Crop cycle monitoring by combining medium and high resolution optical imagery

F.J. García-Haro[1], C. Gevaert[2], M. Campos-Taberner[1]
[1]*Facultat de Física, Universitat de València, Dr. Moliner, 50. 46100-Burjassot, Spain.*
[2] *Department of Earth Observation Science, Faculty ITC, University of Twente, P.O. Box 6, Enschede 7500AA, Enschede, The Netherlands.*
j.garcia.haro@uv.es

ABSTRACT -

This work is focussed on evaluating data fusion methods between medium-resolution imagery such as MODIS and high-resolution datasets such as Landsat. A recently developed method, namely Spatial and Temporal Reflectance Unmixing Model (STRUM) is evaluated. The study area is the rice district in Valencia (Spain) and the overall objective consists in obtaining time series of Landsat-like reflectance imagery using MODIS time series and a limited number of Landsat 8 imagery. The results are suitable to monitor an agronomic season with enhanced spatial resolution within the context of ERMES (EU FP7). This will contribute to produce timely information to force model simulation to the conditions observed by satellite monitoring at a local scale.

1 INTRODUCTION

Information of the phenological development is a fundamental element in crop monitoring because it describes the actual state of the cultivated species/varieties and their relation with the pedo-climatic conditions. Multitemporal satellite imagery may provide occurrence estimation of the phenological stages in a spatially distributed model. However, factors such as inadequate spatial or temporal resolution and cloud cover have limited the effectiveness of utilizing satellite imagery.

Landsat mission provides one of the most extensively used high-resolution data sets. Landsat TM, ETM+ and OLI sensors have a spatial resolution of 30 m for the multispectral bands, which is adequate for many environmental applications. However, Landsat satellites have a revisit time of 16 days, 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 have a daily revisit period, but a lower spatial resolution which limits its effectiveness for fine-scale environmental applications.

In order to overcome this problem we focussed on evaluating data fusion methods between MODIS and Landsat. Two prevalent data fusion methods are the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) (Gao et al. 2006) and unmixing-based data fusion (Zurita-Milla, 2009, 2011; Amorós-López, 2013). In this study we apply a recently developed algoritm, namely the Spatial and Temporal Reflectance Unmixing Model (STRUM), which combines the strengths of both algorithms (Gevaert and García, 2015). The potential of this method is demonstrated using Landsat and MODIS imagery.

The ability of fused images to capture phenological variations is also assessed using temporal NDVI profiles. The study area is the rice district in Valencia (Spain) and the overall objective consists in obtaining time series of reflectance imagery during an agronomic season with enhanced spatial resolution. Improvement is achieved by downscaling time series of MODIS products to disaggregate the product medium resolution using the information about their pixel composition provided by high resolution data.

2 METHODOLOGY

STRUM is an unmixing based method inspired by STARFM principles, i.e. 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.

$$\textbf{\textit{residual}} = M(t_k) - M(t_0) \qquad (1)$$

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.

Unmixing-based image fusion applies four steps to solve the linear mixing model (see figure 1):

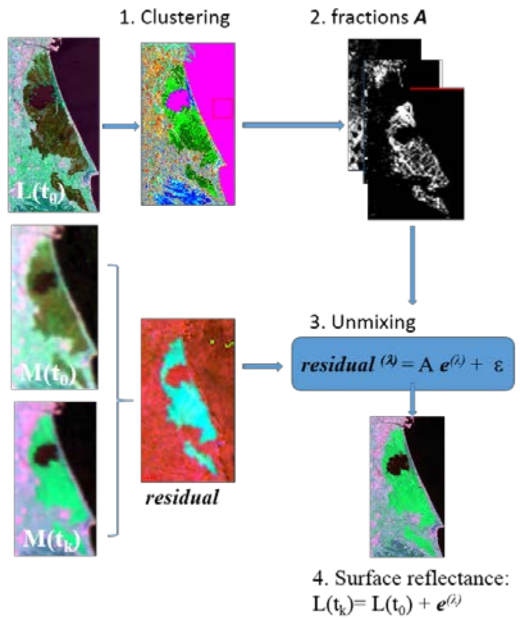


Figure 1. Flow chart of the algorithm for FVC determination

(1) Cluster the high-resolution dataset to define the endmembers. The present study applies the k-means algorithm to identify *k* spectral clusters.
(2) Apply a sliding window of [*n*×*n*] MODIS pixels to the clustered image to record the endmember fractions **A**, a [n^2×*k*] matrix with the abundances of each endmember within each medium-resolution pixel. The use of a sliding window allows for spectral differences between pixels of the same cluster in different locations.
(3) Unmix the medium-resolution residual. The aim is to solve for $\boldsymbol{e}^{(\lambda)}$ a [*k*×*1*] column vector that contains temporal changes of each endmember spectra in MODIS band (λ). $\boldsymbol{residual}^{(\lambda)}$ is a [n^2×*1*] column vector residual values (Ec. 1) of each MODIS pixel in the *n*×*n* moving window for the MODIS band λ which is currently being unmixed. This is achieved by minimizing the residuals (ε) of the linear model (Eq. 2).

$$\boldsymbol{residual}^{(\lambda)} = \mathrm{A}\,\boldsymbol{e}^{(\lambda)} + \varepsilon \qquad (2)$$

We propose a Bayesian approach, which allows specifying the variances σ_0^2 and σ_r^2 assigned to the prior of $\boldsymbol{e}^{(\lambda)}$ and the noisy reflectance data $\boldsymbol{r}^{(\lambda)}$, respectively. For a detailed description, we refer to Murphy (2012). If the precision assigned to the prior is strong relative to the data strength (σ_0 is large relative to σ_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 $\sigma_{ratio}=\sigma_r/\sigma_0$.

(4) Create 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:

$$\mathrm{L}(t_k) = \mathrm{L}(t_0) + \boldsymbol{e}^{(\lambda)} \qquad (3)$$

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

3. RESULTS

3.1 Study area

The study area is in the Albufera National park region near Valencia, Spain (39.33ºN, 0.36ºW). It is a rectangular area of 1008 km^2, which corresponds to 808×1399 Landsat pixels. The lands occupied by the cultivation of rice amounts to 16,000 ha, which is the area that has remained stable in recent years. The area has a typical Mediterranean climate, mild, with an average annual humidity of 65%. The average annual temperature is 17° C. Their mean values ranging from 11° C in January and 27° C in August. The mean annual precipitation is approximately 430 mm. Precipitation tends to be intense and concentrated in autumn.

Figure 2. Location of the study area

3.2 Satellite image preprocessing

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: May 3, May 19, June 4, June 20, July 15, July 31 and September, 1, 2014.

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. MCD43A4 images corresponding to the same periods 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 et al. 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 (see Figure 3). For each band, slope and intercept parameters were obtained using orthogonal linear regression, since it outperformed the ordinary least squares regression.

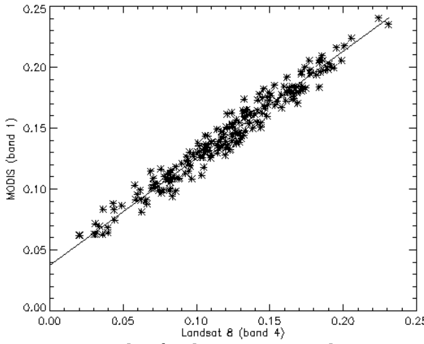


Figure 3. Example of radiometric normalization in red waveband, corresponding to Landsat8/OLI June 4 (target) vs. MCD43 May, 25 – June, 10 (reference). Solid line: orthogonal regression.

3.3 Data fusion optimization

Optimal input parameters were identified using the Landsat 8/OLI image on June 4, 2014 and the MODIS composite from May 25-June 10^{th} as base images, and the MODIS composite from July 12-28^{th} to represent the prediction date (see Figure 4). The number of clusters was varied using the values k=10, 20, 40, and 80. The window size was also varied from w=5 to 41 MODIS pixels in steps of 4. The weight of the *a priori* endmember information was varied, using the values σ_{ratio}=0.01, 1, 2, and 5.

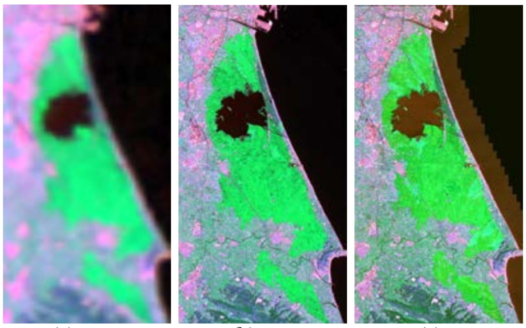

(a) (b)

Figure 4. The base pair ($M(t_0)$, $L(t_0)$) used to predict surface reflectance for each $M(t_k)$. (a) MODIS composite from May, 25 – June, 10^{th} ; (b) Landsat 8/OLI image on June 4.

The accuracy was evaluated taking into account the *Erreur Relative Globale Adimensionnelle de Synthèse* (ERGAS) (Wald, 2002) and the correlation coefficient as a quality indicators. ERGAS 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. The optimal parameters in the current study were defined as w=9, k=20 and σ_{ratio} =1.0, although other combinations are also possible.

3.4 Creation of fused Landsat 8/OLI Reflectance

The application of these optimal parameters produced surface reflectances very similar to Landsat imagery, as indicated in Figure 5. These results correspond to a case with significant temporal change and suggest that a significant time between the base image date and prediction date did not significantly affect the ability to correctly predict Landsat-8 reflectance. Even better results were found in cases with little temporal change.



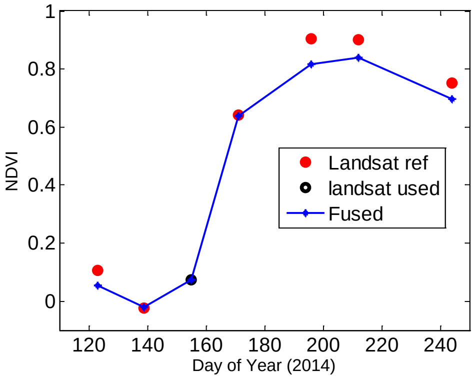
(a) (b) (c)

Figure 5. Results of data fusion in a case with significant temporal change: using June 3th as a base date to predict reflectance on September 1st. (a) The MODIS composite of August 21-September 6, 2014; (b) the predictions using STRUM. (b) a reference (unused) Landsat image on September 1st. All images are displayed with the band combination R:Red, G:NIR, B: SWIR.

3.5 Derivation of temporal profiles

In order to assess whether the fused images are suitable for studying vegetation dynamics, the final experiment consisted of applying the STRUM algorithm to create NDVI temporal profiles of 8 fused images. The NDVI time series were assessed by examining seasonal variations over representative agricultural fields. The aim of the experiment was to fill gaps in the Landsat 8 temporal profiles by using only one Landsat 8 image to create the time series and 8 MODIS data. This simulates realistic situations in which few high-resolution images are available. Furthermore, it allows unused Landsat images to act as a testing dataset to analyze the quality of the fused NDVI predictions.

The reconstruction of the Landsat-like rice cycle (May-September 2014) is illustrated in Figure 1. The Landsat image on June 4 and the MODIS composite from May 25-June 10th were used as base images.



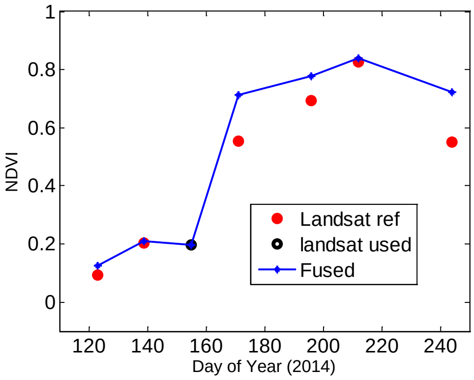


Figure 6. Temporal profiles of the STRUM NDVI predictions for a scenario using a single Landsat image to create the temporal profiles. Results correspond to two representative agricultural fields.

Results confirm that STRUM has the potential to transfer temporal information from MODIS to Landsat in order to capture phenological variations. The method clearly outperforms existing literature methods at observing temporal dynamics in situations where limited high-resolution images are available. For example, Gevaert and García-Haro (2015) found that STARFM fails to capture phenological variations when having only 1 or 2 input Landsat 8 images and predictions are highly dependent on the number input images.

3 CONCLUSIONS

Multi-sensor data fusion provides an opportunity to obtain imagery products which transcend physical sensor limitations. The STRUM algorithm has been recently proposed to produce Landsat-like reflectances while preserving the spatial patterns found in Landsat images. This paper explores the potential of this method to monitor an agronomic season of rice fields with enhanced spatial resolution, while incorporating information regarding temporal variation in reflectances obtained from the input MODIS imagery.

The STRUM produced surface reflectances very similar to Landsat imagery, even in challenging scenarios with a significant time between the base image date and prediction date. The ability of fused images to capture phenological variations has been also demonstrated. Temporal profiles of STRUM NDVI closely resembled Landsat NDVI profiles in experiments simulating situations where few input high-resolution images are available.

The results of this study confirm previous findings that STRUM is well suited for data fusion applications requiring Landsat-like surface reflectances, such as gap-filling and cloud replacement. The method has potential to generate high resolution imagery to monitor the rice agronomic season and provide timely information to force model simulation to the conditions observed by satellite monitoring at a local scale.

4 ACKNOWLEDGMENTS

Funding support of ERMES (EU FP7-Space-2013, Contract 606983), RESET CLIMATE (CGL2012–35831) and LSA SAF (EUMETSAT) projects is acknowledged.

5 REFERENCES

Amorós-López, J., Gómez-Chova, L., Alonso, L., Guanter, L., Zurita-Milla, R., Moreno, J., & Camps-Valls, G., 2013, Multitemporal fusion of Landsat/TM and ENVISAT/MERIS for crop monitoring. *International Journal of Applied Earth Observation and Geoinformation*, ***23***, 132–141.

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***, 1025–1036.

Chavez, P. S. J., 1988, An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, **24**, 459–479.

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***, 2207–2218.

Gevaert, C.M. and F.J. García-Haro, 2015, A Comparison of STARFM and an Unmixing-based Algorithm for Landsat and MODIS Data Fusion, *Remote sensing of Environment*, **156**, 34-44.

Murphy, K., 2012, *Machine learning: a probabilistic perspective* (p. 1098). Cambridge, Massachusetts, USA: MIT Press.

Nielsen, A. A., 2007, The regularized iteratively reweighted MAD method for change detection in multi- and hyperspectral data. *IEEE Transactions on Image Processing*, ***16***, 463–78.

Roy, D. P., Ju, J., Lewis, P., Schaaf, C., Gao, F., Hansen, M., & Lindquist, E., 2008, Multi-temporal MODIS–Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. *Remote Sensing of Environment*, ***112***, 3112–3130.

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***, 393–409.

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**, 1874–1885.

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**, 4308–4317.

Wald, L., 2002, *Data Fusion: Definitions and Architectures : Fusion of Images of Different Spatial Resolutions* (p. 198). Paris, France: Presses des MINES.