

# A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion



Caroline M. Gevaert<sup>a,b,\*</sup>, F. Javier García-Haro<sup>a</sup>

<sup>a</sup> *Departament de Termodinàmica, Facultat de Física, Universitat de València, Dr. Moliner 50, 46100, Burjassot, Valencia, Spain*

<sup>b</sup> *Department of Earth Observation Science, Faculty ITC, University of Twente, P.O. Box 6, Enschede 7500AA, Enschede, The Netherlands*

## ARTICLE INFO

### Article history:

Received 24 February 2014

Received in revised form 4 September 2014

Accepted 7 September 2014

Available online xxxx

### Keywords:

Data fusion

NDVI temporal profiles

STRUM

STARFM

Unmixing-based data fusion

## ABSTRACT

The focus of the current study is to compare data fusion methods applied to sensors with medium- and high-spatial resolutions. Two documented methods are applied, the spatial and temporal adaptive reflectance fusion model (STARFM) and an unmixing-based method which proposes a Bayesian formulation to incorporate prior spectral information. Furthermore, the strengths of both algorithms are combined in a novel data fusion method: the Spatial and Temporal Reflectance Unmixing Model (STRUM). The potential of each method is demonstrated using simulation imagery and Landsat and MODIS imagery. The theoretical basis of the algorithms causes STARFM and STRUM to produce Landsat-like reflectances while preserving the spatial patterns found in Landsat images, and the unmixing-based method to produce MODIS-like reflectances. The ability of fused images to capture phenological variations is also assessed using temporal NDVI profiles. Temporal profiles of STARFM NDVI closely resembled Landsat NDVI profiles. However, the unmixing-based method and STRUM produced a more accurate reconstruction of the NDVI trajectory in experiments simulating situations where few input high-resolution images are available. STRUM had the best performance as it produced surface reflectances which had the highest correlations to reference Landsat images. The results of this study indicate that STRUM is more suitable for data fusion applications requiring Landsat-like surface reflectances, such as gap-filling and cloud masking, especially in situations where few high-resolution images are available. Unmixing-based data fusion is recommended in situations which downscale the spectral characteristics of the medium-resolution input imagery and the STARFM method is recommended for constructing temporal profiles in applications containing many input high-resolution images.

© 2014 Elsevier Inc. All rights reserved.

## 1. Introduction

In satellite design, a trade-off must be made between spatial, spectral, and temporal resolutions. Therefore, many satellites with a high spatial resolution (hereafter referred to as “high-resolution” imagery) have restricted temporal and spectral resolutions when compared to satellites of a medium spatial resolution (hereafter, “medium-resolution”) (Emelyanova, McVicar, Van Niel, Tao Li, & Van Dijk, 2013). The Landsat mission provides one of the most extensively used high-resolution data sets due to its economic accessibility, 40-year continuous historical record, and technical specifications (Middleton et al., 2013). Landsat TM, ETM+ and OLI sensors have a spatial resolution of 30 m for the multi-spectral bands, which is adequate for many environmental applications.

However, Landsat satellites have a revisit time of 16 days (Gao, Masek, Schwaller, & Hall, 2006), and an average of 35% of the images are plagued by cloud cover (Roy et al., 2008). On the other hand, medium-resolution MODIS TERRA imagery has a daily revisit period, but a lower spatial resolution which limits its effectiveness for fine-scale environmental applications.

By applying multi-sensor data fusion, or spatial downscaling, two data sets are combined to create a result which exceeds the physical limitations of the individual input data sets (Lunetta, Lyon, Guindon, & Elvidge, 1998), and contains more information than the original input images (Ehlers, 1991; Pohl & Van Genderen, 1998). For example, Landsat and MODIS imagery can be fused to create a data set with a 30 m spatial resolution and a daily revisit period. Previous studies have applied multi-sensor data fusion between medium- and high-resolution imagery for applications such as phenology analysis (Bhandari, Phinn, & Gill, 2012; Feng et al., 2013; Hwang, Song, Bolstad, & Band, 2011; Walker, de Beurs, Wynne, & Gao, 2012), forest disturbance mapping (Arai, Shimabukuro, Pereira, & Vijaykumar, 2011; Hilker et al., 2009; Xin, Olofsson, Zhu, Tan, & Woodcock, 2013), the

\* Corresponding author at: Department of Earth Observation Science, Faculty ITC, University of Twente, P.O. Box 6, Enschede 7500AA – Enschede, The Netherlands. Tel.: +31 53 487 44 44.

E-mail address: [c.m.gevaert@utwente.nl](mailto:c.m.gevaert@utwente.nl) (C.M. Gevaert).

estimation of biophysical parameters (Anderson et al., 2011; Gao, Anderson, Kustas, & Wang, 2012; Singh, 2011), and public health (Liu & Weng, 2012). Recent developments in four-dimensional (sample, line, time, and wavelength) spatial-temporal imagery may also benefit from data fusion algorithms (Mello et al., 2013; Villa, Chanusot, Benediktsson, Jutten, & Dambreville, 2013).

The spatial and temporal adaptive reflectance fusion model (STARFM) (Gao et al., 2006) is perhaps the most widely-used data fusion algorithms for Landsat and MODIS imagery (Emelyanova et al., 2013). It is one of the few data fusion methods which result in synthetic Landsat-like surface reflectances (Singh, 2011). This method is particularly useful for detecting gradual changes over large land areas, such as phenology studies (Gao et al., 2006; Hilker et al., 2009). A number of studies have suggested improvements for the STARFM algorithm (Hilker et al., 2009; Roy et al., 2008; Zhu, Chen, Gao, Chen, & Masek, 2010).

A second set of data fusion algorithms is based on unmixing techniques. Traditional spectral unmixing methods rely on the linear spectral mixture model to extract end members and abundances on a sub-pixel scale (Bioucas-Dias et al., 2012). In the case of unmixing-based data fusion, the number of endmembers and abundances is obtained from the high-resolution data set, and the spectral signature of the endmembers is unmixed from the medium-resolution data set. This method has previously been applied to Landsat and Medium Resolution Imaging Spectrometer (MERIS) data (Amorós-López et al., 2013; Zurita-Milla, Kaiser, Clevers, Schneider, & Schaepman, 2009; Zurita-Milla, Gómez-Chova, Guanter, Clevers, & Camps-Valls, 2011; Zurita-Milla et al., 2009). Zurita-Milla et al. (2009) restricted the unmixing process to positive spectral reflectances below a defined upper limit. However, as this process was executed separately for each spectral band it may result in unrealistic endmember spectra. Amorós-López et al. (2013) utilized a regularization term in the cost function to restrict the variance of the endmember spectra from a pre-defined spectrum per class. This paper addresses these issues by presenting an alternative Bayesian unmixing method. Including the Bayesian theorem into fusing processes has been previously developed to optimally weight the spectral and panchromatic information in image-sharpening techniques (Fasbender, Radoux, & Bogaert, 2008). Our proposed Bayesian method provides a fast and easy to implement solution, which describes data fusion uncertainties in a clear probabilistic framework.

The main advantage of unmixing-based methods is that, unlike STARFM-based methods, they do not require high-resolution and medium-resolution data to have corresponding spectral bands. This allows for two additional possibilities. Firstly, unmixing-based data fusion can be used to downscale extra spectral bands and/or biophysical parameters to increase the spectral resolution of the high-resolution data sets. Secondly, auxiliary data sets such land cover may supplement or replace high-resolution imagery in the grouping of spectrally similar pixels into clusters (Zurita-Milla et al., 2011).

The current paper makes a comparative assessment of the STARFM and unmixing-based data fusion methods. Although the unmixing algorithm is similar to previous works (Amorós-López et al., 2013; Zurita-Milla, Clevers, & Schaepman, 2008), we propose an efficient Bayesian solution to the unmixing problem which produces realistic estimated spectra. Furthermore, we propose a new data fusion method, the Spatial and Temporal Reflectance Unmixing Model (STRUM), which combines the main features of STARFM (i.e. maintaining the spatial fine details and spectral characteristics of Landsat) and unmixing-based fusion methods (i.e. incorporating the temporal signatures of MODIS) to obtain temporally stable synthetic imagery at Landsat spatial resolution. The accuracy of each method is quantitatively assessed using simulated imagery as well as Landsat and MODIS reflectance products. In particular, the potential to transfer temporal information from MODIS to Landsat in order to capture phenological variations (i.e. NDVI profiles) is demonstrated.

## 2. Algorithm theory

### 2.1. STARFM

The STARFM algorithm described by Gao et al. (2006) was applied in the current study. The algorithm is based on the premise that both Landsat and MODIS imagery observe the same reflectance, biased by a constant error. This error depends on the characteristics of a pixel, and is systematic over short temporal intervals. Therefore, if a base Landsat–MODIS image pair is available on the same date, this error can be calculated for each pixel in the image. These errors can then be applied to the MODIS imagery of a prediction date to obtain a Landsat-like prediction image of that date.

This is done by following four steps. Firstly, MODIS data are reprojected and resampled to the Landsat imagery. Secondly, a moving window is applied to the Landsat imagery to identify similar neighboring pixels. Thirdly, a weight  $W_{ijk}$  is assigned to each similar neighbor based on: (i) the spectral difference between surface reflectances of the base Landsat–MODIS image pair, (ii) the temporal difference of the pixel's value in both MODIS images, and (iii) the spatial Euclidean distance between the neighbor and the central pixel.

The fourth and final step consists of calculating the surface reflectance of the central pixel. The algorithm is characterized in Eq. (1), where  $L(x_{\omega/2}, y_{\omega/2}, t_k)$  represents the central pixel of the moving window for the Landsat image prediction,  $M(x_{\omega/2}, y_{\omega/2}, t_k)$  is the value of the MODIS pixel on the prediction date, and  $L(x_{\omega/2}, y_{\omega/2}, t_0)$  and  $M(x_{\omega/2}, y_{\omega/2}, t_0)$  are the values of the Landsat and MODIS pixels of the base pair images:

$$L(x_{\omega/2}, y_{\omega/2}, t_k) = \sum_{i=1}^{\omega} \sum_{j=1}^{\omega} \sum_{k=1}^n W_{ijk} * (M(x_i, y_j, t_k) + L(x_i, y_j, t_0) - M(x_i, y_j, t_0)) \quad (1)$$

For a more detailed description of the STARFM algorithm, we refer to Gao et al. (2006).

### 2.2. Unmixing-based data fusion

An unmixing-based data fusion method was developed, primarily based on previous works (Amorós-López et al., 2013; Zurita-Milla et al., 2008), but introducing some modifications. Unmixing-based image fusion applies four steps to solve the linear mixing model: (1) clustering the high-resolution data set to define the end members, (2) calculating the fractions, or abundances, of each endmember within each medium-resolution pixel, (3) unmixing the medium-resolution pixel, and (4) assigning reflectance spectra to the high-resolution pixels (Fig. 1).

In Step 1, the present study applies the  $k$ -means algorithm to identify  $k$  spectral clusters. Other studies have used different clustering algorithms such as ISODATA (Zurita-Milla et al., 2008) and fuzzy  $k$ -means (Amorós-López et al., 2013). Auxiliary information, such as temporal variance or land cover data, may supplement or replace the high-resolution image in this step.

The second step applies a sliding window of  $[n \times n]$  MODIS pixels to the clustered image to record the endmember fraction matrix  $A$ , a  $[n^2 \times k]$  matrix with  $n^2$  rows (one for each pixel within the neighborhood) and  $k$  columns (one for each endmember). The use of a sliding window allows for spectral differences between pixels of the same cluster in different locations.

In Step 3, the unmixing is performed by solving the linear mixing model (Eq. 2), which is done separately for each MODIS band ( $\lambda$ ). The aim is to solve for  $\mathbf{e}^{(\lambda)}$  a  $[k \times 1]$  column vector that contains the reflectance of each of the  $k$  endmembers.  $\mathbf{r}^{(\lambda)}$  is a  $[n^2 \times 1]$  column vector containing the surface reflectance of each MODIS pixel in the  $n \times n$  moving window for the MODIS band  $\lambda$  which is currently being unmixed.

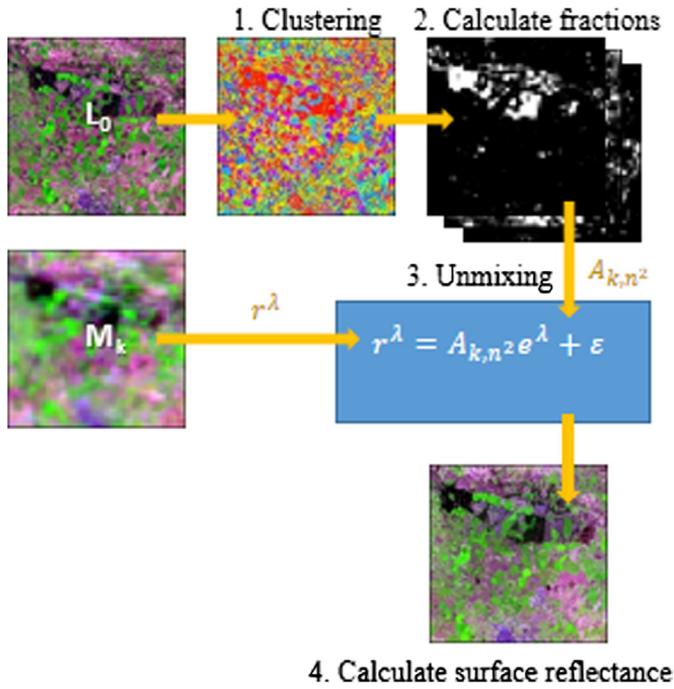


Fig. 1. Theoretical workflow of the unmixing-based data fusion algorithm.

This is achieved by minimizing the residuals ( $\varepsilon$ ) of the linear model (Eq. 2).

$$\mathbf{r}^{(\lambda)} = A \mathbf{e}^{(\lambda)} + \varepsilon \quad (2)$$

Two considerations are particularly important in the unmixing step: the relation between the number of clusters and the moving window size and restricting the unmixing phase. Regarding the former, spectral unmixing is very sensitive to co-linearity problems in which high correlations between endmembers lead to inversion of ill-posed matrices (García-Haro, Sommer, & Kemper, 2005). Though it is mathematically feasible to use as many classes ( $k$ ) as low resolution pixels in the neighborhood ( $n^2$ ), highly correlated data exert a significant influence on the estimates. Hence,  $n^2$  should be considerably larger than the number of classes  $k$ . The idea consists of restricting the set of possible classes to a reduced number ( $k'$ ) of appropriate classes that can be different for each sliding window, thereby allowing a more accurate decomposition. This is achieved by discarding classes with fractional abundance not exceeding a value of 0.10 in at least one of the MODIS pixels of the neighborhood.

The second consideration is that unmixing may produce unrealistic estimated spectra if the spectral shape of the endmembers is not imposed in any way. There are a number of strategies which address this problem. Zurita-Milla et al. (2008) proposed restraining the linear unmixing method by ensuring that the reflectance values must be positive and not exceeding an appropriate upper limit. Amorós-López et al. (2013) included a spectral regularization term in the form of a cost function which precludes that endmembers identified in the window differ considerably from pre-defined endmember spectra obtained from the input images.

In this paper, we propose an alternative solution, based on Bayesian theory, to constrain the estimation of the endmembers  $\mathbf{e}^{(\lambda)}$ . By assuming a prior probability of the values to be inferred  $p(\mathbf{e}^{(\lambda)})$  the Bayesian rule can be applied to combine data  $\mathbf{r}^{(\lambda)}$  and prior  $p(\mathbf{e}^{(\lambda)})$  information. In this study, we assumed a Gaussian distribution with mean  $\mathbf{e}_0$  and covariance matrix  $\Sigma_0$ , i.e.  $p(\mathbf{e}^{(\lambda)}) = \mathcal{N}(\mathbf{e}_0, \Sigma_0)$ . We also assumed that the data  $\mathbf{r}^{(\lambda)}$  are noisy observations also following a Gaussian distribution, with covariance matrix  $\Sigma_r$ . Under these assumptions, the posterior

distribution for the solution,  $p(\mathbf{e}^{(\lambda)} | \mathbf{r}^{(\lambda)})$ , follows a normal distribution with expectation value  $\mu_e$  and covariance matrix  $\Sigma_e$ , i.e.:

$$p(\mathbf{e}^{(\lambda)} | \mathbf{r}^{(\lambda)}, A) \approx \mathcal{N}(\mathbf{e} | \mu_e, \Sigma_e). \quad (3)$$

Applying the Bayes' theorem for conditional probabilities gives us the following expressions (Murphy, 2012):

$$\Sigma_e = [\Sigma_0^{-1} + A^T \Sigma_r^{-1} A]^{-1} \quad (4)$$

$$\mu_e = \Sigma_e [A^T \Sigma_r^{-1} \mathbf{r}^{(\lambda)} + \Sigma_0^{-1} \mathbf{e}_0]. \quad (5)$$

Note that  $\mu_e$  is the minimum mean square error estimator of the endmembers.  $\mu_e$  is expressed as the sum of two terms. The first term is linear in  $\mathbf{r}^{(\lambda)}$ , being the unconstrained estimator, whereas the second term is linear in  $\mathbf{e}_0$ . Thus the predicted value for the endmembers results from a compromise between the information coming from the measurement  $\mathbf{r}^{(\lambda)}$  and the prior knowledge  $\mathbf{e}_0$  that we have regarding the endmembers.

In this study, we used spherical covariance matrices for simplicity, defining  $\Sigma_0$  and  $\Sigma_r$  as follows:

$$\Sigma_0 = \sigma_0^2 I(k) \quad \Sigma_r = \sigma_r^2 I(n^2). \quad (6)$$

where  $I(k)$  and  $I(n^2)$  represent the  $[k \times k]$  and  $[n^2 \times n^2]$  identity matrices, respectively, and  $\sigma_0^2$  and  $\sigma_r^2$  present the variances assigned to the prior of  $\mathbf{e}^{(\lambda)}$  and the noisy data  $\mathbf{r}^{(\lambda)}$ . If the precision assigned to the prior is strong relative to the data strength ( $\sigma_0$  is large relative to  $\sigma_r$ ), more emphasis is put on the prior and vice versa. The relative importance of the prior is controlled in this study by optimizing the ratio  $\sigma_{\text{ratio}} = \sigma_r / \sigma_0$ . By reducing  $\sigma_{\text{ratio}}$ , more weight is put on the prior.

The prior end member means  $\mathbf{e}_0$  are selected among MODIS pixels with high abundance levels for each class, similar to the approach presented by Amorós-López et al. (2013). Selecting end members directly from the image may reduce biases caused by atmospheric conditions and radiometric corrections. For each class, the purity of the MODIS pixel presenting the highest end member abundance is retained. This allows us refine the method using a class specific precision, i.e. replacing the spherical covariance matrix (Eq. 6) by a diagonal covariance matrix in which each term is regulated by the purity of the class. The reliability of the prior is thus assumed to be small for classes with lower homogeneity.

The fourth and final step creates the fused image. The unmixed spectrum of the relevant endmember is assigned to each pixel at Landsat scale. This is done separately for each MODIS pixel through the use of a moving window, thus taking into account spectral variability within each cluster, or end member, throughout the scene.

### 2.3. STRUM

Steps 1 and 2 in STRUM are identical to the previous unmixing method but steps 3 and 4 are inspired by STARFM principles, assuming that MODIS and Landsat surface reflectance are radiometrically and temporally comparable. Similar to the STARFM, this method requires three input images: a Landsat and MODIS image on the same base date ( $t_0$ ), and a MODIS image on the prediction date ( $t_k$ ). A residual image is defined as the difference between the two MODIS images, i.e. residual =  $r(t_k) - r(t_0)$ . This is utilized in step 3 as the input medium-resolution data set and unmixed using the same Bayesian approach described above.

$$\text{residual}^{(\lambda)} = A \mathbf{e}^{(\lambda)} + \varepsilon \quad (7)$$

The aim is to solve for  $e^{(\lambda)}$ , a column vector that contains temporal changes of each endmember spectra in MODIS band ( $\lambda$ ). The definition of this residual implies that STRUM requires corresponding spectral bands between the high- and medium-resolution data, as opposed to the unmixing method which can be applied to all the spectral bands of the medium-resolution data.

Step 4 creates the fused image. The temporal change of the relevant endmember is assigned to each Landsat pixel in the window to its class label, and added to the input Landsat image on the base date:

$$L(t_k) = L(t_0) + e^{(\lambda)} \quad (8)$$

The fused results therefore provide Landsat-like reflectances containing information regarding temporal variation in surface reflectances obtained from the input MODIS imagery.

### 3. Materials and methods

#### 3.1. Study area

The study area is in the Barrax region near Albacete, Spain (39.9° N, 2.5° W). It is a square area of 576 km<sup>2</sup>, which corresponds to 800 × 800 Landsat pixels. Typical elevations lie around 700 m, and it has a Mediterranean climate. It is located in one of the driest regions of Europe, with mean annual rainfall below 400 mm (Su et al., 2008). Approximately 90% of the study area is used for agricultural purposes, 7% consists of forests and semi-natural areas, and the remaining 3% are artificial areas such as urban fabric. Crops in the region consist of cereals, alfalfa, sunflower, potato, sugar beet and vegetables (Amorós-López et al., 2013).

#### 3.2. Experiment with Simulation Imagery

The three data fusion methods were first applied to simulated imagery in order to eliminate the interference of confounding factors such as radiometric and geometric inconsistencies between sensors. To create the simulated imagery, a number of polygons were digitized over the input Landsat imagery to represent agricultural fields. Eight spectral classes were defined and assigned realistic spectral values corresponding to eight land use classes (bare soil, asphalt, water, forest and four crops in various stages of development), which were obtained from the Landsat imagery. The polygons were transformed to a 6-band raster representing the high-resolution image. This image was then resampled at a ratio of 15:1 to create the medium-resolution image, thus obtaining a spatial ratio similar to that of MODIS (463 m) to Landsat (30 m) imagery. A total of four sample images were created, two MODIS-like images and two Landsat-like images for two different dates:  $t_0$  and  $t_k$ .

The base image pair on day  $t_0$  was combined with the medium-resolution image on prediction date  $t_k$  to predict the high-resolution image on the prediction date. This simulates the application of data fusion in a situation with temporal change. The results of all three methods were compared to the simulated high-resolution images on  $t_k$  by calculating the RMSE, bias ( $\rho_{\text{Landsat}} - \rho_{\text{fused}}$ ) and Pearson's correlation coefficient ( $r$ ) between the prediction and high-resolution images.

#### 3.3. Data fusion applications using satellite imagery

##### 3.3.1. Satellite image pre-processing

The current study uses Landsat 8/OLI imagery for the high-resolution input data. The Landsat 8 images were downloaded as a Level 1T product and corrected to TOA reflectance using the parameters provided in the metadata file. The images were atmospherically corrected using the dark object subtraction (DOS) method (Chavez, 1988). As the study area lies in two Landsat paths, images are available every 8 days rather than the usual interval of 16. However, only six

corresponding cloud-free Landsat 8/OLI images were available within the time frame: April 14, May 23, June 1, June 24, July 3, and July 19, 2013.

The moderate resolution imaging spectroradiometer (MODIS) sensor, on board satellites Terra and Aqua, is used for medium-resolution imagery. The MODIS MCD43A4 Nadir BRDF-Adjusted Reflectance product with a 463 m resolution was utilized, as previous studies have indicated that MODIS BRDF products provide more accurate data fusion results than daily MODIS surface reflectances (Roy et al., 2008; Walker et al., 2012). This product is a 16-day composite of both Aqua and Terra satellites, produced every 8 days. Seventeen consecutive MCD43A4 images between March 30, 2013 and August 5, 2013 were used.

All images were subsetted to the study area, low-quality pixels identified by the quality flag information were removed, and the images were reprojected to UTM coordinates. Furthermore, the geometric co-registration of each Landsat–MODIS image pair was optimized. This was done by determining the optimal offset which maximized the Pearson's correlation coefficient between all spectral bands of the MODIS image and resampled Landsat image.

Inspection of the data revealed significant differences between the Landsat and MODIS reflectance data sets, caused by differences in image processing chains and spectral band difference effects (Teillet, Fedosejevs, Thome, & Barker, 2007). In order to reduce this bias, a radiometric normalization method was applied, which assumes a linear relationship between the reflectance of both images. The critical aspect is the determination of suitable invariant pixels (at MODIS scale) upon which to base the normalization. We applied iteratively reweighted multivariate alteration detection (IRMAD) transformation (Nielsen, 2007) to select pixels with a high no-change probability (i.e., >0.95). The method is completely automatic and compares favorably with normalization using hand-selected invariant features (Canty & Nielsen, 2008). We regressed the (resampled) Landsat images onto the corresponding MODIS images at no-change locations. For each band, slope and intercept parameters were obtained using orthogonal linear regression, since it outperformed the ordinary least squares regression.

##### 3.3.2. Data fusion optimization

Sensitivity to input parameters, distribution of prediction errors and the quality of the results were analyzed for all three fusion methods. Firstly, optimal input parameters were identified for each algorithm using the Landsat 8/OLI image on July 3, 2013 and the MODIS composite from June 26 to July 12th as base images, and the MODIS composite from July 12 to 28th to represent the prediction date. For the unmixing-based method, the number of clusters was varied using the values  $k = 10, 20, 40,$  and  $80$ . The window size was also varied from  $\omega = 5$  to 41 MODIS pixels in steps of 4. The weight of the *a priori* endmember information was varied, using the values  $\sigma_{\text{ratio}} = 0.01, 1, 2,$  and  $5$ . The input parameters of STRUM are equal to those of the unmixing-based algorithm and were therefore not calculated separately. Regarding the STARFM algorithm, there were two main parameters of interest. The search distance refers to the maximum distance between the pixel being predicted and the similar neighbor. This was assigned values of 250 m, 500 m, 750 m, 1000 m, 1500 m, 2000 m, and 3000 m. Secondly, the number of spectral slices influences the initial selection of similar neighbors within the moving window. The number of spectral slices refers to the number of land cover classes expected to be present in the image (Zhu et al., 2010). Twice the standard deviation divided by the number of spectral slices determines the maximum spectral difference between the central pixel and the potentially similar neighbor. Thus, a higher number of spectral slices imply a stricter selection. In this study, 10, 20 40 and 80 spectral slices were utilized.

The results were compared to the actual Landsat image on July 19, 2013 using the RMSE, Pearson's correlation coefficient, and the Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS). The ERGAS assumes that the fused result must be spectrally similar to the medium-

resolution image (Wald, 2002) and spatially similar to the high-resolution image (Lillo-Saavedra, Gonzalo, Arquero, & Martinez, 2005), and is calculated as follows:

$$\text{ERGAS}_{\text{spectral}} = 100 \frac{h}{l} \sqrt{\frac{1}{N_{\text{ban}}} \sum_{i=1}^{N_{\text{ban}}} (\text{RMSE}_i / M_i)^2} \quad (9)$$

$$\text{ERGAS}_{\text{spatial}} = 100 * \frac{h}{l} * \frac{\text{RMSE}}{M} \quad (10)$$

where  $h$  is the spatial resolution of the high-resolution image,  $l$  is the spatial resolution of the low-resolution image,  $N_{\text{ban}}$  is the number of spectral bands,  $M$  is the mean spectral value of all bands, and  $M_i$  is the mean spectral value of band  $i$ . Lower ERGAS values indicate a higher fusion quality.

After defining the optimal input parameters, the quality of the optimal fused images was further analyzed to identify possible dependencies on spectral region and spatial distribution. The spatial distribution was analyzed by assessing the linear relationship between the RMSE of fused reflectances (referred to the Landsat observations) and the normalized difference vegetation index (NDVI) of the Landsat observations.

### 3.3.3. Temporal profile analyses

In order to assess whether the fused images are suitable for studying vegetation dynamics, the final experiment consisted of applying both algorithms to create NDVI temporal profiles of 17 fused images at 8-day intervals. The NDVI is a commonly-used index to estimate vegetation greenness which reflects both density and condition. Although it has known limitations (Jiang et al., 2006) and NDVI values show discrepancies between sensors (Steven, Malthus, Baret, Xu, & Chopping, 2003; Van Leeuwen, Orr, Marsh, & Herrmann, 2006), it has often been used for the development of temporal profiles for vegetation analysis (Gu, Li, Huang, & Okin, 2009; Hmimina et al., 2013) and agricultural monitoring (Lunetta, Shao, Ediriwickrema, & Lyon, 2010).

The NDVI time series were assessed by examining seasonal variations over four representative agricultural fields. The aim of the experiment was to fill gaps in the Landsat 8 temporal profiles by using MODIS data with more frequent coverage. After evaluating an ideal scenario in which many Landsat images were available, the same tests were carried out using only one or two Landsat 8 images to create the time series. This simulates realistic situations in which few high-resolution images are available. Furthermore, it allows unused Landsat images to act as a testing data set to analyze the quality of the fused NDVI predictions. Finally, NDVI predictions of the first series including all input Landsat images were resampled to the MODIS resolution and compared with the NDVI profiles of the original MODIS imagery over three homogeneous (at MODIS resolution) sites.

## 4. Results

### 4.1. Experiment with simulated imagery

A visual comparison of the simulation results indicates that all three methods accurately predict the corresponding Landsat-like image (Fig. 2). Both colors and fine spatial details are correctly captured in the fused images. Errors are higher when no homogeneous medium-resolution pixels are available within the search distance. This is especially visible in the STARFM method (Fig. 2h), where the narrow beige areas between the larger agricultural fields appear slightly reddish. The STARFM method also produced errors around the borders of objects which had similar reflectance values in the base date, but follow differing temporal patterns. For example, the two circular fields in the upper right corner of the simulation imagery were both vegetated on the base date, but the right field had been “harvested” before the second date. In

the left border of the harvested field, some pixels are bright green rather than taupe (Fig. 2c). This is due to the selection of similar neighbors from the field to the left (which is still vegetated). Similar errors are visible in pixels of the irregular field in the bottom left of the image which border the vegetated field. The unmixing-based method and STRUM do not display such errors.

Quantitative indicators indicate STARFM provided the most accurate predictions in the simulation imagery tests (Table 1). It closely replicated the actual simulation image, as the correlation, RMSE and ERGAS are almost perfect. STARFM method performed better than the unmixing-based method, although the performance of the latter was also very good.

### 4.2. Optimization of input parameters

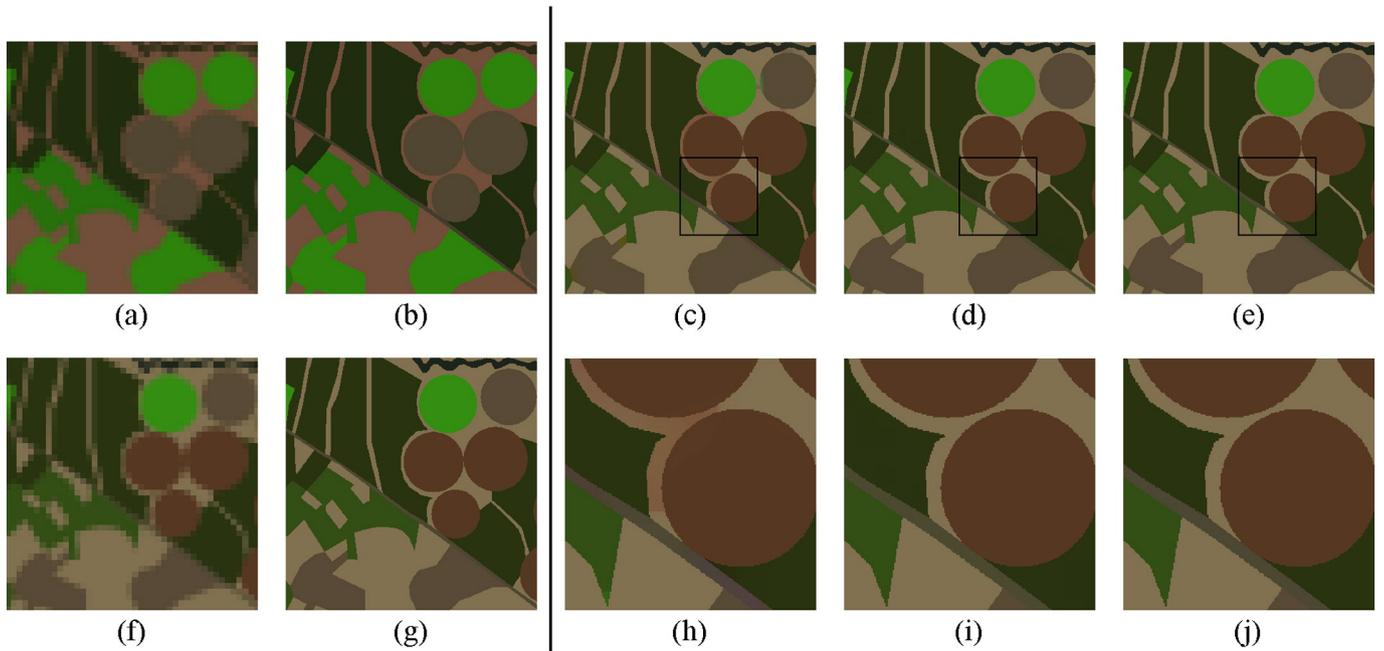
#### 4.2.1. STARFM

Both search distance ( $\omega$ ) and the number of spectral slices ( $m$ ) influence which pixels are utilized to form the linear system of equations. Changing these parameters obtained slightly different STARFM products. The RMSE was lower for higher amounts of spectral slices. Increasing search distance was related to a decreasing RMSE, although for 80 spectral slices the RMSE becomes rather insensitive to the search distance. Larger search distances greatly increased computational costs, so the optimal parameters were defined as  $m = 80$  spectral slices and  $\omega = 750$  meters. Using these optimal parameters (Fig. 3) produced surface reflectances very similar to Landsat imagery, as indicated by Table 2.

Using optimal input parameters, the errors of each spectral band were calculated (Table 3). Most Landsat/MODIS band combinations provided similar errors, although the combinations with larger wavelengths (TM4/MODIS2, TM5/MODIS6, and TM7/MODIS7) provided higher errors. Regarding spatial distribution of errors, no significant correlation was found between prediction errors and NDVI of the Landsat image ( $p < 0.001$ ), which reveals that vegetation density has no influence on the accuracy of the STARFM predictions within the study area.

#### 4.2.2. Unmixing-based data fusion

Fig. 4 presents the relation between variations of the input parameters on the data fusion quality (represented by ERGAS) of the unmixing-based method. The results indicate that high numbers of clusters combined with small moving-window sizes lead to poor-quality fused results. The accuracy was higher for smaller numbers of clusters (lower ERGAS and higher correlation coefficient). Few clusters (i.e.  $k = 10$  or 20) produced better indicator values, and reduced the sensitivity to variations in moving-window size. In order to preserve the spectral variability while maintaining fusion quality, a small value of neighborhood size (i.e. less than 20) was preferred. Furthermore, it was observed that the inclusion of *a priori* spectral information allows to significantly improve the results. For example, a significant reduction of  $\text{ERGAS}_{\text{spectral}}$  is observed when assigning equal variance to prior information and measurements ( $\sigma_{\text{ratio}} = 1$ ) compared with a nearly unconstrained estimate and uninformative prior ( $\sigma_{\text{ratio}} = 0.01$ ). Taking into account the ERGAS values in Fig. 4 and a visual inspection of the fused images, the optimal parameters in the current study were defined as  $\omega = 9$ ,  $k = 20$  and  $\sigma_{\text{ratio}} = 1.0$ , although other combinations are also possible. These input parameters produced a fused product which contained spectral information more accurately portraying the MODIS reflectances than Landsat, as the correlation coefficient, RMSE, and ERGAS indicate (Table 2). This is reflected in Fig. 3, which illustrates that the STARFM output is spectrally very similar to the Landsat image whereas the unmixed-based output retains mainly the spectral characteristics of the MODIS image. Compared with the unmixing-based method, the STARFM and STRUM algorithms accurately preserve most of the fine spatial detail in the Landsat image. The computation time per fused image in Fig. 3 was 68.9 s (STARFM), 15.8 s (STRUM), and 15.7 s (unmixing-based method) on a computer with an Intel® Core™ i7-2670QM processor and 8.00 GB RAM.



**Fig. 2.** The simulated medium-resolution imagery for base date  $t_0$  (a) and prediction date  $t_k$  (f), the simulated high-resolution imagery for  $t_0$  (b) and  $t_k$  (g), the STARFM (c), unmixing-based (d), and STRUM (e) results of the simulation tests. (h), (i), and (j) provide a closer look to a subset of the results of STARFM, unmixing-based fusion and STRUM respectively. All images are displayed with the band combination R:SWIR, G:NIR, and B:Red.

STRUM most accurately predicted the actual Landsat 8 observations (Table 2). The correlation to MODIS imagery was slightly lower than the original unmixing-based method, although it was still higher than that of the STARFM. Visually, this method also produces a fused data set which is highly comparable to the reference Landsat image (Fig. 3).

Table 3 indicates that the quality of the fused product for each spectral band was satisfactory. Each band had a higher correlation and lower RMSE for the MODIS imagery than the Landsat imagery. The TM4/MODIS2 combination had the lowest correlations and highest RMSE to both the Landsat and MODIS imagery. The correlations of the other spectral bands were similar, although the RMSE increased with the wavelength. The spectral distribution of STRUM prediction errors followed a similar pattern.

#### 4.3. Derivation of temporal profiles

The first scenario used all available input Landsat 8 images to create the NDVI profile. The results demonstrate that the STARFM algorithm precisely captured Landsat 8/OLI NDVI values (Fig. 5). The unmixing-based method and STRUM slightly overestimated Landsat NDVI at low vegetation densities, while slightly underestimating the Landsat NDVI at higher vegetation densities.

Temporal profiles were also derived using only one or two Landsat input images. The results of these scenarios illustrate that STARFM predictions are highly dependent on the number input image(s) (Fig. 6). STARFM always used the temporally closest image base pair for predictions. The point where the input image changed (between DOY = 161 and DOY = 169) is obvious in the STARFM temporal profiles. When using only one input image for the temporal profiles, STARFM failed to

capture phenological variations. However, the unmixing-based method and STRUM generally follow the phenological trends indicated by the Landsat 8 imagery, even when only one input Landsat image is used. Furthermore, the STARFM method produces negative NDVI values in some situations (e.g. Fig. 6e and f), which is likely due to the fact that the algorithm processes each spectral band separately.

The final temporal profile experiment consisted of comparing the resampled fused images to the MODIS NDVI over homogenous (at the MODIS resolution) sites. Fig. 7 indicates that all three methods captured relative phenological variations indicated by the MODIS images. The unmixing-based algorithm is the only algorithm that does not utilize Landsat reflectance to obtain predictions, which is likely why this algorithm is slightly better at reproducing MODIS NDVI observations than the other two methods.

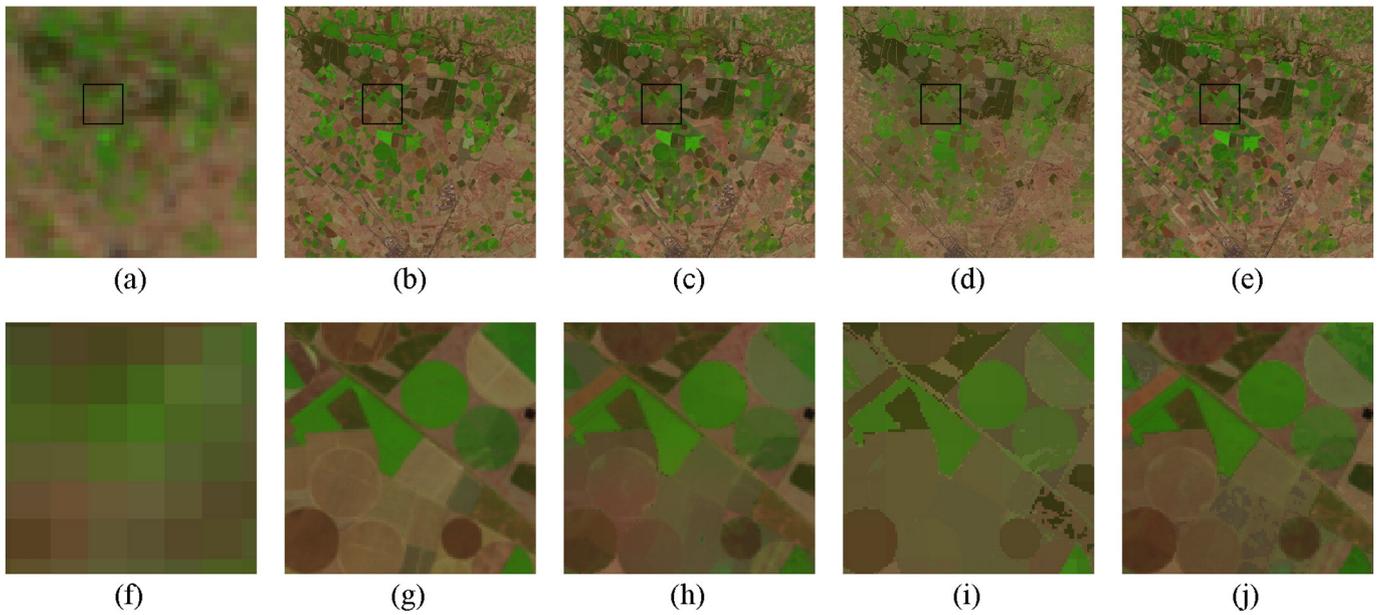
## 5. Discussion

### 5.1. Error distributions and optimal input parameters

The results presented in the current study indicate that all three fusion methods can successfully fuse MODIS and Landsat imagery. The low RMSE and biases obtained through the STARFM method are comparable to those obtained by other studies (Emelyanova et al., 2013; Gao et al., 2006). However, the inclusion of *a priori* spectral information in the current unmixing-based algorithm enhances the data fusion quality when compared to the ERGAS values obtained by similar studies. The current methodology obtained a spectral ERGAS of 0.52 and a spatial ERGAS of 1.13, whereas similar studies obtained 0.95 and 1.90 (Zurita-Milla et al., 2008), 0.58 and 2.10 (Zurita-Milla et al., 2009), and 1.35

**Table 1**  
Quality indicators of data fusion applied to simulation imagery.

Method	$r_{\text{Landsat}}$	$r_{\text{MODIS}}$	$\text{RMSE}_{\text{Landsat}}$	$\text{RMSE}_{\text{MODIS}}$	$\text{ERGAS}_{\text{spatial}}$	$\text{ERGAS}_{\text{spectral}}$
STARFM	0.998	0.999	0.009	0.007	0.293	0.224
Unmixing-based	0.980	0.979	0.016	0.015	0.668	0.624
STRUM	1.000	1.000	0.000	0.000	0.004	0.002



**Fig. 3.** Results of data fusion using July 3, 2013 as a base date to predict reflectance on July 19th. The top row presents MODIS composite of July 12–28, 2013 (a), the reference Landsat image on July 19th (b), and the predictions using STARFM (c), the unmixing-based algorithm (d), and STRUM (e). The second row displays a detailed subset of each image to enhance the differences. All images are displayed with the band combination R:SWIR, G:NIR, and B:Red.

and 2.45 (Amorós-López et al., 2013) respectively when applied to MERIS and Landsat imagery of various agricultural areas. It should be noted that the imagery utilized in the other studies is not identical to that of the current study, so comparison of ERGAS values is only indicative. STRUM obtained a slightly lower spectral ERGAS of 0.66 while improving the spatial ERGAS 1.02. One further advantage of the proposed unmixing methods is the use of a Bayesian framework, which provided clear uncertainty ranges and efficient computation.

Predictions in the NIR region (TM4/MODIS2) had the lowest correlation to both Landsat and MODIS images for all three methods. The RMSE, however, increased at longer wavelengths for all methods. Regarding the relation between error distribution and vegetation cover, the STARFM showed no significant distributions. However, both unmixing-based methods displayed a slightly positive trend between vegetation cover and prediction error in all regions except the NIR. The overestimation of red reflectance and underestimation of NIR causes the slight overestimation of low Landsat NDVI values and slight underestimation of high Landsat NDVI values which was visible in the

time profiles. This is likely due to the mixing of MODIS pixels which smooth the extremes values of Landsat reflectances, such as averaging vegetation densities over a larger area. Further investigations regarding the distribution of prediction errors, and if such errors are spatially and temporally consistent, could play a significant role in algorithm improvement. Another line of research could compare optimal data fusion input parameters identified in the current study with other studies to possibly identify a set of generally applicable input parameters. This could significantly speed up the application of data fusion to new areas.

## 5.2. Comparison of data fusion algorithms

The theoretical basis of the algorithms causes STARFM and STRUM to produce Landsat-like reflectances, and the unmixing-based method to produce MODIS-like reflectances. Such differences must be taken into account when comparing the results of each data fusion method. By applying the IRMAD radiometric normalization technique, biases between sensors were minimized. However, this does not eliminate the problem

**Table 2**

Global indicators of fused images produced using optimal input parameters.

Data fusion method	$r_{\text{Landsat}}$	$r_{\text{MODIS}}$	$\text{RMSE}_{\text{Landsat}}$	$\text{RMSE}_{\text{MODIS}}$	$\text{ERGAS}_{\text{spatial}}$	$\text{ERGAS}_{\text{spectral}}$
STARFM	0.953	0.975	0.044	0.029	1.155	0.685
Unmixing-based	0.949	0.989	0.045	0.021	1.133	0.521
STRUM	0.963	0.978	0.039	0.027	1.018	0.662

**Table 3**

Data fusion quality indicators using the optimum input parameters for each spectral band.

Spectral bands	STARFM				Unmixing-based				STRUM			
	Landsat		MODIS		Landsat		MODIS		Landsat		MODIS	
	$r$	RMSE	$r$	RMSE	$r$	RMSE	$r$	RMSE	$r$	RMSE	$r$	RMSE
TM1/MODIS3	0.910	0.013	0.945	0.008	0.836	0.016	0.957	0.007	0.925	0.012	0.936	0.009
TM2/MODIS4	0.906	0.021	0.936	0.014	0.858	0.025	0.958	0.012	0.921	0.019	0.940	0.013
TM3/MODIS1	0.916	0.033	0.942	0.022	0.882	0.038	0.959	0.018	0.936	0.029	0.938	0.022
TM4/MODIS2	0.829	0.041	0.899	0.023	0.764	0.046	0.921	0.021	0.817	0.043	0.876	0.026
TM5/MODIS6	0.830	0.067	0.877	0.044	0.829	0.062	0.947	0.028	0.872	0.057	0.892	0.040
TM7/MODIS7	0.870	0.062	0.899	0.042	0.857	0.061	0.950	0.029	0.907	0.052	0.907	0.039

completely and therefore this study utilized simulated imagery where MODIS-like reflectances are obtained directly from the Landsat-like reflectances to support the comparison of the three data fusion methods. The results of simulation imagery experiments suggest that STRUM provides the most accurate predictions, followed by STARFM and the unmixing-based algorithm. The simulation tests also highlight a number of causes of prediction errors. For example, similar neighbors are selected separately for each spectral band in the STARFM algorithm, which may cause the fused spectra to be unrealistic. This is supported by the negative NDVI values of low vegetation cover obtained by some time series (Fig. 6e and f).

The unmixing-based method identified similar neighboring pixels more accurately. This is visible in the simulation test, where the unmixing method identified slight differences in spectral reflectance which anticipated differing temporal tendencies. Furthermore, results of the temporal profiles supported this finding by revealing that unmixing-based predictions were less dependent on the number of input high-resolution images. STRUM maintained this good performance with low dependence on the availability of Landsat imagery, but furthermore provided surface reflectance data which accurately reflected actual Landsat observations. The computation time of the unmixing-based method and STRUM was also much lower.

The limitations of STARFM in identifying temporal variations in surface reflectances, as displayed in Fig. 6, are also due to the formulation of the STARFM algorithm. The weight for each spectrally similar neighbor is inversely proportional to the difference in reflectance measured in the MODIS image between the base and the prediction dates. Although this was intended to stabilize the algorithm, it also causes it to be biased towards pixels representing little temporal change. The STRUM method avoids this issue by using clustering methods to select similar pixels.

Gridding effects and the point spread function (PSF) cause the sensor observation in any grid cell to be only partially derived from the location of the cell (Gómez-chova et al., 2011; Tan et al., 2006). This may influence unmixing-based fusion, especially for pixels along the borders of differing land covers. In this paper, the use of Nadir BRDF-Adjusted

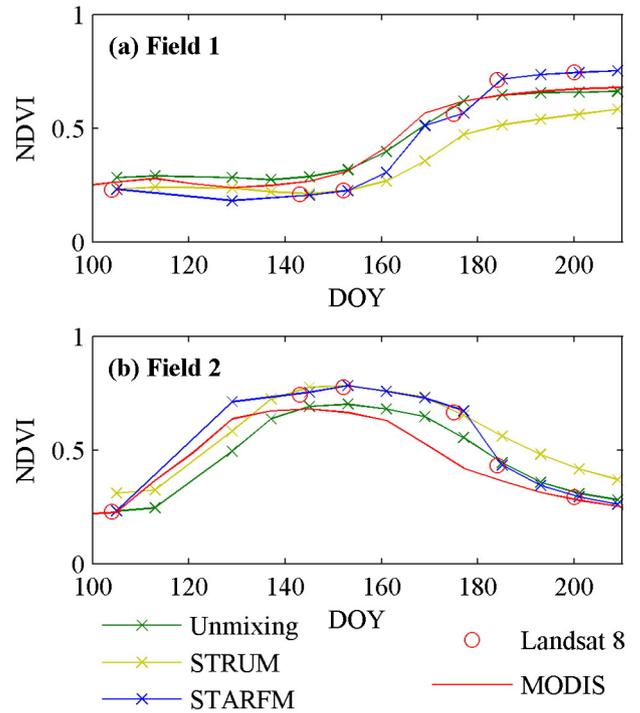


Fig. 5. Temporal profiles of the average NDVI of two agricultural fields resulting from the STARFM, unmixing-based fusion and STRUM using all available input Landsat images, compared to the NDVI calculated from actual Landsat and MODIS imagery (for interpretation of the references to color in this figure legend, the reader is referred to the web version of the article).

reflectance (MCD43A4) may lower these influences, which could be more significant when using daily MODIS images due to different image acquisition angles. The high quality of the unmixing-based results (Table 2) suggests that remaining gridding effects do not seriously

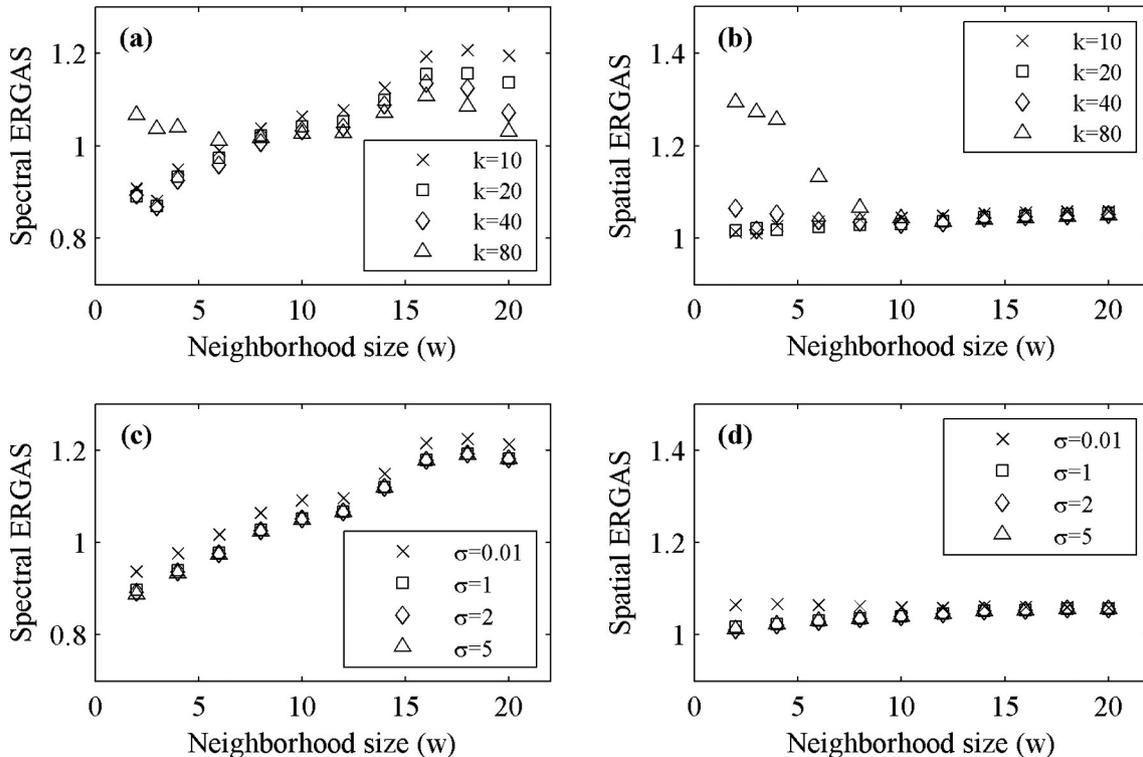
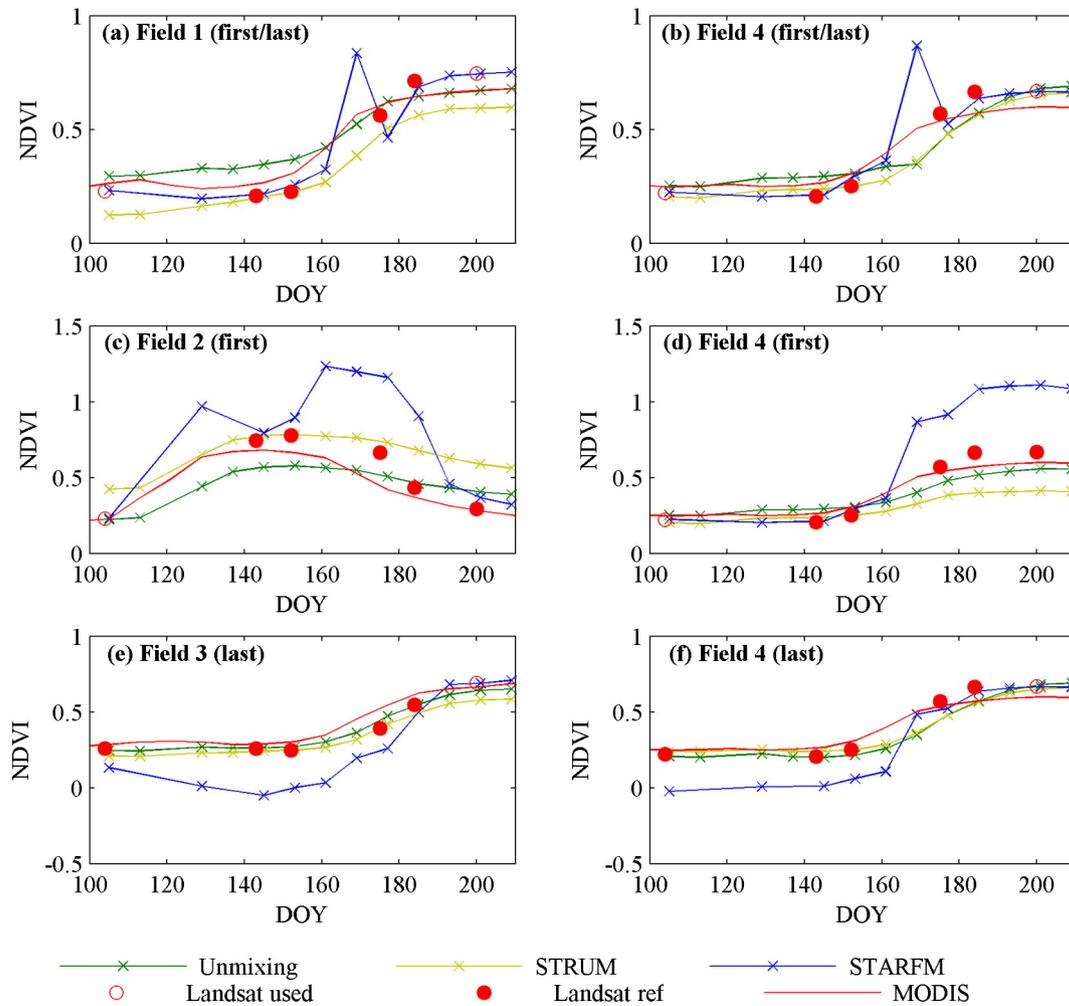


Fig. 4. Response of the spectral ERGAS (a) and spatial ERGAS (b) to changes in the moving window size ( $w$ ) and the number of clusters ( $k$ ) for a fixed  $\sigma_{ratio} = 0.01$ , and the response of spectral ERGAS (c) and spatial ERGAS (d) to changes in the moving window size ( $w$ ) and a priori weighting ( $\sigma_{ratio}$ ) for a fixed  $k = 20$ .



**Fig. 6.** A selection of temporal profiles of STARFM, unmixing-based, and STRUM NDVI predictions for: scenarios using both the first and last Landsat image to create the temporal profiles (a) and (b), and scenarios using only the first (c), (d) or last (e), (f) Landsat image to construct the time series (for interpretation of the references to color in this figure legend, the reader is referred to the web version of the article).

deteriorate fusion quality, although the extent to which PSF plays a role, and whether the improved accuracy outweighs the additional processing costs, should be subject to further study.

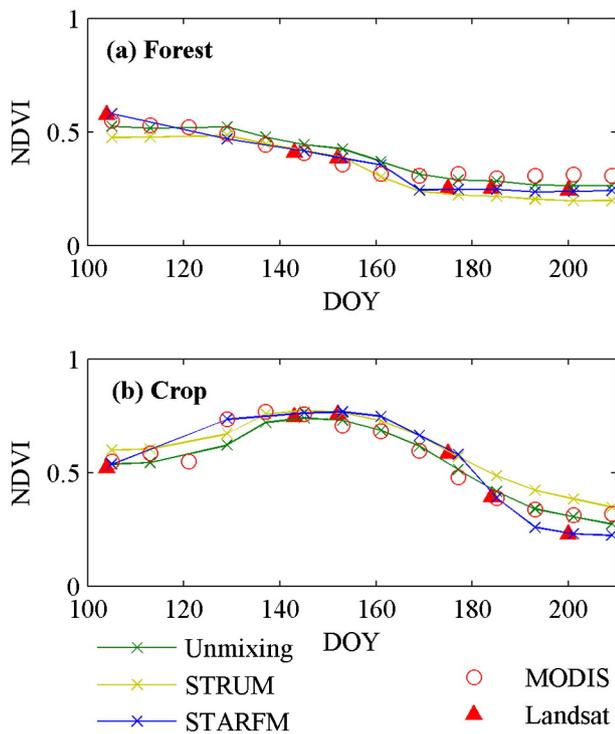
## 6. Conclusions

This study performs a benchmarking comparison and quality assessment of spatial downscaling through the data fusion methods: STARFM, a recent Landsat–MODIS data fusion method that has gained much popularity over the last few years, and two unmixing-based methods. The two unmixing-based methods rely on a Bayesian approach that optimally incorporates the available prior spectral information to constrain the unmixing process. The first unmixes MODIS imagery directly. The second, STRUM, is a novel approach which unmixes the temporal differences in MODIS imagery, and then adds it back to the Landsat reflectance in a similar manner as the STARFM. To our knowledge, this is the first study in which the unmixing-based method is applied to MODIS imagery and directly compared to STARFM. All three methods are compared using Landsat and MODIS reflectance products as well as simulated imagery, and provide high-quality fused data sets. A multitemporal analysis (using 6 Landsat 8/OLI and 17 MODIS images) allowed assesses the ability of fused images to reconstruct vegetation dynamics through NDVI temporal profiles.

Moreover, the current study underlines the difficulty of comparing conceptually different data fusion methods. This limitation was partially

overcome by using simulated imagery. In the test simulating temporal change, both algorithms boasted a similar performance. However, the STARFM exhibited prediction errors in a few cases due to unsuitable neighbor selection in areas following differing temporal patterns, whereas the unmixing method had similar performances in situations with temporal change. This finding was consistent with the application to satellite imagery. Although the STARFM accurately predicted Landsat surface reflectances, the unmixing-based method produced better results when few reference Landsat images are available. These results suggest that the STARFM is more sensitive to temporal change than the unmixing-based method. The application of STRUM further minimizes cross-sensor surface reflectance biases and results in the highest correlations to reference Landsat-8 data of all three methods.

To conclude, the selection of the optimal data-fusion method depends greatly on the intended application. STRUM is expected to have the best performance in most applications, such as gap-filling, cloud-replacement, and observing temporal dynamics in situations where limited high-resolution images are available. Other unmixing-based methods maintain the spectral information of the medium-resolution image. STARFM may have a higher performance in situations where many high-resolution images are available. Although the current study applied data fusion to Landsat and MODIS images, the techniques described here can easily be applied to other sensors. Further research could continue analyzing algorithm performance in various situations to facilitate the selection of the most appropriate algorithm in future applications.



**Fig. 7.** Temporal NDVI profiles of homogenous MODIS pixels, compared to Landsat, STARFM, unmixing-based and STRUM NDVI values resampled to MODIS resolution (463 m) (for interpretation of the references to color in this figure legend, the reader is referred to the web version of the article).

## Acknowledgment

We would like to thank Dr. Feng Gao and Dr. Jeff Masek (NASA Landsat Ecosystem Disturbance Adaptive Processing System) for making the STARFM C code available through the LEDAPS website (<http://ledaps.nascom.nasa.gov/index.html>). Funding support of ERMES (EU FP7-Space-2013, Contract 606983), RESET CLIMATE (CGL2012–35831) and LSA SAF projects is acknowledged. Thanks are due to anonymous reviewers for their very inspiring comments on a previous version of the paper.

## References

- Amorós-López, J., Gómez-Chova, L., Alonso, L., Guanter, L., Zurita-Milla, R., Moreno, J., et al. (2013). Multitemporal fusion of Landsat/TM and ENVISAT/MERIS for crop monitoring. *International Journal of Applied Earth Observation and Geoinformation*, 23, 132–141. <http://dx.doi.org/10.1016/j.jag.2012.12.004>.
- Anderson, M. C., Kustas, W. P., Norman, J. M., Hain, C. R., Mecikalski, J. R., Schultz, L., et al. (2011). Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery. *Hydrology and Earth System Sciences*, 15(1), 223–239. <http://dx.doi.org/10.5194/hess-15-223-2011>.
- Arai, E., Shimabukuro, Y. E., Pereira, G., & Vijaykumar, N. L. (2011). Remote sensing. *Remote Sensing*, 3, 1943–1956. <http://dx.doi.org/10.3390/rs3091943>.
- Bhandari, S., Phinn, S., & Gill, T. (2012). Preparing Landsat Image Time Series (LITS) for monitoring changes in vegetation phenology in Queensland, Australia. *Remote Sensing*, 4(12), 1856–1886. <http://dx.doi.org/10.3390/rs4061856>.
- Bioucas-Dias, J. M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Member, S., et al. (2012). Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(2), 354–379. <http://dx.doi.org/10.1109/JSTARS.2012.2194696>.
- Canty, M. J., & Nielsen, A. a. (2008). Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted MAD transformation. *Remote Sensing of Environment*, 112(3), 1025–1036. <http://dx.doi.org/10.1016/j.rse.2007.07.013>.
- Chavez, P.S. J. (1988). An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, 24(3), 459–479. [http://dx.doi.org/10.1016/0034-4257\(88\)90019-3](http://dx.doi.org/10.1016/0034-4257(88)90019-3).
- Ehlers, M. (1991). Multisensor image fusion techniques in remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 46(1), 19–30. [http://dx.doi.org/10.1016/0924-2716\(91\)90003-E](http://dx.doi.org/10.1016/0924-2716(91)90003-E).

- Emelyanova, I. V., McVicar, T. R., Van Niel, T. G., Tao Li, L., & Van Dijk, A. I. J. M. (2013). Remote sensing of environment assessing the accuracy of blending Landsat – MODIS surface reflectances in two landscapes with contrasting spatial and temporal dynamics : A framework for algorithm selection. *Remote Sensing of Environment*, 133, 193–209. <http://dx.doi.org/10.1016/j.rse.2013.02.007>.
- Fasbender, D., Radoux, J., & Bogaert, P. (2008). Bayesian data fusion for adaptable image pansharpening. *IEEE Transactions on Geoscience and Remote Sensing*, 46(6), 1847–1857. <http://dx.doi.org/10.1109/TGRS.2008.917131>.
- Feng, M., Sexton, J. O., Huang, C., Masek, J. G., Vermote, E. F., Gao, F., et al. (2013). Global surface reflectance products from Landsat: Assessment using coincident MODIS observations. *Remote Sensing of Environment*, 134, 276–293. <http://dx.doi.org/10.1016/j.rse.2013.02.031>.
- Gao, F., Anderson, M. C., Kustas, W. P., & Wang, Y. (2012). Simple method for retrieving leaf area index from Landsat using MODIS leaf area index products as reference. *Journal of Applied Remote Sensing*, 6(1), 063554. <http://dx.doi.org/10.1117/1.JRS.6.063554>.
- Gao, F., Masek, J., Schwaller, M., & Hall, F. (2006). On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, 44(8), 2207–2218. <http://dx.doi.org/10.1109/TGRS.2006.872081>.
- García-Haro, F. J., Sommer, S., & Kemper, T. (2005). A new tool for variable multiple endmember spectral mixture analysis (VMESMA). *International Journal of Remote Sensing*, 26(10), 2135–2162. <http://dx.doi.org/10.1080/01431160512331337817>.
- Gómez-chova, L., Zurita-milla, R., Alonso, L., Amorós-lópez, J., Guanter, L., & Camps-valls, G. (2011). Gridding artifacts on medium-resolution satellite image time series: MERIS case study. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 49(7), 2601–2611. <http://dx.doi.org/10.1109/TGRS.2011.2108660>.
- Gu, J., Li, X., Huang, C., & Okin, G. S. (2009). A simplified data assimilation method for reconstructing time-series MODIS NDVI data. *Advances in Space Research*, 44(4), 501–509. <http://dx.doi.org/10.1016/j.asr.2009.05.009>.
- Hilker, T., Wulder, M. a., Coops, N. C., Linke, J., McDermid, G., Masek, J. G., et al. (2009). A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sensing of Environment*, 113(8), 1613–1627. <http://dx.doi.org/10.1016/j.rse.2009.03.007>.
- Hmimina, G., Dufrene, E., Pontailleur, J. -Y., Delpierre, N., Aubinet, M., Caquet, B., et al. (2013). Evaluation of the potential of MODIS satellite data to predict vegetation phenology in different biomes: An investigation using ground-based NDVI measurements. *Remote Sensing of Environment*, 132, 145–158. <http://dx.doi.org/10.1016/j.rse.2013.01.010>.
- Hwang, T., Song, C., Bolstad, P. V., & Band, L. E. (2011). Downscaling real-time vegetation dynamics by fusing multi-temporal MODIS and Landsat NDVI in topographically complex terrain. *Remote Sensing of Environment*, 115(10), 2499–2512. <http://dx.doi.org/10.1016/j.rse.2011.05.010>.
- Jiang, Z., Huete, A.R., Chen, J., Li, J., Yan, G., & Zhang, X. (2006). Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction. *Remote Sensing of Environment*, 101(1), 366–378.
- Lillo-Saavedra, M., Gonzalo, C., Arquero, A., & Martinez, E. (2005). Fusion of multispectral and panchromatic satellite sensor imagery based on tailored filtering in the Fourier domain. *International Journal of Remote Sensing*, 26(6), 1263–1268. <http://dx.doi.org/10.1080/01431160412331330239>.
- Liu, H., & Weng, Q. (2012). Enhancing temporal resolution of satellite imagery for public health studies: A case study of West Nile Virus outbreak in Los Angeles in 2007. *Remote Sensing of Environment*, 117, 57–71. <http://dx.doi.org/10.1016/j.rse.2011.06.023>.
- Lunetta, R. S., Lyon, J. G., Guindon, B., & Elvidge, C. D. (1998). North American landscape characterization dataset development and data fusion issues. *Photogrammetric Engineering & Remote Sensing*, 64(8), 821–829.
- Lunetta, R. S., Shao, Y., Ediriwickrema, J., & Lyon, J. G. (2010). Monitoring agricultural cropping patterns across the Laurentian Great Lakes Basin using MODIS-NDVI data. *International Journal of Applied Earth Observation and Geoinformation*, 12(2), 81–88. <http://dx.doi.org/10.1016/j.jag.2009.11.005>.
- Mello, M., Vieira, C., Rudorff, B., Aplin, P., Santos, R., & Aguiar, D. (2013). STARS: A new method for multitemporal remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 51(4), 1897–1913. <http://dx.doi.org/10.1109/TGRS.2012.2215332>.
- Middleton, E. M., Ungar, S. G., Mandl, D. J., Ong, L., Frye, S. W., Campbell, P. E., et al. (2013). The Earth Observing One (EO-1) satellite mission: Over a decade in space. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(2), 243–256. <http://dx.doi.org/10.1109/JSTARS.2013.2249496>.
- Murphy, K. (2012). *Machine learning: A probabilistic perspective*. Cambridge, Massachusetts, USA: MIT Press, 1098 (Retrieved from <http://dl.acm.org/citation.cfm?id=2380985>).
- Nielsen, A. A. (2007). The regularized iteratively reweighted MAD method for change detection in multi- and hyperspectral data. *IEEE Transactions on Image Processing*, 16(2), 463–478 (Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/17269639>).
- Pohl, C., & Van Genderen, J. L. (1998). Review article. Multisensor image fusion in remote sensing: Concepts, methods and applications. *International Journal of Remote Sensing*, 19, 823–854. <http://dx.doi.org/10.1080/014311698215748>.
- Roy, D. P., Ju, J., Lewis, P., Schaaf, C., Gao, F., Hansen, M., et al. (2008). Multi-temporal MODIS-Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. *Remote Sensing of Environment*, 112(6), 3112–3130. <http://dx.doi.org/10.1016/j.rse.2008.03.009>.
- Singh, D. (2011). Generation and evaluation of gross primary productivity using Landsat data through blending with MODIS data. *International Journal of Applied Earth Observation and Geoinformation*, 13(1), 59–69. <http://dx.doi.org/10.1016/j.jag.2010.06.007>.
- Steven, M.D., Malthus, T. J., Baret, F., Xu, H., & Chopping, M. J. (2003). Intercalibration of vegetation indices from different sensor systems. *Remote Sensing of Environment*, 88(4), 412–422. <http://dx.doi.org/10.1016/j.rse.2003.08.010>.

- Su, Z., Timmermans, W., Gieske, a., Jia, L., Elbers, J. a., Oliso, a., et al. (2008). Quantification of land–atmosphere exchanges of water, energy and carbon dioxide in space and time over the heterogeneous Barrax site. *International Journal of Remote Sensing*, 29(17–18), 5215–5235, <http://dx.doi.org/10.1080/01431160802326099>.
- Tan, B., Woodcock, C. E., Hu, J., Zhang, P., Ozdogan, M., Huang, D., et al. (2006). The impact of gridding artifacts on the local spatial properties of MODIS data: Implications for validation, compositing, and band-to-band registration across resolutions. *Remote Sensing of Environment*, 105(2), 98–114, <http://dx.doi.org/10.1016/j.rse.2006.06.008>.
- Teillet, P.M., Fedosejevs, G., Thome, K., & Barker, J. L. (2007). Impacts of spectral band difference effects on radiometric cross-calibration between satellite sensors in the solar-reflective spectral domain. *Remote Sensing of Environment*, 110(3), 393–409, <http://dx.doi.org/10.1016/j.rse.2007.03.003>.
- Van Leeuwen, W. J.D., Orr, B. J., Marsh, S. E., & Herrmann, S. M. (2006). Multi-sensor NDVI data continuity: Uncertainties and implications for vegetation monitoring applications. *Remote Sensing of Environment*, 100(1), 67–81, <http://dx.doi.org/10.1016/j.rse.2005.10.002>.
- Villa, a, Chanussot, J., Benediktsson, J. a, Jutten, C., & Dambreville, R. (2013). Unsupervised methods for the classification of hyperspectral images with low spatial resolution. *Pattern Recognition*, 46(6), 1556–1568, <http://dx.doi.org/10.1016/j.patcog.2012.10.030>.
- Wald, L. (2002). *Data fusion: Definitions and architectures: Fusion of images of different spatial resolutions*. Paris, France: Presses des MINES, 198.
- Walker, J. J., de Beurs, K. M., Wynne, R. H., & Gao, F. (2012). Evaluation of Landsat and MODIS data fusion products for analysis of dryland forest phenology. *Remote Sensing of Environment*, 117, 381–393, <http://dx.doi.org/10.1016/j.rse.2011.10.014>.
- Xin, Q., Olofsson, P., Zhu, Z., Tan, B., & Woodcock, C. E. (2013). Remote sensing of environment toward near real-time monitoring of forest disturbance by fusion of MODIS and Landsat data. *Remote Sensing of Environment*, 135, 234–247, <http://dx.doi.org/10.1016/j.rse.2013.04.002>.
- Zhu, X., Chen, J., Gao, F., Chen, X., & Masek, J. G. (2010). An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sensing of Environment*, 114(11), 2610–2623, <http://dx.doi.org/10.1016/j.rse.2010.05.032>.
- Zurita-Milla, R., Clevers, J. G. P. W., & Schaepman, M. E. (2008). Unmixing-based Landsat TM and MERIS FR data fusion. *IEEE Geoscience and Remote Sensing Letters*, 5(3), 453–457, <http://dx.doi.org/10.1109/LGRS.2008.919685>.
- Zurita-Milla, R., Gómez-Chova, L., Guanter, L., Clevers, J. G. P. W., & Camps-Valls, G. (2011). Multitemporal unmixing of medium-spatial-resolution satellite images: A case study using MERIS images for land-cover mapping. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11), 4308–4317, <http://dx.doi.org/10.1109/TGRS.2011.2158320>.
- Zurita-Milla, R., Kaiser, G., Clevers, J. G. P. W., Schneider, W., & Schaepman, M. E. (2009). Downscaling time series of MERIS full resolution data to monitor vegetation seasonal dynamics. *Remote Sensing of Environment*, 113(9), 1874–1885, <http://dx.doi.org/10.1016/j.rse.2009.04.011>.