Unmixing-Based Landsat TM and MERIS FR Data Fusion

Zurita-Milla, R ; Clevers, J G P W ; Schaepman, Michael E

DOI: https://doi.org/10.1109/LGRS.2008.919685

Originally published at:
Zurita-Milla, R; Clevers, J G P W; Schaepman, Michael E (2008). Unmixing-Based Landsat TM and MERIS FR Data Fusion. Geoscience and Remote Sensing Letters, 5(3):453-457.
DOI: https://doi.org/10.1109/LGRS.2008.919685

Unmixing-Based Landsat TM and MERIS FR Data Fusion

Raul Zurita-Milla, Jan G. P. W. Clevers, and Michael E. Schaepman, *Senior Member*

***Abstract*—An unmixing-based data fusion technique is used to generate images that have the spatial resolution of Landsat Thematic Mapper (TM) and the spectral resolution provided by the Medium Resolution Imaging Spectrometer (MERIS) sensor. The method requires the optimization of the following two parameters: the number of classes used to classify the TM image and the size of the MERIS "window" (neighborhood) used to solve the unmixing equations. The ERGAS index is used to assess the quality of the fused images at the TM and MERIS spatial resolutions and to assist with the identification of the best combination of the two parameters that need to be optimized. Results indicate that it is possible to successfully downscale MERIS full resolution data to a Landsat-like spatial resolution while preserving the MERIS spectral resolution.**

***Index Terms*—ERGAS, fusion quality, Landsat, linear mixing model, Medium Resolution Imaging Spectrometer (MERIS), spatial unmixing.**

I. Introduction

During the last few years, data fusion methods have received more and more attention from the remote sensing community because of the increasing need to integrate the vast amount of data that are being collected by Earth observation satellites. As a result, a large number of data fusion methods have been developed (see, for example, [1]–[4] for a review). In this letter, we focus on the implementation and evaluation of the so-called unmixing-based data fusion approach [5]. The aim of this data fusion approach is to combine two images acquired over the same area but at different spatial resolutions to produce an image with the spatial resolution of the high spatial resolution image and the spectral resolution of the low spatial resolution image. Often, the selected low spatial resolution image has a better spectral resolution than the high spatial resolution image. As a result, the fused image has (potentially) more information than each of the original images. A simplified version of this data fusion approach has been used by Minghelli-Roman *et al.* [6], [7] to combine Medium Resolution Imaging Spectrometer (MERIS) full resolution (FR) and Landsat Enhanced Thematic Mapper (ETM) data for coastal water monitoring. In their approach, only one parameter, namely, the number of classes used to classify the high resolution image, needs to be optimized, because they solved the unmixing equations for the whole image at once. A large number of classes (typically > 100) are needed to achieve good results with this method [6], [7]. However, such a large number of classes are not always feasible or realistic. For instance, if an existing land cover classification is used to get the high spatial resolution information, then the number of classes is limited and, in most cases, well below 100. Furthermore, solving the unmixing equations for the whole image at once might severely hamper the quality and usability of the fused images, because all pixels belonging to one class will get the same spectral signature. In other words, if we apply the method as described by Minghelli-Roman *et al.* [6], [7], we implicitly reject all the within-class variability. Therefore, we believe that using a neighborhood should be preferred over simultaneously solving the unmixing for all the pixels present in the scene. For that reason, here, we implement a detailed version of the unmixing-based fusion algorithm where two parameters, the number of classes used to classify the Thematic Mapper (TM) image and the size of the MERIS FR neighborhood used to solve the unmixing equations, need to be optimized. The ERGAS index [3] is used to support the optimization of these two parameters and to quantitatively assess the quality of the fused images.

Finally, this letter presents a case study that uses MERIS FR and Landsat TM data over land, because several studies have proven the potential of MERIS for this kind of applications [8]–[10]. If the proposed data fusion approach proves to be successful, the resulting fused images could be used to improve land cover maps and/or to monitor ecosystems at high spatial and spectral resolutions.

II. Methodology

The study area covers approximately 40 km $\times$ 60 km of the central part of The Netherlands (52.19° N, 5.91° E). A Landsat-5 TM image from July 10, 2003 and a MERIS FR level 1b image acquired on July 14, 2003 were available over this area. The TM image was georeferenced to the Dutch national coordinate system (RD) using a cubic convolution resampling method and a pixel size of 25 m. The digital numbers of the TM image were converted into radiances (in watts per square meter per steradian per micrometer) using the latest calibration coefficients [11] to ensure that both the TM and the MERIS image are in the same radiometric units. The MERIS FR level 1b image (300 m pixel size and radiances in watts per square meter per steradian per micrometer) was first corrected for the smile effect [12]. Then, an image-to-image coregistration was performed in order to ensure the best

Manuscript received June 2, 2007; revised January 21, 2008. The work of R. Zurita-Milla was supported by the Dutch SRON GO program under Grant EO-061.

The authors are with the Centre for Geo-Information, Wageningen University, 6700 AA Wageningen, The Netherlands (e-mail: raul.zurita-milla@wur.nl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2008.919685

possible match between the two images. In this process, the TM image was used as a reference, and a nearest neighbor resampling was used not to modify the original MERIS pixel values. The pixel sizes of the TM and MERIS FR sensors were preserved, which implies that 144 TM pixels are inside each MERIS FR pixel. Subsequently, the TM and the MERIS FR images were fused using an unmixing-based data fusion approach. The method consists of the following four main steps [5].

First, the high spatial resolution image is used to identify the main components (i.e., spectral groups) of the study area. For this purpose, the TM image was classified into nc unsupervised classes using the ISODATA classification rule. In this letter, five nc values were used: 10, 20, 40, 60, and 80.

Second, a sliding window of $k \times k$ MERIS FR pixels is applied to each of the TM classified images to generate class-proportion matrices. These matrices contain the proportions of each of the nc classes that fall within each of the MERIS FR pixels that are inside the $k \times k$ window. In this letter, 14 window sizes (from now on, referred to as neighborhoods) were tested: from $k = 5$ to $k = 53$ in steps of four.

Third, the spectral information of all the classes present in the $k \times k$ neighborhood are unmixed using the proportion matrices and their corresponding MERIS FR radiance values. Here, it is important to notice that the unmixing is solved for each low resolution band independently. Therefore, care must be taken to select a neighborhood size (k^2) larger than or equal to the number of classes present in the neighborhood, because each MERIS FR pixel provides only one (mixing) equation. Although the unmixing is solved for all the classes present in the neighborhood, only the spectral information of the classes present in the central pixel of the neighborhood is kept because that is the pixel that is being effectively unmixed.

Finally, each of the TM unsupervised classes present in the central pixel of the neighborhood is replaced by its corresponding unmixed MERIS signal. By repeating this operation for all the MERIS FR pixels, for all MERIS bands, and for all the possible combinations of nc and k, a series of fused images is generated.

Fig. 1 shows the four steps of the unmixing-based data fusion approach and presents a matrix–vector notation for the third step (i.e., the unmixing). This notation should be interpreted as follows:

$$\mathbf{L}^{i,k} = \mathbf{P}^{k,nc} \cdot \mathbf{S}^{i,k,nc} + \mathbf{E}^{i}, \qquad i = 1, 2, \ldots, N \tag{1}$$

where

- $\mathbf{L}^{i,k}$ is a $(k^2 \times 1)$ vector that contains the values of band i for all the MERIS FR pixels present in the neighborhood k;
- $\mathbf{P}^{k,nc}$ is a $(k^2 \times nc)$ matrix containing the proportions of the TM unsupervised classes that fall inside each of the MERIS FR pixels present in the neighborhood k;
- $\mathbf{S}^{i,k,nc}$ is the $(nc \times 1)$ unknown vector of unmixed spectral information (band-i radiances) for each of the classes present in k;

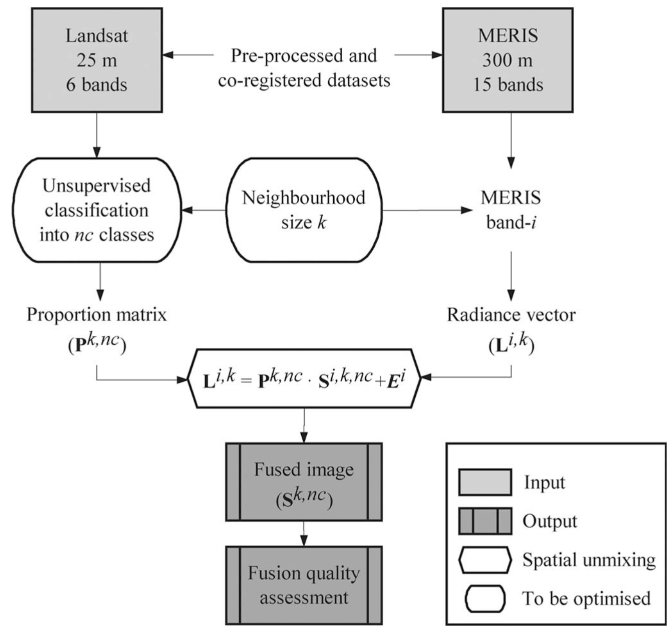

Fig. 1. Scheme of the general methodology.

- $\boldsymbol{S}^{k,nc}$ is the fused image after iterating over all MERIS pixels and all MERIS bands;
- $\mathbf{E}^{i}$ is a $(k^2 \times 1)$ vector of residual errors;
- N is the total number of bands of the low-resolution image.

This formulation of the unmixing-based data fusion indirectly implies that the number of classes used to classify the TM image (nc) and the size of the MERIS FR neighborhood (k) need to be optimized. nc needs to be optimized, because it depends on the spectral variability of the scene (heterogeneous scenes will most likely require a larger nc value than homogeneous ones). k also needs to be optimized, because it has a great impact on the spectral quality of the fused image. On the one hand, k should be kept as small as possible so that the fused image is spectrally dynamic and consistent with the variability recorded by the low spatial resolution sensor. On the other hand, k should be sufficiently large to provide enough equations to solve the unmixing. In other words, (1) is a system of k^2 equations (one equation per low resolution pixel in the neighborhood) with up to nc unknowns (depending on the number of classes present in such a neighborhood). This means that k^2 must be greater than or equal to the number of classes inside the MERIS neighborhood. However, if we use very large k values, the output image will have low spectral variability, because each system of equations results in a unique solution. For instance, if the size of the neighborhood matches the size of the scene ($k =$ image size), then all the pixels of one class identified with Landsat TM will have the same spectral response independently of their position within the scene. Using $k =$ image size, therefore, results in a fused image with a low spectral dynamic range, where each of the classes is represented by an approximation of its mean spectral response. The latter approach was the one used by Minghelli-Roman *et al.* [6], [7]. Although it is computationally fast (we only need to solve one system of equations), here, we prefer to also optimize the size of the neighborhood k such that we can account for the natural variability of the components present in the scene.

Finally, a constrained least squares method was used to retrieve $\mathbf{S}^{i,k,nc}$ from (1). The use of a constrained method is justified, because the solution should fulfill the following two conditions: 1) the radiance values must be positive and 2) the radiance values cannot be larger than the MERIS radiance-saturation values [13].

A. Data Fusion Quality and Optimization of nc and k

A quantitative assessment of the quality of the fused images was done at the level of the TM and of the MERIS spatial resolution. This assessment was used to support the selection of the best combination of nc and k.

Bearing in mind that any fused image should be as identical as possible to the original low resolution image once degraded back to its original resolution (coherence property [14]), we degraded the fused images $\boldsymbol{S}^{k,nc}$ to 300 m using a mean filter. After this, we assessed the quality of the degraded fused images by comparing them with the original MERIS FR image. The ERGAS index [3] was used for this comparison

$$\mathrm{ERGAS} = 100\frac{h}{l}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\mathrm{rmse}_i^2/M_i^2\right)} \qquad (2)$$

where

- h is the resolution of the high spatial resolution image (TM);
- l is the resolution of the low spatial resolution image (MERIS FR);
- N is the number of spectral bands involved in the fusion;
- rmse_i is the root mean square error computed between the degraded fused image and the original MERIS image (for the band i);
- M_i is the mean value of the band i of the reference image (MERIS).

The ERGAS index equals zero when the degraded fused image (300 m) is equal to the original MERIS FR image. Therefore, low ERGAS values indicate high image fusion quality.

If we assume that spectrally corresponding bands are highly correlated for images that have been acquired nearly at the same date, then the ERGAS index can also be used to evaluate the quality of the fused images at 25 m. This ERGAS will be named $\mathrm{ERGAS}_{\mathrm{TM}}$ (because the Landsat TM image will be used as a reference), whereas the ERGAS computed at 300 m will be referred to as $\mathrm{ERGAS}_{\mathrm{M}}$ (MERIS used as a reference). Other terms, like spatial and spectral ERGAS, have been identified in literature to indicate that the ERGAS index is computed at different spatial resolutions [15]–[18].

The expression used to compute the $\mathrm{ERGAS}_{\mathrm{TM}}$ is basically the same as (2) except for the following: 1) the rmse_i is computed between Landsat TM bands 1–4, and their spectrally corresponding fused bands (3, 5, 7, and 13, respectively) and 2) M_i corresponds to the mean of the band i of the TM image. The $\mathrm{ERGAS}_{\mathrm{TM}}$ index never reaches a zero value, because the bands that were used for the calculation of this index have slightly different characteristics (band centers and bandwidths). Despite this, its values can be used to assess the quality of the fused images, because—similar to the $\mathrm{ERGAS}_{\mathrm{M}}$—the lower the $\mathrm{ERGAS}_{\mathrm{TM}}$, the better the quality of the fused image.

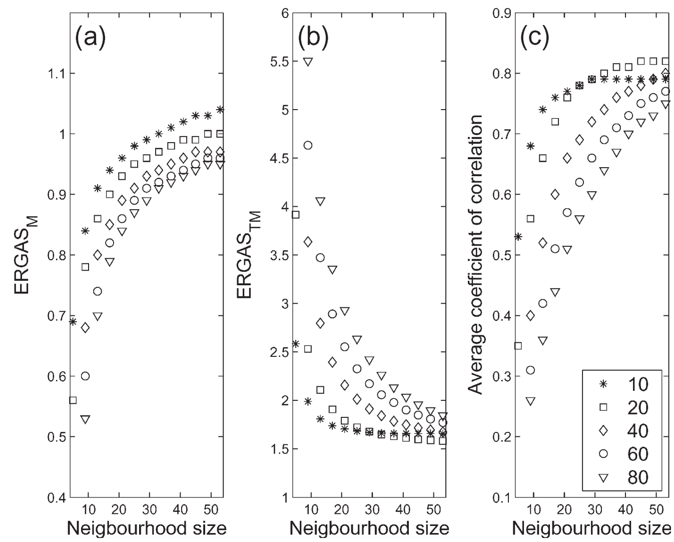

Fig. 2. Results for (a) $\mathrm{ERGAS}_{\mathrm{M}}$, (b) $\mathrm{ERGAS}_{\mathrm{TM}}$, and (c) coefficient of correlation. Each symbol represents the number of classes used to classify the TM image. Notice that the neighborhood size of five does not provide sufficient equations to solve the unmixing when the TM image is classified into 40, 60, and 80 classes.

In order to better understand, evaluate, and benchmark the values obtained for the $\mathrm{ERGAS}_{\mathrm{TM}}$, the average coefficient of correlation ($\bar{r}$) was also computed at 25 m. First, the coefficient of correlation was computed for the four pairs of bands used to compute the $\mathrm{ERGAS}_{\mathrm{TM}}$. Then, these values were averaged to produce a single $\bar{r}$ value for each of the fused images.

III. Results and Discussion

A. TM and MERIS FR Data Fusion

Fig. 2 shows the ERGAS indexes and the $\bar{r}$ values for all fused images that were generated for the different combinations of nc and k.

Most fused images yielded low ERGAS values [Fig. 2(a) and (b)], which means that the unmixing-based data fusion succeeded in synthesizing the spectral information of the MERIS FR image at a high spatial resolution. However, relatively high $\mathrm{ERGAS}_{\mathrm{TM}}$ values (> 3) were found for the images unmixed using small k values. This might indicate that the solution of the unmixing equations is not stable when few equations are used and that regularization methods might be needed in these cases. Poor $\bar{r}$ values (< 0.45) were found when unmixing with small k values [Fig. 2(c)], whereas high $\bar{r}$ values (> 0.75) were always associated with low $\mathrm{ERGAS}_{\mathrm{TM}}$ values (< 2). Because of this opposite behavior, we conclude that the information given by the $\mathrm{ERGAS}_{\mathrm{TM}}$ and $\bar{r}$ is equivalent. For this reason, we mainly discuss the results obtained using the ERGAS indexes.

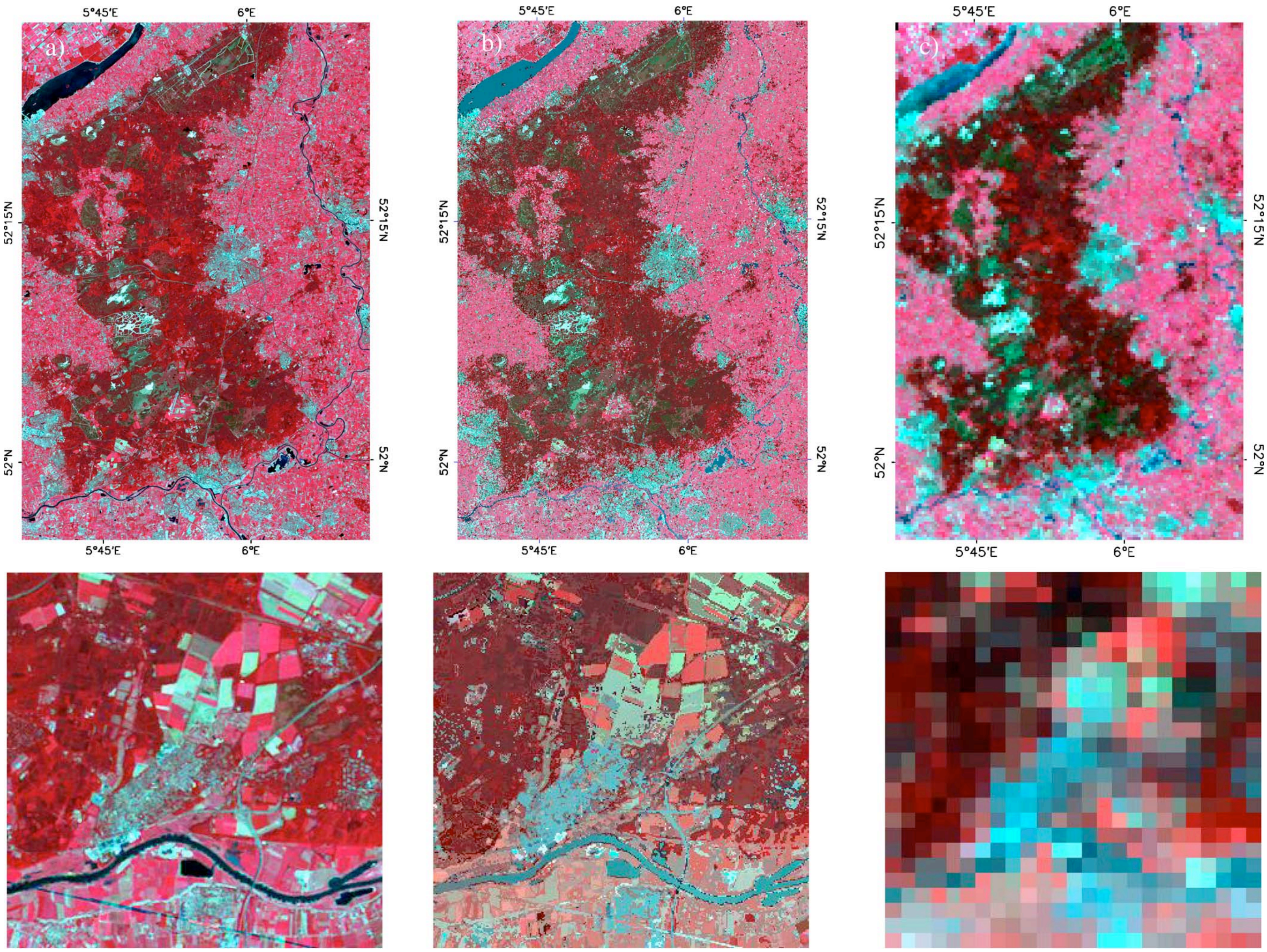

Fig. 3. RGB color composite of (a) bands 4, 3, and 2 of the TM image, (b) bands 13, 7, and 5 of the fused image obtained for $nc = 60$ and $k = 45$, and (c) bands 13, 7, and 5 of the original MERIS FR image. Upper row shows the whole study area, whereas the lower row shows a 25 × 25 pixel subset.

Two additional observations can be made from Fig. 2. First, the ERGAS indexes are inversely correlated: the $\mathrm{ERGAS_M}$ decreases when increasing the number of classes, and it increases with larger neighborhood sizes, whereas the $\mathrm{ERGAS_{TM}}$ presents the opposite behavior. This means that there is a tradeoff between the quality of the fused images at 25 and at 300 m and that we cannot find an optimum combination of nc and k that minimizes both ERGAS values. Secondly, both ERGAS indexes and the average coefficient of correlation show a saturation behavior. This means that increasing nc or k beyond the values that were tested in this letter will not improve the quality of the fused images.

The selection of the best fused image is not straightforward, because there is no combination of nc and k that simultaneously minimizes the two ERGAS indexes. However, from Fig. 2(a), we recognize that the range of variation of the $\mathrm{ERGAS_M}$ is rather small (the smoothing effect caused by increasing the window size is apparently not very important). Therefore, we could select as the best fused image the one that first minimizes the $\mathrm{ERGAS_{TM}}$ and then the $\mathrm{ERGAS_M}$. Nevertheless, a large number of fused images potentially meet this criterion. Furthermore, a visual check of these fused images showed that, indeed, they are very similar. As an illustration, Fig. 3 shows an RGB color composite of the fused image obtained with $nc = 60$ and $k = 45$ (upper row—whole study area; lower row—a 25 × 25 pixel subset). For comparison purposes, an RGB color composite of the original TM and MERIS FR images is also shown in Fig. 3.

In general, the fused image preserves well the spatial patterns found in the TM image while remaining spectrally similar to the MERIS FR image. However, some deviating pixels can be seen at the boundary between objects (e.g., river shorelines). These pixels correspond to mixed pixels, and they are difficult to unmix because the TM unsupervised classification is rather noisy in those areas and because they cover a very small fraction of the neighborhood under study.

IV. Conclusion

In this letter, we have studied the applicability of the linear mixing model to fuse a Landsat TM and a MERIS FR level 1b image. The method, known as unmixing-based data fusion, requires the optimization of the following two parameters: the number of classes used to classify the TM image nc and the size of the MERIS neighborhood k used to solve the unmixing equations. Several combinations of nc and k have been tested.

The ERGAS index computed at 25 and at 300 m, together with the average coefficient of correlation $(\bar{r})$ computed at 25 m, was used to assess the quality of the fused images and to assist in the identification of the best fused image. The results of the fusion quality assessment indicate the following: 1) the unmixing-based data fusion approach presented here succeeded in preserving the spectral information of MERIS: low values were found for the $\mathrm{ERGAS}_{\mathrm{M}}$ since all the spectral information present in the fused image comes from MERIS; in addition, high $\bar{r}$ values were found between the first four TM bands and its spectrally corresponding fused bands and 2) there is always a tradeoff between the quality of the fused images at 25 and at 300 m, because the ERGAS indexes are inversely correlated. Therefore, we suggest using a specific application (e.g., land cover mapping) to better identify the best fused image. For illustration purposes, we selected as the best fused image the one obtained for $nc = 60$ and $k = 45$, because it is one of the images that minimizes the $\mathrm{ERGAS}_{\mathrm{TM}}$ while keeping one of the lowest possible $\mathrm{ERGAS}_{\mathrm{M}}$ values. Nevertheless, further work is required in order to estimate the real possibilities and limitations of this data fusion method. Special attention should be paid to the coregistration of the high and low spatial resolution images and to the criteria used to select the best fused image.

The unmixing-based data fusion approach may also be used to fuse MERIS FR time series with one or more TM images. This multitemporal data fusion exercise will be of great interest for land cover mapping and for monitoring vegetation dynamics (e.g., in terms of fraction of absorbed photosynthetically active radiation, leaf area index, or chlorophyll content) at high spatial, spectral, and temporal resolutions. Nevertheless, it is important to realize that these fused images will only be an approximation of what the MERIS sensor would be measuring if it had a spatial resolution of 25 m. In addition, possible landscape changes between the dates of the Landsat TM acquisition and the MERIS images might further affect the quality of the fused images [7], since the number and location of land cover classes may change if the time span becomes too wide.

Acknowledgment

The authors would like to thank three anonymous reviewers for their valuable comments and recommendations.

References

[1] C. Pohl and J. L. Van Genderen, “Multisensor image fusion in remote sensing: Concepts, methods and applications,” *Int. J. Remote Sens.*, vol. 19, no. 5, pp. 823–854, 1998.

[2] M. Ehlers, “Multisensor image fusion techniques in remote sensing,” *ISPRS J. Photogramm. Remote Sens.*, vol. 46, no. 1, pp. 19–30, 1991.

[3] L. Wald, *Data Fusion Definitions and Architectures: Fusion of Images of Different Spatial Resolutions*. Paris, France: Ecole des Mines de Paris, 2002.

[4] D. L. Hall and J. Llinas, “An introduction to multisensor data fusion,” *Proc. IEEE*, vol. 85, no. 1, pp. 6–23, Jan. 1997.

[5] B. Zhukov, D. Oertel, F. Lanzl, and G. Reinhackel, “Unmixing-based multisensor multiresolution image fusion,” *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1212–1226, May 1999.

[6] A. Minghelli-Roman, M. Mangolini, M. Petit, and L. Polidori, “Spatial resolution improvement of MeRIS images by fusion with TM images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 7, pp. 1533–1536, Jul. 2001.

[7] A. Minghelli-Roman, L. Polidori, S. Mathieu-Blanc, L. Loubersac, and F. Cauneau, “Spatial resolution improvement by merging MERIS-ETM images for coastal water monitoring,” *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 2, pp. 227–231, Apr. 2006.

[8] M. M. Verstraete, B. Pinty, and P. J. Curran, “MERIS potential for land applications,” *Int. J. Remote Sens.*, vol. 20, no. 9, pp. 1747–1756, Jun. 1999.

[9] J. G. P. W. Clevers, S. M. De Jong, G. F. Epema, F. D. Van der Meer, W. H. Bakker, A. K. Skidmore, and K. H. Scholte, “Derivation of the red edge index using the MERIS standard band setting,” *Int. J. Remote Sens.*, vol. 23, no. 16, pp. 3169–3184, Aug. 2002.

[10] J. G. P. W. Clevers, M. E. Schaepman, C. A. Mücher, A. J. W. de Wit, R. Zurita-Milla, and H. M. Bartholomeus, “Using MERIS on Envisat for land cover mapping in The Netherlands,” *Int. J. Remote Sens.*, vol. 28, no. 3/4, pp. 637–652, 2007.

[11] G. Chander and B. Markham, “Revised Landsat-5 TM radiometric calibration procedures and postcalibration dynamic ranges,” *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 11, pp. 2674–2677, Nov. 2003.

[12] R. Zurita-Milla, J. G. P. W. Clevers, M. E. Schaepman, and M. Kneubuehler, “Effects of MERIS L1b radiometric calibration on regional land cover mapping and land products,” *Int. J. Remote Sens.*, vol. 28, no. 3/4, pp. 653–673, 2007.

[13] ESA EOHelpdesk, *MERIS Lsat Values Over Land*, 2006. personal communication.

[14] L. Wald, T. Ranchin, and M. Mangolini, “Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images,” *Photogramm. Eng. Remote Sens.*, vol. 63, no. 6, pp. 691–699, 1997.

[15] M. Lillo-Saavedra, C. Gonzalo, A. Arquero, and E. Martinez, “Fusion of multispectral and panchromatic satellite sensor imagery based on tailored filtering in the Fourier domain,” *Int. J. Remote Sens.*, vol. 26, no. 6, pp. 1263–1268, Mar. 2005.

[16] M. Lillo-Saavedra and C. Gonzalo, “Spectral or spatial quality for fused satellite imagery? A tradeoff solution using the wavelet à trous algorithm,” *Int. J. Remote Sens.*, vol. 27, no. 7, pp. 1453–1464, Apr. 2006.

[17] M. González-Audícana, X. Otazu, O. Fors, and J. Alvarez-Mozos, “A low computational-cost method to fuse IKONOS images using the spectral response function of its sensors,” *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1683–1691, Jun. 2006.

[18] M. Lillo-Saavedra and C. Gonzalo, “Multispectral images fusion by a joint multidirectional and multiresolution representation,” *Int. J. Remote Sens.*, vol. 28, no. 18, pp. 4065–4079, Jan. 2007.