

Article

## A Spectral Unmixing Model for the Integration of Multi-Sensor Imagery: A Tool to Generate Consistent Time Series Data

Georgia Doxani <sup>1,\*</sup>, Zina Mitraka <sup>2</sup>, Ferran Gascon <sup>1</sup>, Philippe Goryl <sup>1</sup> and Bojan R. Bojkov <sup>1</sup>

<sup>1</sup> European Space Agency, Via Galileo Galilei, Frascati 00044, Italy;

E-Mails: Ferran.Gascon@esa.int (F.G.); Philippe.Goryl@esa.int (P.G.);

Bojan.Bojkov@esa.int (B.R.B.)

<sup>2</sup> Foundation for Research and Technology Hellas, N.Plastira 100, Vassilika Vouton,

Heraklion 70013, Greece; E-Mail: mitraka@iacm.forth.gr

\* Author to whom correspondence should be addressed; E-Mail: Georgia.Doxani@esa.int;

Tel.: +39-06-941-80458; Fax: +39-06-941-80302.

Academic Editors: Olivier Hagolle, Benjamin Koetz, Olivier Arino, Sylvia Sylvander, Anton Vrieling and Prasad S. Thenkabail

Received: 29 May 2015 / Accepted: 20 October 2015 / Published: 23 October 2015

---

**Abstract:** The Sentinel missions have been designed to support the operational services of the Copernicus program, ensuring long-term availability of data for a wide range of spectral, spatial and temporal resolutions. In particular, Sentinel-2 (S-2) data with improved high spatial resolution and higher revisit frequency (five days with the pair of satellites in operation) will play a fundamental role in recording land cover types and monitoring land cover changes at regular intervals. Nevertheless, cloud coverage usually hinders the time series availability and consequently the continuous land surface monitoring. In an attempt to alleviate this limitation, the synergistic use of instruments with different features is investigated, aiming at the future synergy of the S-2 MultiSpectral Instrument (MSI) and Sentinel-3 (S-3) Ocean and Land Colour Instrument (OLCI). To that end, an unmixing model is proposed with the intention of integrating the benefits of the two Sentinel missions, when both in orbit, in one composite image. The main goal is to fill the data gaps in the S-2 record, based on the more frequent information of the S-3 time series. The proposed fusion model has been applied on MODIS (MOD09GA L2G) and SPOT4 (Take 5) data and the experimental results have demonstrated that the approach has high potential. However, the different acquisition characteristics of the sensors, *i.e.* illumination and viewing geometry, should be

taken into consideration and bidirectional effects correction has to be performed in order to reduce noise in the reflectance time series.

**Keywords:** data fusion; spatial-spectral unmixing; multi-sensor; multi-temporal; temporal weights

---

## 1. Introduction

The improved observation capabilities of the upcoming Copernicus S-2 and -3 missions are expected to fulfill the current operational monitoring requirements for environmental and security purposes. Providing long-term data in a wide range of spectral, spatial and temporal resolutions, the Sentinel satellites will be essential for future mapping and monitoring applications related to various research fields, such as agriculture, forestry, inland and coastal water quality, wetlands, land degradation, *etc.* [1]. The need for developing new processing methods and tools related to time series of high resolution has been recently pointed out by the scientific community, towards the full exploitation of the significant scientific, operational and commercial potential of the Sentinel data. In particular, one of the recommendations in the recent S-2 workshops was the synergy between S-2 with other Sentinels and/or other satellite missions, in order to increase the observation capabilities, especially during cloudy periods and regions. Considering this, the synergistic use of multi-sensor data is investigated in this research with the aim of increasing image availability in the cases of cloud coverage. Image fusion methods have been investigated thoroughly in the last few years, combining in one composite image the improved spectral, spatial and temporal features of multiple datasets acquired by different sensors and on different dates. The developed methodology in this research is based on the one introduced by Zhukov *et al.* [2] in order to fuse the thermal band with the corresponding reflective bands of Landsat Thematic Mapper (TM). The unmixing process is accomplished in the following steps: (1) the classification of the higher spatial resolution image (HSpaR); (2) the estimation of each class contribution to the signal of the lower spatial resolution image (LSpaR); (3) the calculation of the pure spectra (endmember) for every class; and (4) the restoration of the unmixed image pixel. The procedure is window-based, meaning that the LSpaR pixel to be unmixed is the central pixel of a window and the unmixing process is performed based on the contextual information of all the pixels in the window. The proposed unmixing methodology demonstrated a significant enhancement in sharpness and radiometric accuracy of the fusion output in comparison to the input data.

The fusion approach of Zhukov *et al.* [2] in its original and alternate forms has found many applications on a variety of imagery. Minghelli-Roman *et al.* [3] combined the bands of simulated Envisat/MERIS and Landsat/TM data to generate an image that combines the best characteristics of each sensor. The experimental results indicated that the approach is advantageous over other unmixing methods that require *a-priori* knowledge of endmembers and their spectral profiles. Nevertheless, its main drawback is the lack of spectral variability among the pixels belonging to the same class, as they all take the same pixel value in the fused image. Zurita-Milla *et al.* [4,5] have also presented a sophisticated implementation of a linear spectral mixture model to downscale a time series of Envisat/MERIS based on the high spatial-resolution information of a land-cover database. The accuracy

assessment of the resulting images indicated the potential of the proposed image fusion analysis. However, the problem of such applications is that updated land-cover databases are not always available.

In order to overcome the aforementioned shortcomings of unmixing approaches, an alternative approach was proposed by Amorós-López *et al.* [6]. More specifically, the suggested methodology employed a soft clustering approach for the classification of the HSpaR image, with the goal of preserving the spectral variability among pixels belonging to the same class within the analyzed window. The clustering algorithm of Self-Organizing Map (SOM) was selected to estimate the membership pixel value to each of the clusters. The fusion procedure followed the processing steps of Zhukov *et al.* [2], but instead of using the “hard” value of a pixel correspondence on only one class, it employs the fuzzy membership values on all possible classes. The proposed fusion process was implemented on Envisat/MERIS Full Resolution and Landsat/TM images, with the qualitative and quantitative assessment to evaluate its performance in data fusion. A similar approach was also followed to obtain a time series of Landsat-like images with high-frequency and the image products were evaluated through a crop monitoring analysis, indicating the accuracy and significance of this methodology in such applications [7].

Furthermore, in order to estimate the daily surface reflectance at Landsat spatial resolution, Gao *et al.* [8] proposed a spatial and temporal adaptive reflectance fusion model (STARFM) to composite Landsat and MODIS surface reflectance. Taking into consideration the spectral, temporal and spatial similarities between the central pixel of a window and those of each of the neighboring pixels, a weight function was determined to define their similarity and contribution to the unmixing procedure. Implementations of this approach showed accurate results, preserving the improved features of the blending data, (*i.e.* Landsat high spatial resolution and MODIS high temporal resolution) [8–11]. However, the algorithm did not succeed in predicting the surface reflectance in two cases: (1) when there were transitory changes between the acquisition dates of Landsat images and thus they were not captured; and (2) when complex land cover areas were involved, and it was difficult to identify “pure” MODIS pixels in the sliding window [8,12]. In order to detect the spatial changes and manage the STARFM limitation concerning the transient changes, the Spatial Temporal Adaptive Algorithm for mapping Reflectance Change (STAARCH) model was introduced by Hilker *et al.* [13]. Based on Tasseled Cap transformations of both Landsat and MODIS reflectance data, the algorithm detects vegetation changes and employs each optimal Landsat-MODIS image pair in the fusion process. In order to address the issue of accurate surface reflectance in the heterogeneous areas, another enhanced version of STARFM, the ESTARFM model, was presented by Zhu *et al.* [12]. Under the assumption of a linear relationship among the endmembers’ reflectances (“pure” pixels) of multi-temporal observations of a short time period and based on the ratio of the reflectance change of the fine resolution pixels (considered as endmembers) to the corresponding mixed coarse pixel, they defined a conversion coefficient. The ESTARFM model was effective in estimating even small object reflectance and, thus, in handling changing and heterogeneous landscapes.

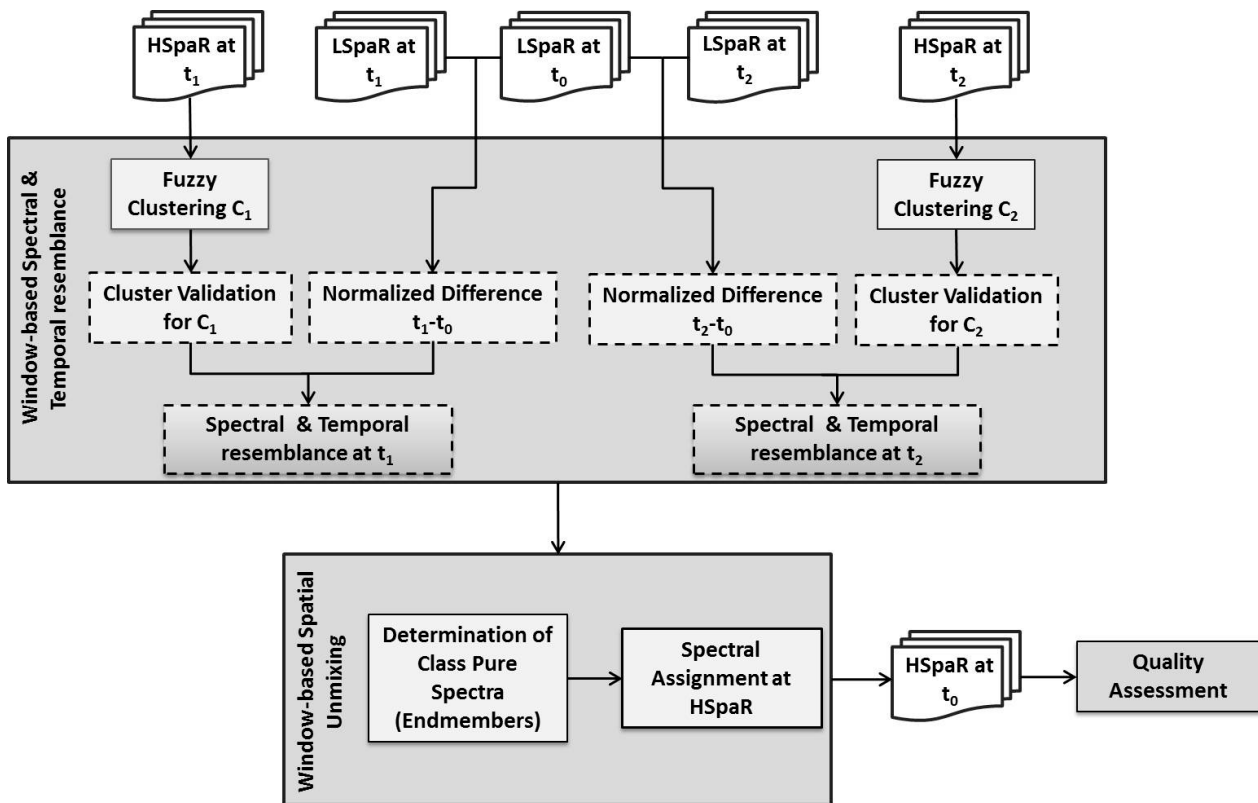
Zhang *et al.* [14] estimated the surface reflectance on the prediction date based on the temporally weighted blending result of the base dates. In their approach, the classification outcome of the stacked HSpaR datasets (two base dates) served for the unmixing of the LSpaR datasets of three dates (the base dates and the prediction date). The temporal weights were employed to handle the resulting unmixed images that generated to the final predicted fused image. In this case, the patch-based classification of the HSpaR data was proposed, which is mainly an object-based unsupervised classification method,

aiming to cope with the “salt-and-pepper” noise effect in the classification output. However, the suggested classification algorithm did not manage to handle the spectral variability of the pixels in the LSpaR analyzed pixel, and the methodology was affected by the aforementioned issue of unmixing approaches based on “hard” clustering algorithms. The other significant drawback of the proposed method is that the window-based unmixing process, which is in general time-demanding, must be implemented three times for each case study. Moreover, if the required test implementations for the window size and clusters’ number definition (trial-and-error processes) are taken into consideration, the computation time increases significantly.

The algorithm developed in the present study is an enhanced version of those proposed in [2,7,14]. The improvements target the clustering approach and the efficient management of the time series. In particular, the definition of the optimal number of clusters in each case study is embedded in the algorithm, with the goal of eliminating user’s involvement in the procedure. Clustering validation measures are engaged to assess the clustering algorithm performance and determine the cluster number that best fits the data. As such, the clustering and unmixing results are enhanced, and the processing time is reduced. In all previous studies such a task required a trial-and-error process, meaning many executions of the entire procedure and definition of the optimal cluster number based on the fusion product quality. Additionally, we propose to classify separately each of the available HSpaR imagery and include them in the unmixing process by assigning temporal weights to them. Thus, the similarity between the LSpaR image pixels of the prediction date and the base dates defines the corresponding contribution of the HSpaR pixel to the unmixing process. All the transition changes throughout the study period are therefore taken into account and not only the ones between the base dates, as in the case of classification of the stacked multitemporal HSpaR dataset, suggested by previous studies [7,14]. The proposed image fusion methodology is described in Section 2, including the clustering process, the automatic definition of the optimal number of clusters and the window-based spatial unmixing. The experimental results on the imagery and their quantitative and qualitative evaluation are analyzed in Section 3. Section 4 summarizes the experimental results and improvements introduced by this study’s fusion methodology.

## 2. Methodology

The overall developed methodology for the unmixing fusion is implemented on a sliding window (Figure 1). A LSpaR image acquired on the date  $t_0$  is unmixed based on the information of the two HSpaR images with the closest acquisition dates before and after  $t_0$ , i.e.  $t_1$  and  $t_2$ . The unmixing process is accomplished within the following steps. (a) *Spectral resemblance estimation*: An unsupervised fuzzy classification (clustering) is applied on HSpaR images, in order to acquire the land cover types  $C_1$  and  $C_2$  at the two base dates  $t_1$  and  $t_2$ . Those classes’ contribution to LSpaR image pixel of  $t_0$  will define the spectral resemblance between the consequent study dates; (b) *Temporal resemblance estimation*: The calculation of temporal weights based on the normalized differences of the LSpaR images between the consequent study dates; (c) *Endmembers determination*: The assessment of classes’ endmembers through a system of a linear mixture equations; (d) *Spectral Assignment at HSpaR*: A HSpaR image pixel at  $t_0$  is generated based on the linear combination of the resulting classes’ endmembers, the classes’ contribution and their respective temporal weights.



**Figure 1.** The overall developed unmixing fusion methodology. The processes with the dash type outline are the ones embedded in the new algorithm in order to improve the time and quality performance.

2.1. Unsupervised Fuzzy Classification

The essential land cover information for the unmixing process resulted from the unsupervised fuzzy classification (fuzzy clustering) of the HSpaR images. According to the concept of the fuzzy theory, every image pixel can have a membership value, ranging from 0 to 1, to all the clusters indicating this way its similarity to the corresponding cluster. Such flexibility makes fuzzy clustering more suitable to handle indistinct boundaries, which are typically met in the natural environment [15,16]. In order to derive the land cover information of the study area, this paper employs the Fuzzy Maximum Likelihood Estimation (FMLE) clustering. When using the FMLE approach, the estimation of membership values depends on an exponential function that employs the covariance matrix  $F_i$  and the prior probability  $a_i$  of selecting  $i^{th}$  cluster.

$$d_{ij}^2(X_j, V_i) = \frac{\sqrt{\det(F_i)}}{a_i} \exp \left[ (X_j - V_i)^T F^{-1} (X_j - V_i) / 2 \right] \tag{1}$$

where  $X_j$  is the  $j$ -bands image vector and  $V = (v_1, v_2, \dots, v_k)$  is the vector of cluster centers (*i.e.* the means of the clusters). The FMLE is able to deal with the problem of large variability in cluster shapes, sizes and densities, but it requires a good initialization. Therefore, it is recommended to initially estimate the cluster centroids with the fuzzy c-means clustering algorithm, in order to obtain an optimal fuzzy partition with the FMLE in the next phase [15].

However, regardless of the algorithm used, the critical issue in clustering is that the land cover types are not known *a priori* (no available ground reference information) and in each case study, the number of clusters must be assumed by the user [15]. A good way to deal with this issue is to apply certain validation criteria with the aim of evaluating the clustering performance within a range of cluster numbers. The optimal data clustering is achieved when the clusters are well-separated and have minimal volume, and the data points are close to the cluster centroids (high membership values) [15,17–19]. A variety of validation indices in the literature measure the separation, partition and compactness of the clusters [17,18,20]. It is noted that none of the indices can be separate or effective to all the data sets, as each of them has different measure context (e.g., compactness, separability, fuzziness), therefore their combination is proposed for defining the optimal cluster number [20]. Hence, in this study certain validity measures are embedded in the clustering procedure determining the ideal number of clusters. More specifically, the following indices were calculated: the Fuzzy Hypervolume Validity (FHV) and the Partition Density (PD), involving the fuzzy covariance matrix of the cluster and the membership; the Partition Index (SC), the ratio of the sum of compactness and separation of the clusters; the Separation Index (S), similar to SC but divided by minimum-distance separation; and Xie and Beni's Index (XB), the ratio of the total variation within clusters and the separation of clusters [19,20]. Generally, it is recommended that when the difference between the validity indices is insignificant, fewer clusters are better [19].

## 2.2. Spectral Unmixing Concept

The unmixing procedure is based on a moving window, in which the spectral reflectances of the pure HSpaR pixels (endmembers) of the central LSpaR pixel are estimated by taking into account all the pixels in the window. Namely, a number of linear equations (Equation (2)), as many as the LSpaR pixels in the window, is defined, where  $S^{LSpaR}$  is the LSpaR pixel value,  $C_k^{LSpaR}$  is the contribution of the land cover types  $k$  of HSpaR in the LSpaR pixel,  $E$  is the pure pixel signal (endmembers) of the land cover types  $k$ , and  $e$  is the residuals of the linear model.

$$S^{LSpaR} = C_k^{LSpaR} \cdot E + e \quad (2)$$

The solution of the linear system is given by the least squares method accounting for the following:

1. The class contribution to the LSpaR image pixel,  $C_k^{LSpaR}$ , which is the proportion of each class in the LSpaR image pixel and is estimated using membership values  $U_k^{HSpaR}$  to the corresponding class  $k$ . Explicitly,  $C_k^{LSpaR}$  is the average of all the class contributions  $U_k^{HSpaR}$  to LSpaR pixel footprint  $N$ , according to Equation (3). The critical issue in this case pertains to ensuring accurate co-registration between the images. In this study, the LSpaR image pixel footprint is resampled on HSpaR image pixel size, considering that the point spread function (PSF) is rectangular. Therefore,  $N$  is the total number of HSpaR image pixels  $i$  within the LSpaR image pixel footprint.

$$C_k^{LSpaR} = \frac{1}{N} \sum_{i \in N} U_{ik}^{HSpaR} \quad (3)$$

2. The spectral difference between the LSpaR image of the predicted date  $t_0$  and that of each of the base dates  $t_1$  and  $t_2$ , in order to handle changes between multitemporal observations efficiently.

Based on these differences, the normalized temporal weights  $T_1$  and  $T_2$  are calculated for every pixel in the window (Equation (4)) and being later involved in the unmixing process.

$$T_{t=1,2} = \frac{1 / \left( \sum_{t \in w}^{i \in w} (|S^{LSpaR}_t - S^{LSpaR}_{t_0}|) \right)}{\sum_{t=1,2}^{i \in w} \left( 1 / \left( \sum_{t \in w}^{i \in w} (|S^{LSpaR}_t - S^{LSpaR}_{t_0}|) \right) \right)} \quad (4)$$

The temporal weights  $T_1$  and  $T_2$  are estimated per window, the weights being assigned to the central pixel of the sliding window with size  $w$ . The least changes occurred between the two dates, the greater the similarity between the corresponding image pixels and thus the higher the value of the temporal weight  $T$  for the selected pixel.

In essence, taking into consideration the available HSpaR and LSpaR images, the classes' contributions  $C^{LSpaR}$  to the LSpaR image pixel together with the temporal weights  $T_1$  and  $T_2$  allow the multitemporal analysis of the spectral resemblance within the imagery. Therefore, the initial Equation (2) becomes:

$$S^{LSpaR} = (T_1 \cdot C_{t_1}^{LSpaR} E_1 + T_2 \cdot C_{t_2}^{LSpaR} E_2) + e \quad (5)$$

In order to avoid the propagation of possible classification or co-registration errors to the endmembers calculation, any class with a contribution less than 5% ( $C_t^{LSpaR} < 0.05$ ) is discarded. The inversion of the equation system (Equation (5)) and the retrieval of the endmembers  $E$  is achieved by implementing the least-squares method on each band separately.

In the final step, the values of the fused image pixels  $S^F$  are derived by assigning the estimated endmembers  $E$  to every HSpaR pixel according to the corresponding class. The class membership values  $U_{tk}^{HSpaR}$  and the temporal weights  $T$  are accounted for, and thus preserving the spectral and temporal variability inside the analyzed window.

$$S^F = (T_1 \cdot U_{t_1 k}^{HSpaR} E_1 + T_2 \cdot U_{t_2 k}^{HSpaR} E_2) \quad (6)$$

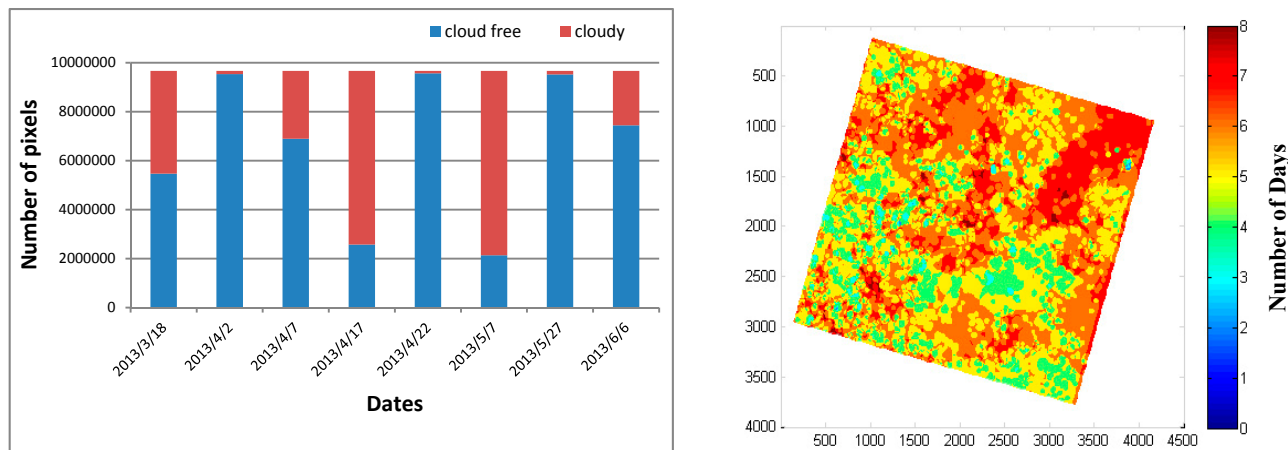
### 3. Experimental Results and Evaluation

#### 3.1. Dataset and Study Area

The HSpaR imagery in our case study is a time series with a five-day revisit derived from SPOT-4 (Take5). The data were provided by European Space Agency (ESA) in collaboration with French Space Agency (CNES) in the framework of an experiment concerning applications and methods, before the launch of S-2. The SPOT Level2A (ORTHO\_SURF\_CORR\_PENTE) dataset provides surface reflectances corrected from atmospheric and terrain effects. The acquisition dates of the imagery are from January to June of 2013. However, due to cloud cover, images are not available every five days throughout the entire period; this limitation is observed to a different extent for all the available sites [21].

The study area (20 × 25 km) is in Tienen (Belgium), a mainly agricultural area and one of the study sites of Joint Experiment for Crop Assessment and Monitoring (JECAM). According to JECAM's requirements for Earth Observation optical data, the spatial resolution should be from 20 m to 5 m. Additionally, during the study period (from February/March to August), the temporal frequency should be every 15 days, although weekly acquisitions would be preferred. The available dates and the

corresponding cloud-free pixels of the entire SPOT4 scene are presented in Figure 2. A partial cloud cover is present in most of the images, there hardly being one clear observation per month throughout the entire period.



**Figure 2.** The image pixels are presented per day in a bar chart of cloud-free and cloudy pixels (**left**). The study area is displayed in pixel coordinate system and in colors according to the availability of cloud free observations per day (**right**).

The LSpaR imagery is a time series of MODIS surface reflectance product (MOD09GA), which was re-projected from the native Sinusoidal projection to the UTM\_WGS84 through the Google Earth Engine re-projection tool. Although the MOD09GA dataset is available on a daily basis, the cloud coverage hinders daily image availability in this particular instance. The two datasets have corresponding bandwidths with the MODIS land bands being narrower than the SPOT-4 ones (Table 1).

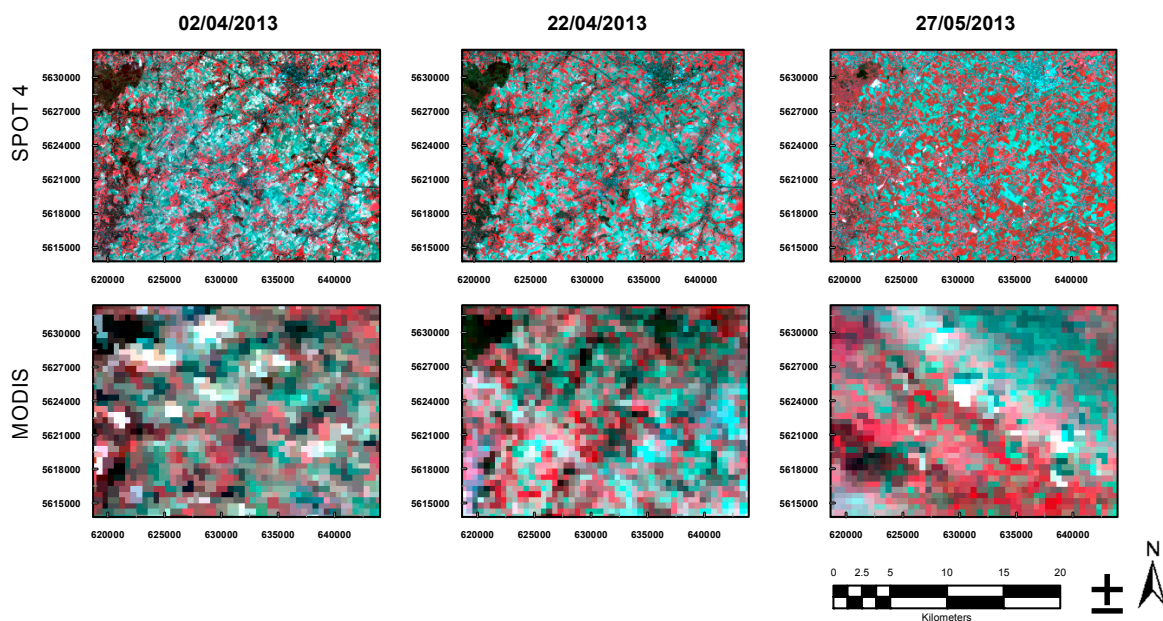
**Table 1.** The bandwidths of MODIS and SPOT-4 bands.

MODIS Bands MOD09GA L2G	MODIS Bandwidth (nm)	SPOT4 Bands Level 2A	SPOT Bandwidth (nm)
Band 3/Blue	459–479		
Band 4/Green	545–565	XS1/Green	500–590
Band 1/Red	620–670	XS2/Red	610–680
Band 2/NIR	841–876	XS3/NIR	790–890
Band 6/SWIR1	1628–1652	SWIR	1530–1750
Band 7/SWIR2	2105–2155		

The time period selected for applying the developed unmixing procedure is from April to May and the corresponding dates of the available cloud free images are presented in Table 2. Assuming that a HSpaR cloud free image is not available on 22 April 2013, the MODIS pixel of the corresponding date is unmixed based on the classification of the available SPOT images before (2 April 2013) and after (27 May 2013) the missing acquisition date. The MODIS images of all study dates are also involved into the unmixing process and specifically for the temporal weights estimation. In order to validate the



developed methodology, the SPOT-4 image of 22 April 2013 was used as a “reference image” in order to assess the accuracy of the unmixing result (Table 2, Figure 3).



**Figure 3.** The study area as depicted on 2 April 2013 and 27 May 2013 (base dates) and on 22 April 2013 (prediction date) on SPOT 4 (**upper row**) and MODIS (**bottom row**) images. The multitemporal images are illustrated in false-color composites (R: NIR, G: Red, B: Green).

**Table 2.** Description of the datasets.

Imagery	Acquisition Dates	Usage	Spatial Resolution Radiometric Values	Geometric Reference	
MODIS MOD09GA L2G	2 April 2013	Unmixing	Surface Reflectance	Map Projection: UTM Ellipsoid Type: WGS84	
	22 April 2013	Unmixing			500 m
	27 May 2013	Unmixing			
SPOT4 (Take 5) Level 2A	2 April 2013	Classification	20 m		
	22 April 2013	Validation			
	27 May 2013	Classification			

### 3.2. Clustering Results and Optimal Cluster Number

The land cover types of the study area were the output of FMLE clustering, with the number of clusters to be defined by applying the validation measures described in Section 2.1. The optimal clusters’ number was explored in the range of cluster number  $k = [5, 40]$  for both HSpaR images and it was based on the best value for each case, *i.e.* local minimum or maximum of each criterion. For the sake of simplicity, in Table 3, only the values of the validation measures for the image of 27 May are presented in the range of  $k = [15, 40]$ . Concerning the FHV and PD, the corresponding local minimum and maximum values were acquired for  $k = 33$ , indicating also the optimum value. Moreover, the value  $k = 33$  can be confirmed as ideal by SE, S and XB, although the corresponding increasing or decreasing trend in these cases is rather monotonic. Similarly, the clustering results analysis of the HSpaR image of 2 April 2013 indicated the cluster value  $k = 34$  as the optimal one. The automated optimal selection of

the clusters' number does not only improve the classification results, but it also facilitates the unmixing procedure through faster processing times and increased accuracy.

**Table 3.** The results of clustering validation in the range of 15 to 40 clusters for the HSPaR image *t2*.

Clusters Criteria	Optimal Value	15	16	17	18	19	20	21	22	23	24	25	26	27
		<b>FHV e-06</b>	min	3.34	3.10	3.90	3.93	3.88	4.81	4.82	4.73	4.80	4.85	5.14
<b>PD e+05</b>	max	14.77	15.21	15.92	16.44	16.87	17.34	18.80	19.28	20.75	21.18	22.54	23.96	24.80
<b>SC</b>	max	0.82	0.87	0.82	0.89	0.92	0.92	0.94	0.93	0.95	0.94	0.93	0.90	0.92
<b>S e-07</b>	min	9.49	10.96	10.59	11.86	11.91	12.30	12.22	12.31	12.41	13.07	13.85	13.35	13.73
<b>XB</b>	min	8.24	7.59	6.49	6.51	5.54	5.14	5.43	4.15	4.24	4.24	4.00	3.61	3.36
Clusters Criteria	Optimal Value	28	29	30	31	32	33	34	35	36	37	38	39	40
		<b>Fhv e-06</b>	min	5.25	5.00	4.74	4.54	4.37	3.91	4.03	4.21	4.39	4.32	4.17
<b>PD e+05</b>	max	25.18	26.86	28.51	29.66	31.17	33.72	30.81	30.91	30.94	30.98	30.06	30.13	39.21
<b>SC</b>	max	0.94	0.95	0.93	0.95	0.94	0.98	0.95	0.96	0.98	1.00	0.98	0.99	0.99
<b>S e-07</b>	min	13.59	13.75	13.13	13.67	13.33	12.92	13.48	13.87	14.01	14.15	13.75	13.99	14.00
<b>XB</b>	min	3.97	3.60	3.41	3.10	3.04	2.57	2.89	3.01	2.99	3.07	2.39	2.39	2.52

### 3.3. Spectral-Unmixing Fusion

The qualitative and quantitative evaluation of the unmixing output was performed based on selected quality measures indicating the similarity between the images. A visual inspection and comparison of the output fused image with the reference SPOT image indicated the areas with similarities and differences, thus the areas where the algorithm performed well or not. Multispectral images, as well as the corresponding Normalized Difference Vegetation Index (NDVI) images were utilized for the visual comparison. Concerning the quantitative assessment, the result was compared to the actual SPOT image using some of the data fusion quality indicators, *i.e.* RMSE, Pearson's correlation coefficient, Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS) and Q4 index. The standard measure of the ERGAS index gives a global indication of the fusion output quality [22] (Equation (7)),

$$ERGAS = 100 \frac{h}{l} \sqrt{\frac{1}{N} \sum_{i=1}^N \frac{RMSE_i^2}{M_i^2}} \tag{7}$$

where, *h* is the resolution of the high spatial resolution image (20 m herein); *l* is the resolution of the low spatial resolution image (500 m herein); *N* is the number of spectral bands involved in the fusion (four or seven herein); *RMSE<sub>i</sub>* is the root mean square error computed between the fused image and the original low resolution image (for the band *i*); and *M<sub>i</sub>* is the mean value of the band *i* of the reference image. Taking into consideration that the resultant fused image should be identical to the original low-resolution image, in terms of spectral information, the ERGAS index value should be close to zero.

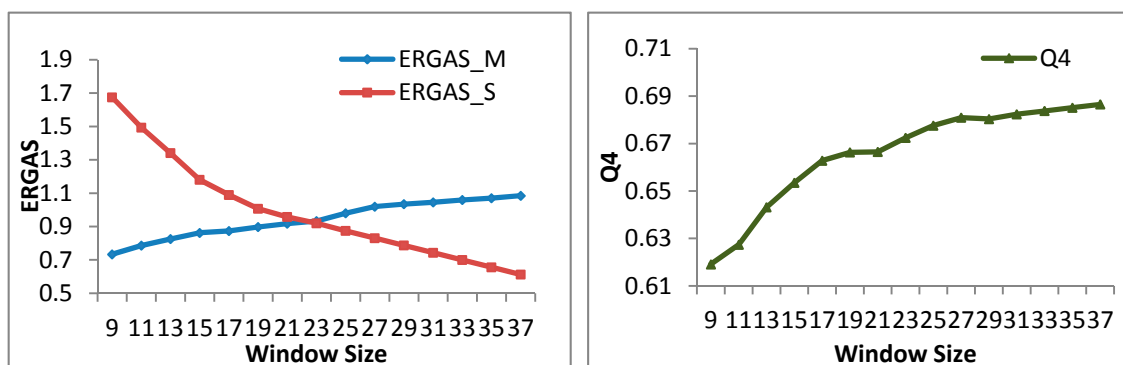
Moreover, the Q4 index was employed as it is considered an appropriate quality measure of a fusion product, when two multispectral images with four bands are involved. It is based on the modulus of the

hypercomplex correlation coefficient, mean bias and contrast variation and it is sensitive to spectral and radiometric variation, since it handles all the bands simultaneously. The Q4 index values range between [0, 1], with the highest value representing the greatest similarity between the images [23]. If the four bands of the two images are expressed and compared as quaternions  $a = a_1 + ia_2 + ja_3 + ka_4$  and  $b = b_1 + ib_2 + jb_3 + kb_4$ , then the Q4 index is defined as

$$Q4 = \frac{|\sigma_a \sigma_b|}{\sigma_a \cdot \sigma_b} \cdot \frac{2 \sigma_a \cdot \sigma_b}{\sigma_a^2 + \sigma_b^2} \cdot \frac{2|\bar{a}||\bar{b}|}{|\bar{a}|^2 + |\bar{b}|^2} \quad (8)$$

All terms of Equation (8) are calculated on  $N \times N$  image blocks, either  $16 \times 16$  or  $32 \times 32$ , and the average of all resulting values is the final value of Q4 over the entire image. The Absolute Average Difference and the Average Difference were also calculated to define the magnitude of difference between the images and their positive or negative deviation.

As the proposed unmixing approach is a window-based process, one of the user-dependent parameters is the window size. In this study the algorithm was tested in a range of window sizes ( $w$ ) [9, 37] and its performance was evaluated by implementing the aforementioned criteria. The assessment in all the cases was conducted using the corresponding bands of the input imagery data, namely the four bands of SPOT 4 and bands 4, 1, 2, and 6 of MODIS imagery (Table 1). The error computation was mainly twofold, at low and high spatial resolution, *i.e.* ERGAS\_M and ERGAS\_S, in order to estimate the respective spectral and spatial distortion. The ERGAS results indicate that the trend of the LSpaR image index is opposite to that of HSpaR image, and thus there is not optimal window size for both cases (Figure 4). Nevertheless, a close inspection of Figure 4 indicates that ERGAS\_M has rather small values without great variations, opposite to ERGAS\_S that decreases significantly, when the window size increases. Therefore, an acceptable option would be to select a window size that minimizes ERGAS\_S and that also keeps ERGAS\_M relatively low. The Q4 index, which was calculated on  $16 \times 16$  image blocks, has similar behavior to ERGAS\_S, meaning that the bigger the window size, the better the index value. Considering all preliminary conditions, the window size  $w = 37$  was chosen in our study, the corresponding unmixing results being presented in Figure 4.



**Figure 4.** The trend curves of ERGAS (for the LSpaR and HSpaR image) and the Q4 index (only applicable on HSpaR) as a function of different window sizes.

In Figure 5, different regions of the study area are presented as subsets of the available SPOT imagery and the fusion product. A visual inspection of the images indicates that the algorithm succeeded in

capturing most temporal changes of the croplands' in the study area (yellow rectangles). Areas with land cover changes from 2 April 2013 to 27 May 2013 are shown as purple rectangles; these are areas that remain unchanged from 2 April 2013 to 22 April 2013, fact that demonstrates that temporal weights are necessary to handle the transition changes throughout the study period. In the case of the stack classified image, proposed in previous studies [7,14], such changes between the base dates are classified as one cluster, either these changes occurred between the first base date and the prediction date or not, and vice versa. For this reason, the separate classification of each image is proposed, in order to consider all the land cover types and use them based on the temporal weights. Nevertheless, the algorithm was not able to accurately estimate the radiometric values of the areas with transition changes among the dates, but with similar land cover types in the base dates (blue rectangles). It should also be noted that the fusion outcome diverges from the original image mainly in the areas with high heterogeneity, where the “salt-and-pepper” effect is noticeable. Such noise is common, when pixel-based classification methodologies are applied on high spatial resolution images. A possible solution to this issue would be an object-based clustering, but in the case of unmixing there would not be any spectral variability inside the objects of the fusion output, and thus it is not recommended in our study. Alternatively, an edge preserving filtering of the images would be suitable for handling spectral variability and reducing the noise present in classification and fusion results [24].

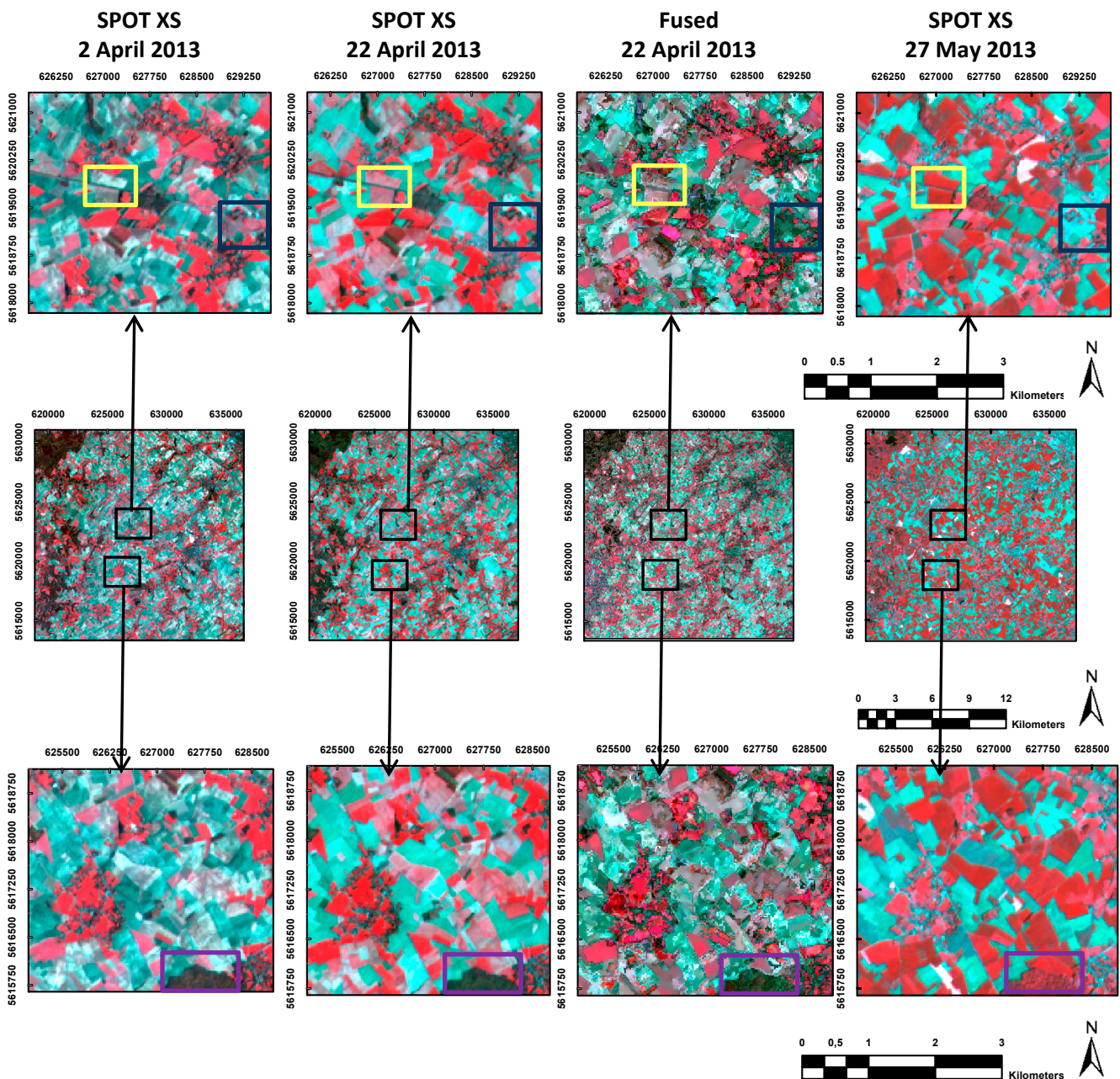
The quantitative results of the fusion quality assessment between the fusion output and the reference SPOT image of 22 April 2013 are shown in Table 4. In all metrics, the prediction errors are higher for the NIR and SWIR bands. A partial explanation is the heterogeneity of the vegetation land cover type and phenological changes that prevent an accurate prediction typically for these bandwidths. Moreover, the relatively low correlation between the surface reflectances of the input data (Figure 6) has an impact on the correlation between the actual and the output pixel values. This high variance can be attributed to the different acquisitions dates and sensors characteristics; as such, in order to perform the data inter-comparison, a directional effect correction is recommended [21,25].

**Table 4.** The quantitative results of the comparison between the unmixing output image and the corresponding SPOT image on 22 April 2013 (for the optimal window size,  $w = 37$ ).

Bands	RMSE	ERGAS_S	CORR	Q4	AvAbsDiff	AvDiff
Green	0.0281		0.7808		0.0229	−0.0131
Red	0.0296		0.7961		0.0241	−0.0009
NIR	0.0458	0.612	0.7443	0.686	0.0413	0.0201
SWIR	0.0496		0.7477		0.0409	−0.0142

The average RMSE of the current methodology is 0.038, while that of similar studies was 0.035 [26], 0.039 [7] and around 0.04 for the different parameters' combination in the approach of Amorós-López *et al.* [6]. Additionally, the Q4 index is 0.68 in our study, whereas a higher value (0.86) was computed by a similar study [7]. The AvAbsDiff for Green, Red, and NIR are 0.023, 0.024 and 0.041 in the current study, compared to 0.0073, 0.009 and 0.017 respectively as derived from the image fusion of Landsat ETM+ and MODIS land data [14]. We should mention that this results' comparison is indicative. The

imagery and the study area are different among the methodologies and thus the inter-comparison of the results would not produce reliable conclusions.

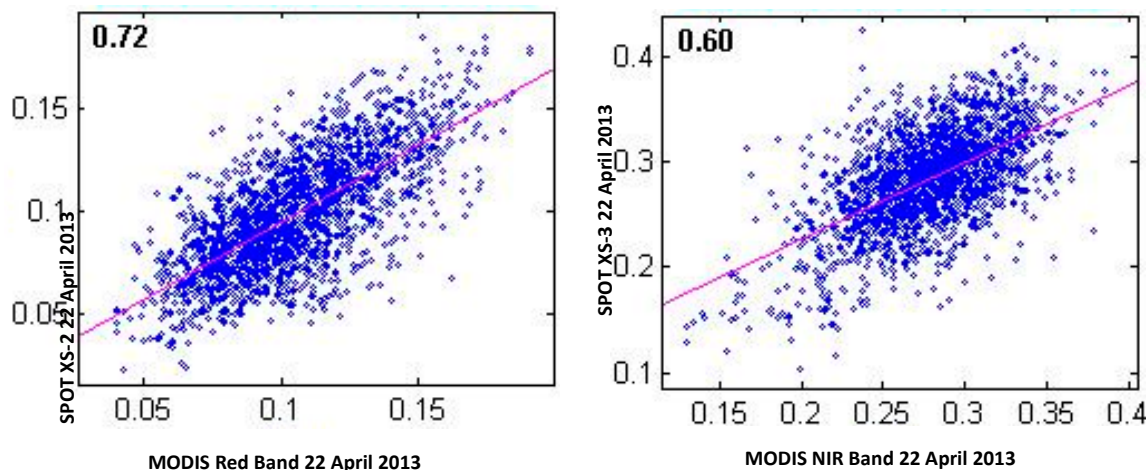


**Figure 5.** The available SPOT imagery on the three acquisition dates and the fusion product on 22 April 2013. Comparison of the images indicates that different types of transition changes are evident during the study period with most phenology changes (yellow rectangles) to be estimated correctly by the proposed algorithm. Accurate results were also produced in the case of a change only between two dates (purple rectangles), while the algorithm failed in estimating the radiometric values on the prediction date, when change was not evident in imagery of the base dates (blue rectangles).

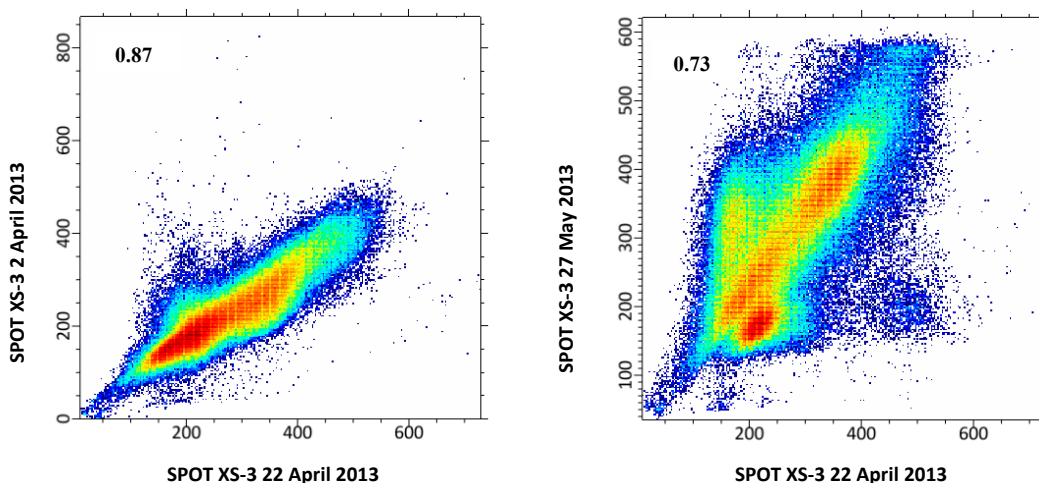
It is also important to be noted that during the study period, significant changes occurred in the area, mainly due to the phenological cycle. Hence, this explains the low correlation between the corresponding

NIR bands of the study dates (Figure 7). Nevertheless, the algorithm was able to handle these changes through implementation of temporal weights, as well as to estimate the reflectance values on the prediction date relatively accurate.

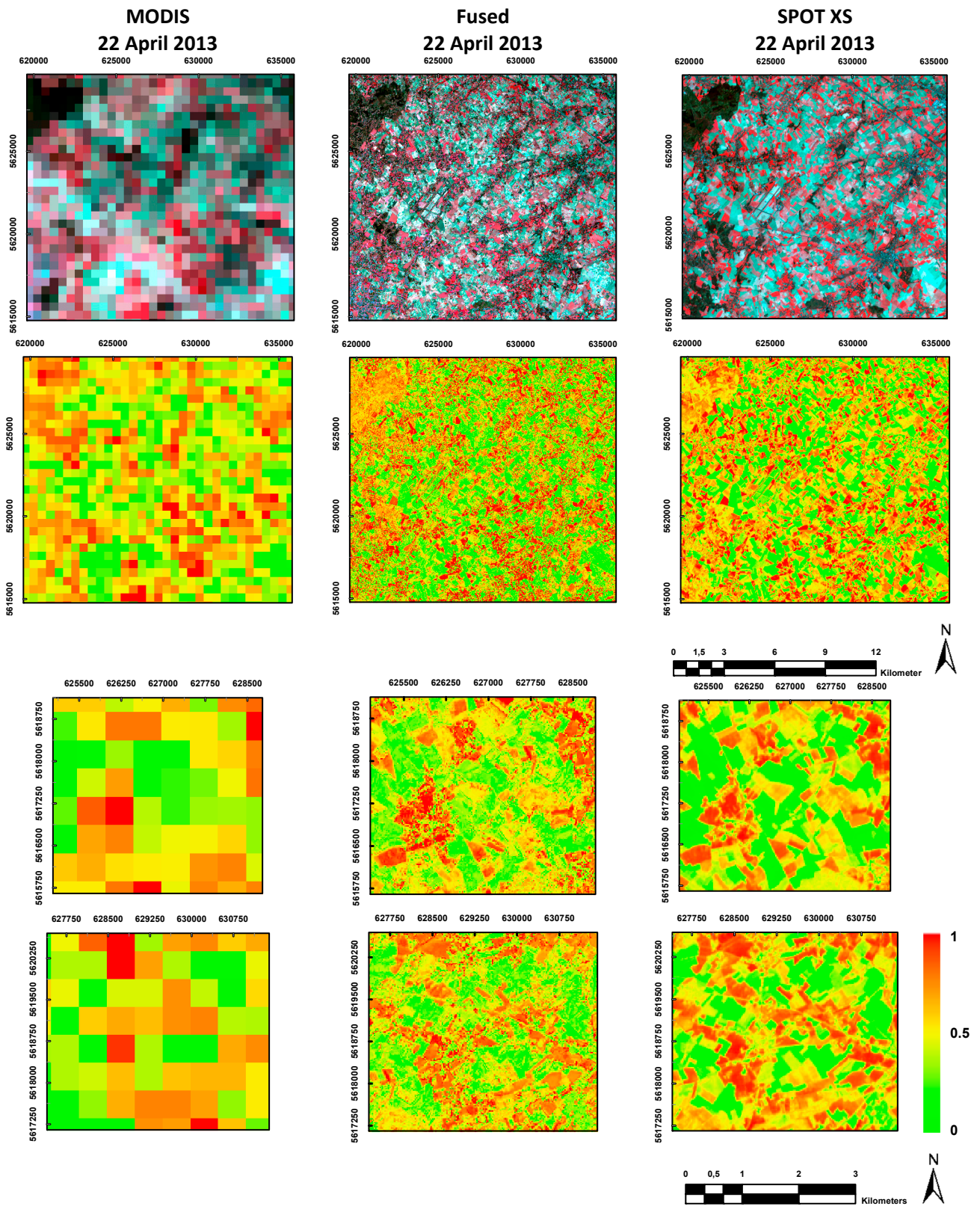
The NDVI results of the input observations and the fusion output are presented in Figure 8. Comparing visually the corresponding image areas, the resulting image has a high resemblance to both images on base dates. The cropland pattern is successfully captured by the predicted NDVI values, although there is a difference in the actual pixel values in certain areas, fact explained by the observed differences in the Red and NIR bands.



**Figure 6.** The scatter plots of the surface reflectance and the corresponding correlation of Red (0.72) and NIR (0.60) bands of the MODIS and SPOT images of the prediction date (22 April 2013). The SPOT pixel values were aggregated to the MODIS pixel size.



**Figure 7.** The scatter plots of SPOT XS-3 bands of the consequent study dates. The low correlation between the bands indicates that important phenology changes occurred in the area mainly during the 22 April 2013 to 27 May 2013 period.



**Figure 8.** The input imagery and the fusion product of the prediction date 22 April 2013 (first row), the corresponding NDVI images (second row) and some NDVI image subsets at large scale (third and fourth rows).

### 3.4. Implementation on the Upcoming Data of Sentinel-2 MSI and Sentinel-3 OLCI

The multispectral/multisensor data integration was tested and evaluated on SPOT and MODIS imagery with the intention of applying to the advanced features of S-2 MSI and S-3 OLCI instruments. Following the corresponding process applied to SPOT data, the S-2 VNIR bands with a spatial resolution of 10 m will be the input data for land cover classification. Improved results are anticipated when vegetated areas are involved, as the red-edge bands with 20 m spatial resolution will be also considered in the clustering process. The land cover information in this case will support the unmixing of the 21 bands of S-3 OLCI with a spatial resolution of 300 m. As the pair of S-3 satellites will enable a short revisit time of less than two days for OLCI, more cloud-free observations are expected to be acquired than with MSI (five days revisit). Therefore, the suggested unmixing fusion would provide data with high spatial, spectral and temporal resolution, which will significantly benefit land services and allow the monitoring of land-cover dynamics, forest cover, photosynthetic activity and soil quality. Nevertheless, the directional effect needs to be corrected, in order to reduce the time series noise considering that the two instruments will have different viewing geometry.

## 4. Conclusions

The unmixing data fusion approach developed in this study aims at filling in the gaps of a time series, when cloud free images are not available. The main improvements introduced by the suggested approach are:

1. The automatic definition of the optimal cluster number with the help of various evaluation criteria, establishing a faster and more robust methodology. The previous approaches required a trial-and-error process in this methodology step, meaning that the unmixing algorithm had to be tested in a range of cluster numbers. By evaluating the unmixing result for each of the different clustering output in those cases, the corresponding cluster number was defined. Estimating the optimal cluster number, without the need to implement the unmixing algorithm in this study, reduces the execution time of the entire process.
2. The temporal weights embedded in the algorithm, enabling the efficient handling of the changes occurred between the study dates. The available HSpaR images are classified separately and the temporal weights define the contribution of every land cover type to the pixel of the LSpaR image to be unmixed. The advantage in this case is twofold; firstly the unmixing algorithm is able to determine the state of the land cover type on the prediction date and secondly this is feasible with only one implementation. In previous approaches, the temporal weights are implemented in the unmixing outputs of the two base dates and subsequent estimations of reflectances are required on the prediction date [14].

The main challenge in the unmixing fusion process is the efficient handling of the available HSpaR images in order to estimate the radiometric values of the prediction date. By implementing the suggested approach, the majority of the changes within the study period, caused mainly by the differences in crop phenology, were estimated effectively. Time efficiency is another advantage of the proposed methodology, as a trial-and-error process is not required for producing a more accurate clustering result, and the fusion output does not need multiple algorithm implementations. However, the approach is



hindered by the salt-and-pepper effect of the classified images, which adds noise in the fusion output. Moreover, handling of the directional dependence of reflectance as a function of satellite viewing and solar geometry is also important in order to reduce noise in the time series generation.

### Acknowledgments

The authors would like to thank the European Space Agency (ESA) and the French Space Agency CNES (Centre National d'Etudes Spatiales) for providing the SPOT-4 (Take 5) time series datasets. Georgia Doxani was supported by ESA's Postdoctoral Research Fellowship programme (RFP).

### Author Contributions

Georgia Doxani was responsible for the design and the implementation of the whole process. She evaluated the experimental results and wrote the paper. Zina Mitrika significantly contributed to the design, application and evaluation of the proposed methodology. Ferran Gascon, Philippe Goryl and Bojan Bojkov monitored the experimental process and contributed to the validation and interpretation of results. They also contributed to the conceptual framework of this study and pointed out its importance in application to the Sentinels. All the authors assisted equally with the manuscript writing by making comments and corrections.

### Conflicts of Interest

The authors declare no conflict of interest.

### References

1. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; *et al.* Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* **2012**, *120*, 25–36.
2. Zhukov, B.; Oertel, D.; Lanzl, F.; Reinhäkel, G. Unmixing-based multi sensor multi-resolution image fusion. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 1212–1226.
3. Minghelli-Roman, A.; Polidori, L.; Mathieu-Blanc, S.; Loubersac, L.; Cauneau, F. Spatial resolution improvement by merging MERIS-ETM images for coastal water monitoring. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 227–231.
4. Zurita-Milla, R.; Kaiser, G.; Clevers, J.G.P.W.; Schneider, W.; Schaepman, M.E. Downscaling time series of MERIS full resolution data to monitor vegetation seasonal dynamics. *Remote Sens. Environ.* **2009**, *9*, 1874–1885.
5. Zurita-Milla, R.; Gómez-Chova, L.; Guanter, L.; Clevers, J.G.P.W.; Camps-Valls, G. Multitemporal unmixing of medium-spatial-resolution satellite images: A case study using MERIS images for land-cover mapping. *IEEE Trans. Geosci. Remote Sens.* **2011**, *11*, 4308–4317.
6. Amorós-López, J.; Gómez-Chova, L.; Guanter, L.; Alonso, L.; Moreno, J.; Camps-Valls, G. Regularized multiresolution spatial unmixing for ENVISAT/MERIS and Landsat/TM image Fusion. *IEEE Geosci. Remote Sens. Lett.* **2011**, *5*, 844–848.

7. Amorós-López, J.; Gómez-Chova, L.; Alonso, L.; Guanter, L.; Zurita-Milla, R.; Moreno, J.; Camps-Valls, G. Multitemporal fusion of Landsat/TM and ENVISAT/MERIS for crop monitoring. *Int. J. Appl. Earth Observ. Geoinf.* **2013**, *23*, 132–141.
8. Gao, F.; Masek, J.; Schwaller, M.; Hall, F. On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2207–2218.
9. Hilker, T.; Wulder, M.A.; Coops, N.C.; Linke, J.; McDermid, G.; Masek, J.; Gao, F.; White, J.C. A new data fusion model for high spatial- and temporal-resolution mapping of forest based on Landsat and MODIS. *Remote Sens. Environ.* **2009**, *113*, 1613–1627.
10. Walker, J.J.; de Beurs, K.M.; Wynne, R.H.; Gao, F. Evaluation of Landsat and MODIS data fusion products for analysis of dryland forest phenology. *Remote Sens. Environ.* **2012**, *117*, 381–393.
11. Jarihani, A.A.; McVicar, T.R.; Van Niel, T.G.; Emelyanova, I.V.; Callow, J.N.; Johansen, K. Blending Landsat and MODIS data to generate multispectral indices: A comparison of “Index-then-Blend” and “Blend-then-Index” approaches. *Remote Sens.* **2014**, *6*, 9213–9238.
12. Zhu, X.; Chen, J.; Gao, F.; Chen, X.; Masek, J.G. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sens. Environ.* **2010**, *114*, 2610–2623.
13. Hilker, T.; Wulder, M.A.; Coops, N.C.; Sritz, N.; White, J.C.; Gao, F.; Masek, J.G.; Stenhouse, G. Generation of dense time series synthetic Landsat data through data blending with MODIS using a spatial and temporal adaptive reflectance fusion model. *Remote Sens. Environ.* **2009**, *113*, 1988–1999.
14. Zhang, W.; Li, A.; Jin, H.; Bian, J.; Zhang, Z.; Lei, G.; Huang, C. An enhanced spatial and temporal data fusion model for fusing Landsat and MODIS surface reflectance to generate high temporal Landsat-like data. *Remote Sens.* **2013**, *5*, 5346–5368.
15. Gath, I.; Geva, A.B. Unsupervised optimal fuzzy clustering. *IEEE Trans. Pattern Anal. Machine Intell.* **1989**, *7*, 773–781.
16. Tso, B.; Mather, P.M. *Classification Methods for Remotely Sensed Data*, 2nd ed.; Taylor and Francis: New York, NY, USA, 2009.
17. Bensaid, A.M.; Hall, L.O.; Bezdek, J.C.; Clarke, L.P.; Silbiger, M.L.; Arrington, J.A.; Murtagh, R.F. Validity-guided (Re)clustering with applications to image segmentation. *IEEE Trans. Fuzzy Syst.* **1996**, *4*, 112–123.
18. Kim, D.W.; Lee, K.H.; Lee, D. On cluster validity index for estimation of the optimal number of fuzzy clusters. *Pattern Recognit.* **2004**, *37*, 2009–2025.
19. Balasko, B.; Abonyi, J.; Feil, B. Fuzzy Clustering and Data Analysis Toolbox (for Use with Matlab). Available Online: <http://www.abonyilab.com/software-and-data/fclusttoolbox/> (accessed on 23 July 2013)
20. Wang, W.; Zhang, Y. On fuzzy cluster validity indices. *Fuzzy Sets Syst.* **2007**, *19*, 2095–2117.
21. Hagolle, O.; Huc, M.; Kadiri, M.; Inglada, J.; Marais-Sicre, C.; Dedieu, G.; Sylvander, S.; Leroy, M.; Meygret, A.; Lonjou, V.; *et al.* SPOT4 (Take 5) time series over 45 sites to get ready for Sentinel-2. Presentation of lessons learned from SPOT4 -Take5 at S2 Symposium, Rome, Italy, 20–22 May 2014.

22. Wald, L. *Data Fusion, Definition and Architectures: Fusion of Image of Different Spatial Resolutions*; Presses des MINES: Paris, France, 2002.
23. Alparone, L.; Baronti, S.; Garzelli, A.; Nencini, F. A global quality measurement of Pan-sharpened multispectral imagery. *IEEE Geosci. Remote Sens. Lett.* **2004**, *4*, 313–317.
24. Inglada, J. Automatic land-cover map production of agricultural areas using supervised classification of SPOT4 (Take5) and Landsat-8 image time series. Presentation from SPOT4 -Take5 users day, CNES Toulouse, France, 18–19 November 2014.
25. Khavarian, H; Atkinson, P.M.; Milton, E.J.; Stewart J.B. Validation of the MODIS reflectance product under UK products. *Int. J. Remote Sens.* **2013**, *34*, 7376–7399.
26. Gevaert, C.M.; García-Haro, F.J. A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion. *Remote Sens. Environ.* **2015**, *156*, 34–44.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).