i.e.*,  $p$ -value  $< 0.05$ .

Method	PMV	$\mu$	HDKLD	BD	EMK	GMK	$\alpha$ GMK
PMV	-	3.52**	8.66**	4.83**	1.93	0.98	1.32
$\mu$		-	7.48**	1.76	1.55	2.28**	4.80**
HDKLD			-	5.68**	8.26**	8.65**	9.77**
BD				-	3.23**	3.95**	6.09**
EMK					-	0.94	3.35**
GMK						-	2.42**
$\alpha$ GMK							-

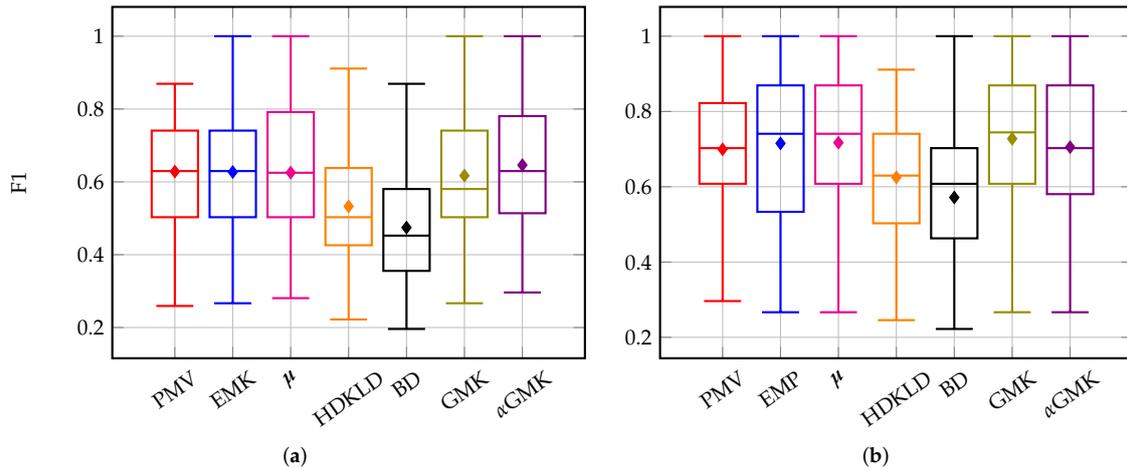
404 In terms of processing load and time, the pixel-based methods are clearly the most demanding.  
 405 Indeed, processing the 160,514 pixels was not possible with SVM, so we had to reduce the number  
 406 of samples. These issues are not faced with object-oriented methods. The fastest methods are  $\mu$  and  
 407 HDKLD, but they did not reach acceptable classification accuracies. The best method in terms of ratio  
 408 accuracy/processing time is  $\alpha$ GMK. It is appropriate for processing a large number of grasslands.

#### 409 4.3.2. Management practices: intra-annual time series

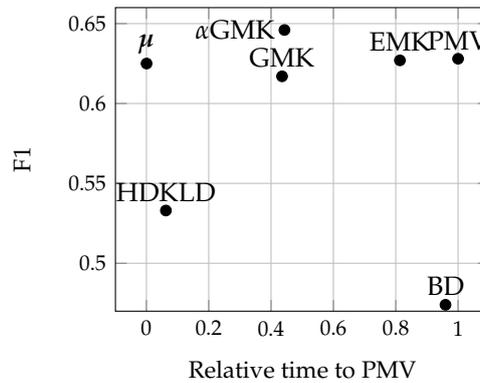
410 The classifications accuracies for management practices are shown in Figure 10 (F1 score) and in  
 411 Figure S3 in the supplementary materials (Kappa coefficient) for year 2013 and for year 2014.

412 In terms of classification accuracy, methods based on divergences (BD and HDKLD) provided  
 413 the worst results. Pixel-based methods, mean modeling method and mean generative kernel methods  
 414 provided similar results in terms of F1 score, except for PMV which was significantly worse than the  
 415 others for the year 2013.  $\alpha$ GMK provided the highest values in 2013 (average F1 of 0.65) but it was not  
 416 significantly better than the others for this dataset. Indeed, due to the very low number of grasslands  
 417 composing this dataset, confusion matrices were quite similar whatever the method. It is therefore  
 418 difficult to compare the methods efficiency in this configuration.

419 Nevertheless, this dataset makes possible the comparison in terms of processing times, because  
 420 the same spectral information was used for all the methods. Figure 11 illustrates the training processing  
 421 time relative to the one of PMV versus the average F1 score for each method. In terms of computational  
 422 time, the pixel-based methods required the largest processing times. BD was also very long, mainly  
 423 because of the shrinkage procedure. Mean modeling was the fastest, followed closely by HDKLD.  
 424  $\alpha$ GMK and GMK were equivalent in terms of computational times. For this configuration with a low  
 425 number of grasslands, the mean modeling was the most efficient in terms of accuracy/processing time  
 426 ratio.



**Figure 10.** Boxplot of F1 score repartitions for classification of management practices using time series of year (a) 2013 and (b) 2014. The line in the box stands for the median whereas the dot stands for the mean.



**Figure 11.** Relative training processing times to PMV and average F1 of each method for intra-annual time series of 2013.

4.27 It is worth noting that the times series of 2014 produced higher classification accuracies (maximum  
4.28 F1 average of 0.73 for GMK) than the time series of 2013 (maximum F1 average of 0.65 for  $\alpha$ GMK).

#### 4.29 4.4. Discussion

4.30 The purpose of this work was to develop a model suitable for the classification of grasslands from  
4.31 dense inter- or intra-annual SITS and robust to the dimension of the data. The proposed method based  
4.32 on a weighted use of the covariance, namely  $\alpha$ GMK, was compared to several competitive methods.

##### 4.33 4.4.1. Methods efficiency

4.34 The methods efficiency are discussed for the old and young grasslands classification, since the  
4.35 results provided with the other dataset are not significantly different, mostly because of the small  
4.36 dataset size.

4.37 The divergence methods (BD and HDKLD) provided the worst results, showing that they are not  
4.38 robust enough to a high dimensional space.

4.39 Although they provided results close to the best results, pixel-based methods (PMV and EMK)  
4.40 are the most demanding in terms of computational time and they do not scale well with the number of  
4.41 pixels. Indeed, they have to process  $N$  pixels instead of  $G$  grasslands with  $G \ll N$ . Therefore, we had

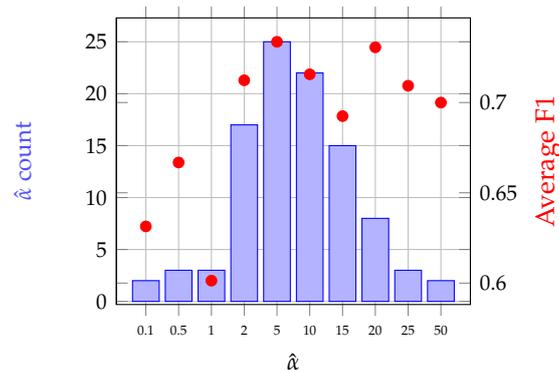
442 to reduce the number of pixels used for the classification. Using them on a large area might be difficult,  
 443 as the **old and young** grasslands dataset showed.

444 Representing grasslands by the estimated distribution of their set of pixels decreases the  
 445 complexity during the SVM process. Therefore, object **level** methods offer a lower computational load  
 446 when compared to empirical mean kernels and pixel-based methods.

447 The mean generative kernel methods performed significantly better than the mean-only method  
 448 ( $\mu$ ). Among them,  $\alpha$ GMK performed better than GMK. It was also one of the most stable methods.

449 In this context, including the covariance information helps to discriminate grasslands. However,  
 450 if the dimensionality is not properly handled, it deteriorates the process (*e.g.*, BD and HDKLD). In this  
 451 case, it is preferable to use the mean values only.  $\alpha$ GMK offers the possibility to weight the influence  
 452 of the covariance information compared to the mean. As a result, it provided better results than the  
 453 mean modeling and than GMK, since it encompasses both.

454 It is furthermore interesting to analyze the optimal values of the weighting parameter  $\alpha$  found  
 455 during the cross-validation and the average of associated F1 scores (Figure 12). The highest F1 scores  
 456 were reached for high values of  $\hat{\alpha}$ . The worst F1 scores were obtained with  $\hat{\alpha} < 2$  and the value  $\hat{\alpha} = 0$   
 457 was never selected. It shows the importance of the covariance information in grasslands modeling: the  
 458 heterogeneity in a grassland must be accounted for and it is not entirely well represented by the mean  
 459 only.



**Figure 12.** Bar plot of  $\hat{\alpha}$  values chosen by cross-validation and the average of associated F1 scores (red dots) for the classification of **old and young** grasslands using  $\alpha$ GMK. NB: The value  $\hat{\alpha} = 0$  was never selected.

#### 4.4.2. Grassland modeling

460 Following on from the methods discussion, the choice of modeling grasslands pixels distribution  
 461 by a Gaussian distribution makes sense in this context. It is particularly well appropriate for  
 462 semi-natural grasslands, which are very heterogeneous, contrary to crops or annual "artificial"  
 463 grasslands which can be assimilated to crops.

464 However, modeling grasslands by the mean only produced equivalent results to the methods  
 465 based on Gaussian modeling for the classification of management practices, **contrary to the old and**  
 466 **young grassland discrimination**. Indeed, management practices are supposed to be uniform at the  
 467 grassland scale. Therefore the mean appears to be sufficient for this application, contrary to the **old and**  
 468 **young** grasslands discrimination, which requires capturing more variations between the grasslands.  
 469 The best modeling might be different depending on the application. Moreover, some grasslands are so  
 470 small that the covariance matrix is too badly estimated.

471 In the proposed kernel, this modeling was made flexible by regularizing the weight given to the  
 472 covariance matrix.  $\alpha$ GMK benefits from its high level of adaptability in front of the object configuration:  
 473 no choice has to be made between a Gaussian or a mean modeling since the method encompasses both.  
 474

475 It also includes several object **level** methods known in the literature. However, this is at the cost of one  
476 more parameter to tune. Therefore, the classification process takes more time than GMK for instance.

477 Above all, although it is the first application of generative mean kernels in remote sensing  
478 classification, the  $\alpha$ -Gaussian mean kernel proved its efficiency and stability in these experiments. **The**  
479 **results suggest** it is appropriate for grasslands classification.

#### 480 4.4.3. Acquisition dates

481 For the management practice classification, using time series of 2014 produced better results  
482 than using 2013. It might be explained by the acquisition dates in the time series. Although 2014  
483 has less images, more clear images were acquired during Spring compared to 2013 which has a  
484 lack of acquisitions in April and May (Figure 3). Indeed, many studies showed that the best season  
485 to discriminate grasslands is during the growing season [36,49,53,54]. Spring is the period of the  
486 vegetation cycle where the management practices begin. Therefore, it is easier to differentiate the  
487 practices during this period. It might thus affect the accuracy of the classification of year 2013.

488 It is not shown in this experiment, but using only one or two years of acquisitions to discriminate  
489 **old from young** grasslands did not produce sufficient classification accuracies. This is the reason  
490 why three years of data were used. Old "permanent" grasslands are supposed to have a more stable  
491 phenology over the years than the young "temporary" grasslands which have been recently sown  
492 (less than five years) [6]. The **young** grasslands phenology is closer to crops in their very first years.  
493 We suppose this makes possible their discrimination with inter-annual SITS. However, the optimal  
494 number of years needed to discriminate these types of grasslands could constitute a research topic.

495 In general, the results could also be enhanced by removing some winter images which can have  
496 a negative influence on the entire annual time series [40]. However, the scope of this study was to  
497 develop a method which is able to use a given time series, without having to process a date selection.

#### 498 4.4.4. Grassland typology

499 On the whole, the classification did not reach high accuracies (F1 maximum average of 0.73 for  
500 management practices and of 0.71 for **old and young** grasslands classification). This can be explained  
501 by the unbalanced dataset with under-representation of grazing and mixed grasslands in the first  
502 application and under-representation of **old** grasslands in the second one. **These classes obtained the**  
503 **lowest producer and user accuracies (cf. Tables S1 and S2 in the supplementary materials) because**  
504 **of their limited number of samples for training the models.** The methods should be tested on a more  
505 balanced dataset of grasslands classes.

506 Moreover, as many times emphasized, semi-natural grasslands (which are present in these  
507 datasets) are characterized by their high level of heterogeneity. Therefore, there might be a large amount  
508 of intra-class variability because of grasslands diversity. The discrimination might be improved by  
509 using more distinct classes: intensively used grasslands against extensively used grasslands, artificial  
510 (monospecific) grasslands against semi-natural grasslands for instance.

#### 511 4.4.5. Comparison with existing works

512 In our knowledge, only the work of Möckel *et al.* [86] relates to the classification of grasslands age  
513 using remote sensing data. They reached a Kappa value of 0.77 in classifying three different grassland  
514 age-classes. However, they used airborne hyperspectral data from a single date. Their recommendation  
515 was to use multitemporal data to improve the classification or to use satellite hyperspectral data to  
516 monitor grasslands over wider areas. Our study was based on using multispectro-temporal satellite  
517 data, but our proposed method would also work with hyperspectral data.

518 As described in the introduction, few studies have been carried out on the analysis of semi-natural  
519 grasslands using high spatio-spectro-temporal resolutions SITS. Usually, methods were pixel-based  
520 and they were applied on a few images or on a precise date selection to avoid dealing with the high

521 dimension of data [42,44,49]. Schuster *et al.* [52] successfully classified grassland habitat using 21  
522 RapidEye images on a pixel basis, but there was no mention of the processing times.

523 At the object level using a time series, grasslands were often represented by their mean NDVI,  
524 such as in [60], who noticed the difficulty to discriminate grasslands from crops because of mean  
525 seasonal NDVI similarities. The closest configuration might be the work of Zillman *et al.* [35], who  
526 used an object-based analysis and spectral reflectances combined with seasonal statistics of vegetation  
527 indices for mapping grasslands across Europe. The seasonal statistics were particularly relevant in the  
528 classification, because they captured well the spectral diversity of the grassland phenology. The use of  
529 these metrics could be considered for discriminating grassland management practices which impact on  
530 the phenology. The authors also concluded that the object-based analysis improves the classification  
531 compared to a pixel-based classification. However, the objects were determined by segmentation.

#### 532 4.5. Prediction of management practices on the land use database grasslands

533 To show the efficiency of  $\alpha$ GMK, we classified all the grasslands from the French agricultural land  
534 use database (RPG) covered by the Formosat-2 time series to predict their management practice in 2014.  
535 All the plots declared as grasslands in 2014, *i.e.*, "permanent grassland" and "temporary grassland"  
536 regardless of their age, were selected. After applying a negative buffer of 8m and rasterizing the  
537 polygons, we removed the plots representing less than 10 Formosat-2 pixels. In the end, there were 797  
538 grassland plots covered by the extent of Formosat-2 for a total of 252,472 pixels.

539 The multispectral SITS of 2014 was used. The SVM was trained on the whole field data  
540 (section 2.3.2) using the same grid search as in the experiments. The parameters chosen after  
541 cross-validation based on F1 score were  $\hat{\alpha} = 5$  and  $\hat{\gamma} = 2^{-15}$ . Then, the model was used to predict the  
542 management practices of the 797 grasslands of the land use database.

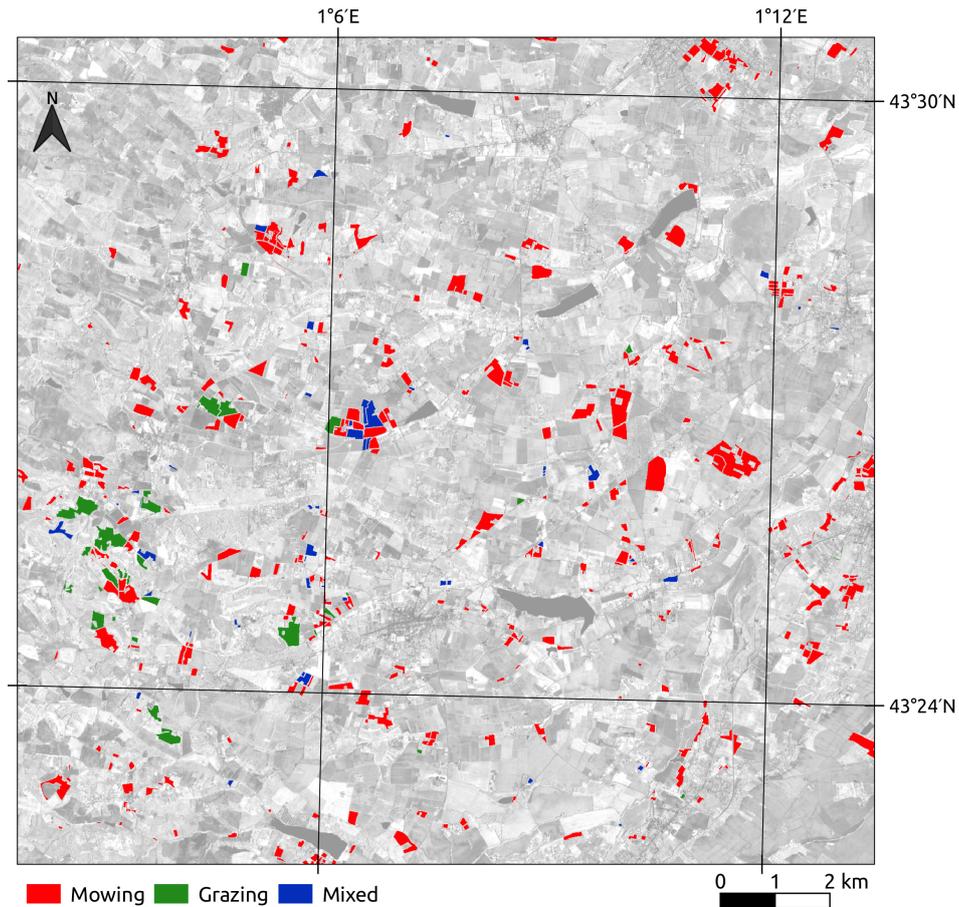
543 The classification accuracy could not be assessed since the true labels of the grasslands are not  
544 known. However, as described in the study site, a spatial distribution of the classes could be expected.  
545 Indeed, grazed and mixed grasslands should be found in the south-west of the site,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

547 An extract of the classification result is shown in Figure 13. It represents the classified grasslands in  
548 their raster format. As expected, most of the grazed and mixed grasslands are located in the south-west  
549 of the image, whereas the north of the image is mostly composed of mown grasslands. Therefore,  
550  $\alpha$ GMK was very likely able to classify with an acceptable accuracy the grasslands management  
551 practices without any *a priori* geographic information. However, specific care should be considered, as  
552 not all the possible management practices were predicted. For instance, grasslands mown twice or  
553 unused grasslands were not in the training dataset, but it does not mean these managements do not  
554 exist in the rest of the data. The method deserves to be tested with an exhaustive grassland typology  
555 to produce more detailed grasslands maps.

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

## 559 5. Conclusion

560 This study aimed at developing a model for the classification of grasslands using satellite image  
561 time series with a high number of spectro-temporal variables. A grassland modeling at the object  
562 level was proposed. To deal with grasslands heterogeneity, their pixels distribution was modeled by  
563 a Gaussian distribution. Then, to measure the similarity between two grasslands, *i.e.*, two Gaussian  
564 distributions, a kernel function based on mean maps was introduced, namely the  $\alpha$ -Gaussian Mean  
565 Kernel. The proposed method was compared to existing pixel-based and object-based classification  
566 methods for the supervised classification of grassland using inter- and intra-annual SITS. The Gaussian  
567 mean kernels provided the highest classification accuracies, showing that the covariance information



**Figure 13.** Extract of the management practices 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.

568 must be accounted for. In terms of processing times, the object-based methods were much faster than  
 569 pixel-based methods.

570 Several contributions have been made in this work. The first lies in the grasslands pixels  
 571 distribution modeling at the object level. A flexible kernel was proposed to encompass both Gaussian  
 572 and mean modeling of grasslands, so no choice has to be made between these two modelings. It can  
 573 therefore be used on homogeneous objects such as artificial grasslands, or on very small objects, as well  
 574 as on heterogeneous semi-natural grasslands. The second contribution is that this kernel is suitable for  
 575 high dimensional data in a small ground sample size context. It enables the use of all the multispectral  
 576 data instead of a single vegetation index or the use of a long time series. Also, it can be used on a  
 577 whole time series without dates selection. Indeed, this new kernel offers very low computational load.  
 578 It can therefore be applied on a large dataset. With this kernel, we were able to process and to classify  
 579 more than 250,000 pixels on a conventional personal computer within a few seconds. Even if it is the  
 580 first application of generative mean kernels in remote sensing classification, the  $\alpha$ GMK proved its  
 581 efficiency and stability in these experiments. It is a good compromise between processing speed and  
 582 accuracy for the classification of grasslands.

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

588 **Supplementary Materials:** The following are available online at [www.mdpi.com/link](http://www.mdpi.com/link), Figure S1: True color  
 589 composite images of the Formosat-2 time series of 2014; Figure S2: Boxplot of Kappa coefficient repartitions for the  
 590 classification of the old and young grasslands; Table S1: Average user accuracy (UA) and producer accuracy (PA)  
 591 (%) over the 100 repetitions for each class, 1: Old, 2: Young; Figure S3: Boxplot of Kappa coefficient repartitions  
 592 for classification of management practices using time series of year (a) 2013 and (b) 2014; Table S2: Average user  
 593 accuracy (UA) and producer accuracy (PA) (%) over the 100 repetitions for each class, 1: Mowing, 2: Mixed, 3:  
 594 Grazing.

595 **Acknowledgments:** This work was partially supported by a CJS INRA-INRIA contract and by the grant Défi  
 596 Mastodons-CNRS. The authors would like to thank CNES and CESBIO for providing the pre-processed Formosat-2  
 597 data. Special thanks to M. Lang for playing a major role in the field work, to D. Dallery for designing the processing  
 598 chain to compute the grasslands age from the RPG, and to R. Carrié for his careful reviewing of the introduction.  
 599 **We would like to thank the anonymous reviewers for their constructive comments.**

600 **Author Contributions:** M.L., M.F. and S.G. conceived the model, M.L., M.F. and S.G. conceived and designed the  
 601 experiments; M.L. performed the experiments; M.L. and M.F. analyzed the data; M.L. and M.F. wrote the paper  
 602 with feedbacks from S.G. and D.S.

603 **Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the design  
 604 of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the  
 605 decision to publish the results.

## 606 Abbreviations

607 The following abbreviations are used in this manuscript:

608	BD	Bhattacharyya Distance
	EMK	Empirical Mean Kernel
	GIS	Geographic Information System
	GMK	Gaussian Mean Kernel
	HDKLD	High Dimensional Kullback-Leibler Divergence
	JMD	Jeffries-Matusita Distance
	KLD	Kullback-Leibler Divergence
609	LAI	Leaf Area Index
	NDVI	Normalized Difference Vegetation Index
	NIR	Near Infrared
	PMV	Pixel Majority Vote
	RBF	Radial Basis Function
	SITS	Satellite Image Time Series
	SVM	Support Vector Machine
	$\alpha$ GMK	$\alpha$ -Gaussian Mean Kernel

## 610 Appendix

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

$$\begin{aligned}
 p_i(\mathbf{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)^{\alpha^{-1}} &= \frac{1}{(2\pi)^{d/2\alpha}} \times \frac{1}{|\boldsymbol{\Sigma}_i|^{1/2\alpha}} \times \exp \left\{ -0.5(\mathbf{x} - \boldsymbol{\mu}_i)^\top (\alpha\boldsymbol{\Sigma}_i)^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) \right\} \\
 &= \frac{(2\pi)^{\frac{d}{2}(1-\frac{1}{\alpha})}}{(2\pi)^{d/2}} \times \alpha^{1/2} \times \frac{|\boldsymbol{\Sigma}_i|^{\frac{1}{2}(1-\frac{1}{\alpha})}}{|\alpha\boldsymbol{\Sigma}_i|^{1/2}} \times \exp \left\{ -0.5(\mathbf{x} - \boldsymbol{\mu}_i)^\top (\alpha\boldsymbol{\Sigma}_i)^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) \right\} \\
 &= \alpha^{1/2} (2\pi)^{\frac{d}{2}(1-\frac{1}{\alpha})} |\boldsymbol{\Sigma}_i|^{\frac{1}{2}(1-\frac{1}{\alpha})} \times p(\mathbf{x}|\boldsymbol{\mu}_i, \alpha\boldsymbol{\Sigma}_i) \\
 &= C(\boldsymbol{\Sigma}_i, \alpha) p(\mathbf{x}|\boldsymbol{\mu}_i, \alpha\boldsymbol{\Sigma}_i).
 \end{aligned} \tag{A1}$$

612 Then, plugging eq. (A1) in eq. (7), we get:

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

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

615 **References**

- 616 1. Eriksson, A.; Eriksson, O.; Berglund, H. Species Abundance Patterns of Plants in Swedish Semi-Natural  
617 Pastures. *Ecography* **1995**, *18*, 310–317.
- 618 2. Cousins, S.A.; Eriksson, O. The influence of management history and habitat on plant species richness in a  
619 rural hemiboreal landscape, Sweden. *Landscape Ecology* **2002**, *17*, 517–529.
- 620 3. Gardi, C.; Tomaselli, M.; Parisi, V.; Petraglia, A.; Santini, C. Soil quality indicators and biodiversity in  
621 northern Italian permanent grasslands. *European Journal of Soil Biology* **2002**, *38*, 103 – 110.
- 622 4. Critchley, C.; Burke, M.; Stevens, D. Conservation of lowland semi-natural grasslands in the UK: a review  
623 of botanical monitoring results from agri-environment schemes. *Biological Conservation* **2004**, *115*, 263 – 278.
- 624 5. Werling, B.P.; Dickson, T.L.; Isaacs, R.; Gaines, H.; Gratton, C.; Gross, K.L.; Liere, H.; Malmstrom, C.M.;  
625 Meehan, T.D.; Ruan, L.; Robertson, B.A.; Robertson, G.P.; Schmidt, T.M.; Schrottenboer, A.C.; Teal, T.K.;  
626 Wilson, J.K.; Landis, D.A. Perennial grasslands enhance biodiversity and multiple ecosystem services in  
627 bioenergy landscapes. *Proceedings of the National Academy of Sciences* **2014**, *111*, 1652–1657.
- 628 6. Austrheim, G.; Olsson, E.G.A. How does continuity in grassland management after ploughing affect plant  
629 community patterns? *Plant Ecology* **1999**, *145*, 59–74.
- 630 7. Norderhaug, A.; Ihse, M.; Pedersen, O. Biotope patterns and abundance of meadow plant species in a  
631 Norwegian rural landscape. *Landscape Ecology* **2000**, *15*, 201–218.
- 632 8. Waldhardt, R.; Otte, A. Indicators of plant species and community diversity in grasslands. *Agriculture,*  
633 *Ecosystems & Environment* **2003**, *98*, 339 – 351. Biotic Indicators for Biodiversity and Sustainable Agriculture.
- 634 9. Hansson, M.; Fogelfors, H. Management of a semi-natural grassland; results from a 15-year-old experiment  
635 in southern Sweden. *Journal of Vegetation Science* **2000**, *11*, 31–38.
- 636 10. Moog, D.; Poschlod, P.; Kahmen, S.; Schreiber, K.F. Comparison of species composition between different  
637 grassland management treatments after 25 years. *Applied Vegetation Science* **2002**, *5*, 99–106.
- 638 11. Zechmeister, H.; Schmitzberger, I.; Steurer, B.; Peterseil, J.; Wrška, T. The influence of land-use practices  
639 and economics on plant species richness in meadows. *Biological Conservation* **2003**, *114*, 165 – 177.
- 640 12. Plantureux, S.; Peeters, A.; McCracken, D. Biodiversity in intensive grasslands: Effect of management,  
641 improvement and challenges. *Agronomy Research* **2005**, *3*, 153–164.
- 642 13. Muller, S. Appropriate agricultural management practices required to ensure conservation and biodiversity  
643 of environmentally sensitive grassland sites designated under Natura 2000. *Agriculture, Ecosystems &*  
644 *Environment* **2002**, *89*, 261 – 266.
- 645 14. Rocchini, D.; Boyd, D.S.; Féret, J.B.; Foody, G.M.; He, K.S.; Lausch, A.; Nagendra, H.; Wegmann, M.;  
646 Pettorelli, N. Satellite remote sensing to monitor species diversity: potential and pitfalls. *Remote Sensing in*  
647 *Ecology and Conservation* **2016**, *2*, 25–36.
- 648 15. Pettorelli, N.; Laurance, W.F.; O'Brien, T.G.; Wegmann, M.; Nagendra, H.; Turner, W. Satellite remote  
649 sensing for applied ecologists: opportunities and challenges. *Journal of Applied Ecology* **2014**, *51*, 839–848.
- 650 16. Newton, A.C.; Hill, R.A.; Echeverría, C.; Golicher, D.; Rey Benayas, J.M.; Cayuela, L.; Hinsley, S.A. Remote  
651 sensing and the future of landscape ecology. *Progress in Physical Geography* **2009**, *33*, 528–546.
- 652 17. Gu, Y.; Wylie, B.K.; Bliss, N.B. Mapping grassland productivity with 250-m eMODIS NDVI and SSURGO  
653 database over the Greater Platte River Basin, USA. *Ecological Indicators* **2013**, *24*, 31 – 36.
- 654 18. Li, Z.; Huffman, T.; McConkey, B.; Townley-Smith, L. Monitoring and modeling spatial and temporal  
655 patterns of grassland dynamics using time-series MODIS NDVI with climate and stocking data. *Remote*  
656 *Sensing of Environment* **2013**, *138*, 232 – 244.
- 657 19. Gu, Y.; Wylie, B.K. Developing a 30-m grassland productivity estimation map for central Nebraska using  
658 250-m MODIS and 30-m Landsat-8 observations. *Remote Sensing of Environment* **2015**, *171*, 291 – 298.
- 659 20. Friedl, M.A.; Michaelsen, J.; Davis, F.W.; Walker, H.; Schimel, D.S. Estimating grassland biomass and Leaf  
660 Area Index using ground and satellite data. *International Journal of Remote Sensing* **1994**, *15*, 1401–1420.
- 661 21. Wylie, B.; Meyer, D.; Tieszen, L.; Mannel, S. Satellite mapping of surface biophysical parameters at the  
662 biome scale over the North American grasslands: A case study. *Remote Sensing of Environment* **2002**, *79*, 266  
663 – 278. Recent Advances in Remote Sensing of Biophysical Variables.
- 664 22. Darvishzadeh, R.; Skidmore, A.; Schlerf, M.; Atzberger, C.; Corsi, F.; Cho, M. LAI and chlorophyll  
665 estimation for a heterogeneous grassland using hyperspectral measurements. *ISPRS Journal of*  
666 *Photogrammetry and Remote Sensing* **2008**, *63*, 409 – 426.

- 667 23. He, Y.; Guo, X.; Wilmshurst, J.F. Reflectance measures of grassland biophysical structure. *International*  
668 *Journal of Remote Sensing* **2009**, *30*, 2509–2521.
- 669 24. Asam, S.; Fabritius, H.; Klein, D.; Conrad, C.; Dech, S. Derivation of leaf area index for grassland  
670 within alpine upland using multi-temporal RapidEye data. *International Journal of Remote Sensing* **2013**,  
671 *34*, 8628–8652.
- 672 25. Schmidlein, S.; Sassini, J. Mapping of continuous floristic gradients in grasslands using hyperspectral  
673 imagery. *Remote Sensing of Environment* **2004**, *92*, 126 – 138.
- 674 26. Ishii, J.; Lu, S.; Funakoshi, S.; Shimizu, Y.; Omasa, K.; Washitani, I. Mapping potential habitats of threatened  
675 plant species in a moist tall grassland using hyperspectral imagery. *Biodiversity and Conservation* **2009**,  
676 *18*, 2521–2535.
- 677 27. Fava, F.; Parolo, G.; Colombo, R.; Gusmeroli, F.; Marianna, G.D.; Monteiro, A.; Bocchi, S. Fine-scale  
678 assessment of hay meadow productivity and plant diversity in the European Alps using field spectrometric  
679 data. *Agriculture, Ecosystems & Environment* **2010**, *137*, 151 – 157. Special section Harvested perennial  
680 grasslands: Ecological models for farming's perennial future.
- 681 28. Oldeland, J.; Wesuls, D.; Rocchini, D.; Schmidt, M.; Jürgens, N. Does using species abundance data improve  
682 estimates of species diversity from remotely sensed spectral heterogeneity? *Ecological Indicators* **2010**,  
683 *10*, 390 – 396.
- 684 29. Feilhauer, H.; Faude, U.; Schmidlein, S. Combining Isomap ordination and imaging spectroscopy to map  
685 continuous floristic gradients in a heterogeneous landscape. *Remote Sensing of Environment* **2011**, *115*, 2513 –  
686 2524.
- 687 30. Duniway, M.C.; Karl, J.W.; Schrader, S.; Baquera, N.; Herrick, J.E. Rangeland and pasture monitoring: an  
688 approach to interpretation of high-resolution imagery focused on observer calibration for repeatability.  
689 *Environmental Monitoring and Assessment* **2012**, *184*, 3789–3804.
- 690 31. Punalekar, S.; Verhoef, A.; Tatarenko, I.V.; van der Tol, C.; Macdonald, D.M.J.; Marchant, B.; Gerard, F.;  
691 White, K.; Gowing, D. Characterization of a Highly Biodiverse Floodplain Meadow Using Hyperspectral  
692 Remote Sensing within a Plant Functional Trait Framework. *Remote Sensing* **2016**, *8*, 112.
- 693 32. Hilker, T.; Natsagdorj, E.; Waring, R.H.; Lyapustin, A.; Wang, Y. Satellite observed widespread decline in  
694 Mongolian grasslands largely due to overgrazing. *Global Change Biology* **2014**, *20*, 418–428.
- 695 33. Cao, R.; Chen, J.; Shen, M.; Tang, Y. An improved logistic method for detecting spring vegetation phenology  
696 in grasslands from MODIS EVI time-series data. *Agricultural and Forest Meteorology* **2015**, *200*, 9 – 20.
- 697 34. Eriksson, O.; Cousins, S.A.; Bruun, H.H. Land-use history and fragmentation of traditionally managed  
698 grasslands in Scandinavia. *Journal of Vegetation Science* **2002**, *13*, 743–748.
- 699 35. Zillmann, E.; Gonzalez, A.; Herrero, E.J.M.; van Wolvelaer, J.; Esch, T.; Keil, M.; Weichelt, H.; Garzón, A.M.  
700 Pan-European Grassland Mapping Using Seasonal Statistics From Multisensor Image Time Series. *IEEE*  
701 *Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **2014**, *7*, 3461–3472.
- 702 36. Ali, I.; Cawkwell, F.; Dwyer, E.; Barrett, B.; Green, S. Satellite remote sensing of grasslands: from observation  
703 to management. *Journal of Plant Ecology* **2016**, *9*, 649–671.
- 704 37. Nagendra, H. Using remote sensing to assess biodiversity. *International Journal of Remote Sensing* **2001**,  
705 *22*, 2377–2400.
- 706 38. Blaschke, T.; Hay, G.J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Feitosa, R.Q.; van der Meer, F.; van der  
707 Werff, H.; van Coillie, F.; Tiede, D. Geographic Object-Based Image Analysis – Towards a new paradigm.  
708 *ISPRS Journal of Photogrammetry and Remote Sensing* **2014**, *87*, 180 – 191.
- 709 39. Poças, I.; Cunha, M.; Pereira, L.S. Dynamics of mountain semi-natural grassland meadows inferred  
710 from SPOT-VEGETATION and field spectroradiometer data. *International Journal of Remote Sensing* **2012**,  
711 *33*, 4334–4355.
- 712 40. Halabuk, A.; Moyses, M.; Halabuk, M.; David, S. Towards Detection of Cutting in Hay Meadows by Using  
713 of NDVI and EVI Time Series. *Remote Sensing* **2015**, *7*, 6107.
- 714 41. Lucas, R.; Rowlands, A.; Brown, A.; Keyworth, S.; Bunting, P. Rule-based classification of multi-temporal  
715 satellite imagery for habitat and agricultural land cover mapping. *ISPRS Journal of Photogrammetry and*  
716 *Remote Sensing* **2007**, *62*, 165 – 185.
- 717 42. Toivonen, T.; Luoto, M. Landsat TM images in mapping of semi-natural grasslands and analysing of  
718 habitat pattern in an agricultural landscape in south-west Finland. *Fennia - International Journal of Geography*  
719 **2003**, *181*, 49 – 67.

- 720 43. Nagendra, H.; Lucas, R.; Honrado, J.P.; Jongman, R.H.; Tarantino, C.; Adamo, M.; Mairota, P. Remote  
721 sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species  
722 diversity, and threats. *Ecological Indicators* **2013**, *33*, 45 – 59. Biodiversity Monitoring.
- 723 44. Price, K.P.; Guo, X.; Stiles, J.M. Optimal Landsat TM band combinations and vegetation indices for  
724 discrimination of six grassland types in eastern Kansas. *International Journal of Remote Sensing* **2002**,  
725 *23*, 5031–5042.
- 726 45. Gamon, J.A.; Field, C.B.; Roberts, D.A.; Ustin, S.L.; Valentini, R. Airbone Imaging Spectrometry Functional  
727 patterns in an annual grassland during an AVIRIS overflight. *Remote Sensing of Environment* **1993**, *44*, 239 –  
728 253.
- 729 46. Corbane, C.; Lang, S.; Pipkins, K.; Alleaume, S.; Deshayes, M.; Millán, V.E.G.; Strasser, T.; Borre, J.V.;  
730 Toon, S.; Michael, F. Remote sensing for mapping natural habitats and their conservation status – New  
731 opportunities and challenges. *International Journal of Applied Earth Observation and Geoinformation* **2015**,  
732 *37*, 7 – 16. Special Issue on Earth observation for habitat mapping and biodiversity monitoring.
- 733 47. Wulder, M.A.; Hall, R.J.; Coops, N.C.; Franklin, S.E. High Spatial Resolution Remotely Sensed Data for  
734 Ecosystem Characterization. *BioScience* **2004**, *54*, 511–521.
- 735 48. Buck, O.; Millán, V.E.G.; Klink, A.; Pakzad, K. Using information layers for mapping grassland habitat  
736 distribution at local to regional scales. *International Journal of Applied Earth Observation and Geoinformation*  
737 **2015**, *37*, 83 – 89. Special Issue on Earth observation for habitat mapping and biodiversity monitoring.
- 738 49. Franke, J.; Keuck, V.; Siegert, F. Assessment of grassland use intensity by remote sensing to support  
739 conservation schemes. *Journal for Nature Conservation* **2012**, *20*, 125 – 134.
- 740 50. Schmidt, T.; Schuster, C.; Kleinschmit, B.; Forster, M. Evaluating an Intra-Annual Time Series for Grassland  
741 Classification - How Many Acquisitions and What Seasonal Origin Are Optimal? *Selected Topics in Applied*  
742 *Earth Observations and Remote Sensing, IEEE Journal of* **2014**, *7*, 3428–3439.
- 743 51. Dusseux, P.; Vertès, F.; Corpetti, T.; Corgne, S.; Hubert-Moy, L. Agricultural practices in grasslands detected  
744 by spatial remote sensing. *Environmental Monitoring and Assessment* **2014**, *186*, 8249–8265.
- 745 52. Schuster, C.; Schmidt, T.; Conrad, C.; Kleinschmit, B.; Förster, M. Grassland habitat mapping by  
746 intra-annual time series analysis – Comparison of RapidEye and TerraSAR-X satellite data. *International*  
747 *Journal of Applied Earth Observation and Geoinformation* **2015**, *34*, 25 – 34.
- 748 53. Psomas, A.; Kneubuhler, M.; Huber, S.; Itten, K.; Zimmermann, N.E. Hyperspectral remote sensing  
749 for estimating aboveground biomass and for exploring species richness patterns of grassland habitats.  
750 *International Journal of Remote Sensing* **2011**, *32*, 9007–9031.
- 751 54. Hill, M.J. Vegetation index suites as indicators of vegetation state in grassland and savanna: An analysis  
752 with simulated SENTINEL-2 data for a North American transect. *Remote Sensing of Environment* **2013**,  
753 *137*, 94 – 111.
- 754 55. Drusch, M.; Bello, U.D.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti,  
755 P.; Martimort, P.; Meygret, A.; Spoto, F.; Sy, O.; Marchese, F.; Bargellini, P. Sentinel-2: ESA's Optical  
756 High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment* **2012**, *120*, 25 – 36.  
757 The Sentinel Missions - New Opportunities for Science.
- 758 56. Laliberte, A.S.; Fredrickson, E.L.; Rango, A. Combining decision trees with hierarchical object-oriented  
759 image analysis for mapping arid rangelands. *Photogrammetric engineering & Remote sensing* **2007**,  
760 *73*, 197–207.
- 761 57. Brenner, J.C.; Christman, Z.; Rogan, J. Segmentation of Landsat Thematic Mapper imagery improves  
762 buffelgrass (*Pennisetum ciliare*) pasture mapping in the Sonoran Desert of Mexico. *Applied Geography* **2012**,  
763 *34*, 569 – 575.
- 764 58. Stenzel, S.; Fassnacht, F.E.; Mack, B.; Schmidlein, S. Identification of high nature value grassland with  
765 remote sensing and minimal field data. *Ecological Indicators* **2017**, *74*, 28 – 38.
- 766 59. Evans, J.; Geerken, R. Classifying rangeland vegetation type and coverage using a Fourier component  
767 based similarity measure. *Remote Sensing of Environment* **2006**, *105*, 1 – 8.
- 768 60. Esch, T.; Metz, A.; Marconcini, M.; Keil, M. Combined use of multi-seasonal high and medium resolution  
769 satellite imagery for parcel-related mapping of cropland and grassland. *International Journal of Applied*  
770 *Earth Observation and Geoinformation* **2014**, *28*, 230 – 237.

- 771 61. Duro, D.C.; Franklin, S.E.; Dubé, M.G. A comparison of pixel-based and object-based image analysis with  
772 selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG  
773 imagery. *Remote Sensing of Environment* **2012**, *118*, 259 – 272.
- 774 62. Sheeren, D.; Fauvel, M.; Josipović, V.; Lopes, M.; Planque, C.; Willm, J.; Dejoux, J.F. Tree Species  
775 Classification in Temperate Forests Using Formosat-2 Satellite Image Time Series. *Remote Sensing* **2016**,  
776 *8*, 734.
- 777 63. Ding, Y.; Zhao, K.; Zheng, X.; Jiang, T. Temporal dynamics of spatial heterogeneity over cropland quantified  
778 by time-series NDVI, near infrared and red reflectance of Landsat 8 OLI imagery. *International Journal of*  
779 *Applied Earth Observation and Geoinformation* **2014**, *30*, 139 – 145.
- 780 64. Pan, Z.; Huang, J.; Zhou, Q.; Wang, L.; Cheng, Y.; Zhang, H.; Blackburn, G.A.; Yan, J.; Liu, J. Mapping  
781 crop phenology using NDVI time-series derived from HJ-1 A/B data. *International Journal of Applied Earth*  
782 *Observation and Geoinformation* **2015**, *34*, 188 – 197.
- 783 65. Cingolani, A.M.; Renison, D.; Zak, M.R.; Cabido, M.R. Mapping vegetation in a heterogeneous mountain  
784 rangeland using Landsat data: an alternative method to define and classify land-cover units. *Remote*  
785 *Sensing of Environment* **2004**, *92*, 84 – 97.
- 786 66. Müller, H.; Rufin, P.; Griffiths, P.; Siqueira, A.J.B.; Hostert, P. Mining dense Landsat time series for  
787 separating cropland and pasture in a heterogeneous Brazilian savanna landscape. *Remote Sensing of*  
788 *Environment* **2015**, *156*, 490 – 499.
- 789 67. Mitchell, T.M. *Machine Learning*; McGraw-Hill, Inc.: New York, NY, USA, 1997.
- 790 68. Donoho, D.L. High-dimensional data analysis: The curses and blessings of dimensionality. AMS  
791 Conference on math challenges of the 21st century, 2000.
- 792 69. Fauvel, M.; Tarabalka, Y.; Benediktsson, J.A.; Chanussot, J.; Tilton, J.C. Advances in spectral-spatial  
793 classification of hyperspectral images. *Proceedings of the IEEE* **2013**, *101*, 652–675.
- 794 70. Hagolle, O.; Huc, M.; Villa Pascual, D.; Dedieu, G. A multi-temporal method for cloud detection, applied  
795 to FORMOSAT-2, VENUS, LANDSAT and SENTINEL-2 images. *Remote Sensing of Environment* **2010**,  
796 *114*, 1747–1755.
- 797 71. Eilers, P.H.C. A Perfect Smoother. *Analytical Chemistry* **2003**, *75*, 3631–3636. PMID: 14570219.
- 798 72. Atzberger, C.; Eilers, P.H. A time series for monitoring vegetation activity and phenology at 10-daily time  
799 steps covering large parts of South America. *International Journal of Digital Earth* **2011**, *4*, 365–386.
- 800 73. Atzberger, C.; Eilers, P.H.C. Evaluating the effectiveness of smoothing algorithms in the absence of ground  
801 reference measurements. *International Journal of Remote Sensing* **2011**, *32*, 3689–3709.
- 802 74. Nitze, I.; Barrett, B.; Cawkwell, F. Temporal optimisation of image acquisition for land cover classification  
803 with Random Forest and MODIS time-series. *International Journal of Applied Earth Observation and*  
804 *Geoinformation* **2015**, *34*, 136 – 146.
- 805 75. Shao, Y.; Lunetta, R.S.; Wheeler, B.; Iiames, J.S.; Campbell, J.B. An evaluation of time-series smoothing  
806 algorithms for land-cover classifications using MODIS-NDVI multi-temporal data. *Remote Sensing of*  
807 *Environment* **2016**, *174*, 258 – 265.
- 808 76. Kullback, S. Letter to the Editor: The Kullback-Leibler distance. *The American Statistician* **1987**, *41*, 340–341.
- 809 77. Richards, J.A.; Jia, X. *Remote Sensing Digital Image Analysis: An Introduction*, 3rd ed.; Springer-Verlag New  
810 York, Inc.: Secaucus, NJ, USA, 1999.
- 811 78. Mehta, N.A.; Gray, A.G. Generative and Latent Mean Map Kernels. *CoRR* **2010**, *abs/1005.0188*.
- 812 79. Gomez-Chova, L.; Camps-Valls, G.; Bruzzone, L.; Calpe-Maravilla, J. Mean Map Kernel Methods for  
813 Semisupervised Cloud Classification. *IEEE Transactions on Geoscience and Remote Sensing* **2010**, *48*, 207–220.
- 814 80. Muandet, K.; Fukumizu, K.; Dinuzzo, F.; Schölkopf, B. Learning from distributions via support measure  
815 machines. *Advances in Neural Information Processing Systems* 25. Curran Associates Inc., 2012, pp. 10–18.
- 816 81. Tarantola, A. *Inverse Problem Theory and Methods for Model Parameter Estimation*; Society for Industrial and  
817 Applied Mathematics: Philadelphia, PA, USA, 2005.
- 818 82. Lopes, M.; Fauvel, M.; Girard, S.; Sheeren, D. High dimensional Kullback-Leibler divergence for  
819 grassland management practices classification from high resolution satellite image time series. 2016  
820 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2016, pp. 3342–3345.
- 821 83. Ledoit, O.; Wolf, M. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of*  
822 *Multivariate Analysis* **2004**, *88*, 365 – 411.
- 823 84. Wilcoxon, F. Individual Comparisons by Ranking Methods. *Biometrics Bulletin* **1945**, *1*, 80–83.

- 824 85. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.;  
825 Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E.  
826 Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.
- 827 86. Möckel, T.; Dalmayne, J.; Prentice, H.C.; Eklundh, L.; Purschke, O.; Schmidlein, S.; Hall, K. Classification  
828 of Grassland Successional Stages Using Airborne Hyperspectral Imagery. *Remote Sensing* **2014**, *6*, 7732.

829 © 2017 by the authors. Submitted to *Remote Sens.* for possible open access publication  
830 under the terms and conditions of the Creative Commons Attribution (CC BY) license  
831 (<http://creativecommons.org/licenses/by/4.0/>).