UNMIXING-BASED LANDSAT AND MERIS IMAGE FUSION FOR LAND COVER MAPPING OVER THE NETHERLANDS

R. Zurita-Milla[1], M.L. Guillen-Climent, J.G.P.W. Clevers, and M.E. Schaepman

Wageningen University. Centre for Geo-Information. PO Box 47. 6700 AA Wageningen, The Netherlands. (raul.zurita-milla, maluz.guillen, jan.clevers, michael.schaepman)@wur.nl

Commission VII, WG VII/6

KEY WORDS: Linear mixing model, spatial unmixing, Landsat, MERIS, fusion quality, ERGAS

ABSTRACT:

Anthropogenic land use activities are significantly contributing to the ecological degradation of the Earth system. Therefore, having actual and reliable land cover information is fundamental to study the impact of such an ecological degradation on our future welfare. High spatial resolution sensors, such as Landsat TM, are typically used to derive land cover information from local to regional scales. However, current high spatial resolution sensors do not provide an appropriate temporal resolution. This is especially true for areas having high cloud coverage throughout all the year. In this respect, The Medium Resolution Imaging Spectrometer, MERIS, aboard the ESA-Envisat environmental satellite delivers data every 2-3 days. This increases the chances of encountering cloud free regions. Nevertheless, MERIS works at a spatial resolution of 300m (full resolution mode), which might not be sufficient to capture the details of highly fragmented landscapes. This is why the synergic use of these 2 sensors was explored in this paper.

An unmixing-based data fusion technique was used to generate images with the spatial resolution of Landsat TM and with the spectral (and eventually temporal) resolution provided by MERIS. More precisely, one Landsat TM and one MERIS full resolution image acquired in July 2003 over The Netherlands were fused using the linear spectral mixing model. First an unsupervised classification of the Landsat TM image was done to obtain the fractional coverages of the different land cover types present in each MERIS pixel. Next, the spectral signatures of each land cover type were retrieved by "inverting" the linear mixing model. This is, MERIS "endmembers" were obtained from the known fractional coverages of each pixel. After that, both the fused and the Landsat TM images were classified to produce maps of the 8 main land cover types over The Netherlands. These maps were subsequently validated using the Dutch land use spatial database (LGN5) as a reference. The paper concludes by describing the potentials and limitations of this multi-sensor approach with respect to the solely use of Landsat TM data.

1. INTRODUCTION

1.1 Remote sensing and land cover mapping

During the last few centuries, human land use activities have significantly contributed to the ecological degradation of our planet (Christensen et al., 1996). Anthropogenic land use activities are responsible for changes in atmospheric composition, the reduction of biodiversity and changes in climate at regional to global scales (Foley et al., 2005; Pielke Sr., 2005). In this context, accurate and up-to-date, land use and land cover information is essential to quantify the real magnitude of these changes and their potential impact on our future welfare.

Earth observation satellites are the main source of data for land cover mapping (Treitz and Rogan, 2004). Remote sensing data has, therefore, been progressively integrated into several national and international initiatives aiming at providing land use and land cover information (Gutman et al., 2004) .

Despite the success achieved by these land use land cover initiatives, our current understanding of the dynamics of land cover change is still far from complete (Foody, 2002). Innovative approaches using the wide variety of remotely sensed data as well as new ways of combining ancillary data are therefore required to further develop the science of land cover mapping (Woodcock and Ozdogan, 2004). In this study we illustrate the use of a relatively new approach for land cover mapping: data fusion.

1.2 Data fusion

Current Earth observation satellites provide data at a wide range of spatial, spectral and temporal resolutions. Because the data gather by different satellites is inherently complementary, it can be combined to generate datasets that have more information that each of the input datasets alone. This process of combining several kinds of imagery is known as data fusion (Park and Kang, 2004). Because fused satellite images contain more information than single-sensor imagery, they generally offer increased interpretation capabilities and more reliable results (Pohl and Van Genderen, 1998).

Data fusion approaches are commonly used to combine high spatial resolution imagery (necessary for an accurate description of the shapes, features and structures of the landscape) with either high spectral resolution imagery (useful to identify objects and to retrieve quantitative information) or with high temporal resolution imagery (essential to monitor vegetation phenophases and to produce maps over areas with persistent cloud coverage).

Several data fusion methods are available in literature. However, most of the them are operator or data type dependent, or they do not properly preserve the spectral

[1] Corresponding author: Raul.Zurita-Milla@wur.nl

information of the input images (Zhang, 2002; Zhang, 2004). An unmixing-based data fusion approach was selected for this study because it tries to preserve the spectral information of the low resolution image as much as possible (Zhukov et al., 1999; Minghelli-Roman et al., 2001). This should facilitate both the elaboration of accurate land cover maps and the use of fused images to monitor vegetation status.

1.3 Research objective

The aim of this study is to evaluate the synergic use of remotely sensed imagery to produce land cover maps over (highly) heterogeneous and frequently cloudy areas. More specifically, this paper illustrates the use of the linear mixing model to fuse Landsat TM and MERIS imagery. The Landsat TM sensor was selected because of its high spatial resolution whereas MERIS was selected because of its high spectral (15 narrow bands) and temporal resolutions (revisit time 2-3 days).

2. MATERIALS AND METHODS

2.1 Study area and remotely sensed data

The study area covers approximately 40 by 60 km of the central part of The Netherlands. This area was selected considering both the heterogeneity of the landscape and the availability of cloud free imagery over the area.

A Landsat TM-5 image acquired the 10th of July 2003 and a MERIS full resolution level 1b image acquired the 14th of July of 2003 were used to illustrate the proposed data fusion approach.

The Landsat TM image was geo-referenced to the Dutch national coordinate system (RD) using a cubic convolution method. A pixel size of 25m was selected during this re-projection so that the output resolution was equal to the resolution of the reference dataset (c.f. section 2.2).

Subsequently, the MERIS full resolution level 1b image was transformed from digital numbers (DN) to top of atmosphere radiances (L_{TOA}) using the metadata provided with the file. Then, the image was corrected for the so-called smile effect (Zurita-Milla et al., 2006). Finally, the image was also re-projected into the RD coordinate system using the ground control points provided with the MERIS file.

The last step of the pre-processing consisted of co-registering the input images. The Landsat TM image was used as a base image for the co-registration. Finally, a spatial subset containing the study area was masked out.

2.2 Reference data

The latest version of the Dutch land use database, LGN5, was used as a reference in this study. This geographical database has a grid structure with a cell size of 25 m and a detailed legend consisting of 39 classes. The LGN5 is based on multi-temporal classification of high resolution satellite data (mainly Landsat imagery from the years 2003 and 2004) and the integration of ancillary data (Hazeu, 2005).

The original 39 classes of the LGN5 were aggregated into the 8 main land cover classes of The Netherlands: grassland, arable land, deciduous forest, coniferous forest, water, built-up, bare soil (including sand dunes), and natural vegetation.

2.3 Unmixing-based fusion

2.3.1 Theoretical background

The selected data fusion approach makes use of the linear mixing model to combine different kinds of imagery. The linear mixing model assumes that there is no multiple scattering occurring within the different land cover types. In this way, the signal received per pixel is "just" a linear combination of the signals corresponding to "pure" land cover types (endmembers) weighted by their area within the pixel (Settle and Drake, 1993).

The linear mixing model is normally used to spectrally unmixing remotely sensed images (Ustin et al., 1993; Adams et al., 1995). In this case, the per-pixel fractional coverages are sought after determining the endmembers present in the image. However, when high spatial resolution imagery is available, the linear mixing model can also be used to spatially unmix low resolution images. This spatial unmixing is also known as unmixing-based data fusion. In this case, the high spatial resolution data is first used to compute the fractional coverages of the different classes present in each low resolution pixel. Then, the fractions are used to look for per-pixel class endmembers. Finally, the high spatial resolution image is used to get the spatial distribution of the classes that have been unmixed, so that the per-pixel endmembers can spatially allocated to create the fused image.

Contrarily to the linear spectral unmixing, which is solved per pixel and for all bands at once, the spatial unmixing problem is solved per band for a given neighbourhood.

Lets use, *i* to indicate a generic low resolution band, *nc* to specify the number of classes in which the high spatial resolution image was classified and *k* to represent size of a square neighbourhood; then, the unmixing-based data fusion approach can be written in a compact matrix-vector notation is as follows:

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

where: $\mathbf{L}^{i,k}$ is a k^2 column vector containing all the (radiance) values of the low resolution pixels in the neighbourhood *k* and the spectral band-*i*.

$\mathbf{P}^{k,nc}$ is a (k^2 x nc) matrix containing the proportions of each class for each low resolution pixels present in *k*.

$\mathbf{S}^{i,k,nc}$ is the per band *nc*-column vector of the endmembers present in a given neighbourhood.

$\mathbf{E}^{i}$ is a k^2 column vector of residual errors.

N is the number of bands of the low resolution image.

The general formulation of the unmixing-based data fusion approach (indirectly) implies the optimalisation of 2 parameters: the number of classes (*nc*) in which the high spatial resolution image is classified and the size of the neighbourhood (*k*) in which the Eq. 1 will be applied.

2.3.2 Case study

Figure 1 illustrates the general methodology used in this study. First the Landsat TM image was classified into 10, 20, 40, 60 and 80 classes using an unsupervised ISODATA classification rule. In this way we obtained the spatial patterns of the study area with different degrees of detail. Then, the classified image was aggregated to match the resolution of MERIS (i.e. 300 by 300m). During this spatial aggregation, the proportion of each

class present in each MERIS-like pixel was computed. Next, a sliding window, representing the neighbourhood size, was used to generate the class proportion matrices ($P^{k,nc}$). Four neighbourhood sizes were used in this study: 5 by 5, 9 by 9, 13 by 13 and 17 by 17.

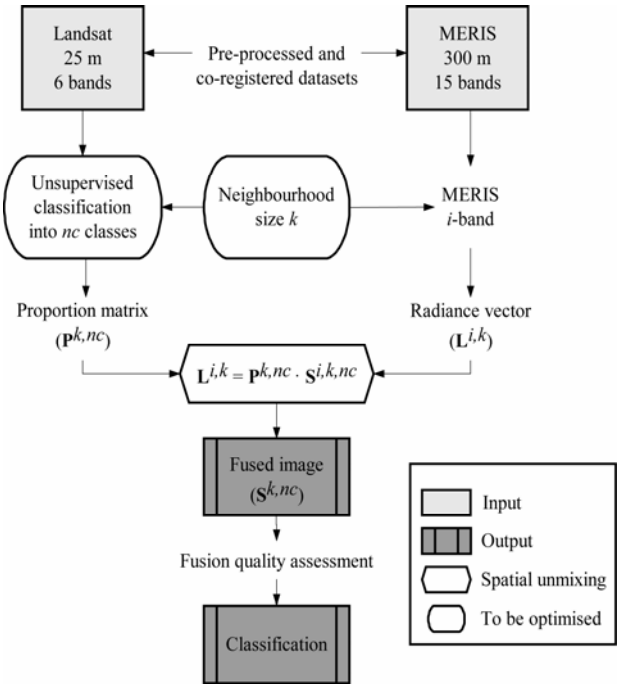

Figure 1. Scheme of the general methodology

On one hand, the neighbourhood size should be kept as small as possible so that the fused image is spectrally dynamic, but on the other hand, it should be as large as possible because it determines the number of equations available to solve the spatial unmixing model. Indeed, the unmixing-based model illustrated in Eq.1 is a system of k^2 equations with nc unknowns. This implies that the number of classes in a given neighbourhood cannot be larger than the number of available equations; otherwise the system of equations would be undetermined. However, the larger the size of the neighbourhood the less variability we will have in the fused image because each neighbourhood has unique set of endmembers. In other words, if we use the whole image as a neighbourhood, then only one set of endmembers can be calculated and this implies that a given class x will have the same spectral signature in the all the pixels of the image.

Using the same sliding window technique as for the preparation of the proportion matrices, we extracted the k^2 elements that conform the (per band) MERIS radiance vector ($\mathbf{L}^{i,k}$) that we needed to solve Eq. 1.

A constrained least-squares method was subsequently used to retrieve the per band vector of the endmembers present in the neighbourhood under study ($\mathbf{S}^{i,nc}$). A constrain method was needed because the endmembers should fulfil the following 2 conditions: i) all the radiance values have to be positive and ii) the radiance values have to be equal or smaller than the (per band) radiance saturation value of the MERIS sensor (www1).

The per band and per neighbourhood endmembers, $\mathbf{S}^{i,nc}$, were finally assigned to the corresponding classes (land ocations) of the central pixel of the neighbourhood.

This spatial unmixing operation was applied to all the low resolution pixels and repeated for all the bands and for all the combinations $nc \times k$, so that a series of fused images was created.

2.4 Data fusion quality

Assessing the quality of fused images is not straightforward because it depends on several factors like the difference in resolution of the input images or the type of landscape under consideration (Thomas and Wald, 2004). In this study, a quantitative quality assessment of the fused images was performed to find the combination of *nc* and *k* that produces the best image (from a spectral point of view).

Bearing in mind that any fused image should be as identical as possible to the original low resolution image once degraded to that resolution, we degraded the fused images to 300m using a mean filter. After this, we assessed the quality of the fused images by comparing the degraded images with the original MERIS images. In this study, the ERGAS index was used to do such a comparison because this index is independent of the units, the number of spectral bands and the resolution of the input images (Wald, 2002). The ERGAS index is computed as follows:

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

where: h is the resolution of the Landsat images.
l is the resolution of MERIS.
N is the number of spectral bands (B_i) involved in the fusion.
M_i is the mean value of each MERIS band. *RMSE* is the root mean square error computed between the (degraded) fused image and the original MERIS image (for the band B_i).

The quality of the fused images was also evaluated in terms of their classification accuracy. This was done to study the relationship between the ERGAS index and the classification accuracy.

2.5 Classification

All fused images were classified using a supervised maximum likelihood classification rule. The LGN5 was used to support the selection of the training samples for the 8 main land cover types of The Netherlands. The original Landsat TM image (6 bands) was also classified using the same training set as for the fused images.

In order to compare classification results of spectrally similar images, we designed a second experiment where the bands 5 and 7 of the Landsat TM image were omitted from the classification. This was done because MERIS does not collect information on this region of the electromagnetic spectrum. Additionally, a cubic convolution resampled MERIS image (300 to 25m) was also classified to evaluate the added value of the fusion process.

Similar to other studies (Zurita-Milla et al., 2006), the MERIS bands 1, 2, 11 and 15 were excluded from all the fused images before its classification. These bands are either very susceptible to atmospheric influences (bands 1 and 2) or they coincide with

absorption features (bands 11 and 15). Therefore, they do not provided relevant information for land cover mapping.
The confusion matrix and the kappa coefficient were used to compare the classification results.

3. RESULTS

3.1 Data fusion

Table 1 illustrates the ERGAS values for all the fused images. Notice that the neighbourhood size of 5 by 5 did not provided enough equations to solve some of the cases. The ERGAS index indicates that the quality of the fused image decreases with an increase in the number of classes used to classified the Landsat TM image and when increasing the neighbourhood size.
All the fused images yielded ERGAS values below 1, and this is consider to be a very good value since a threshold of 3 has been defined as the upper limit for a good fusion quality (Wald, 2002). These very low ERGAS values indicate that the unmixing-based data fusion succeeded in preserving the spectral information of the MERIS image. Indeed, the unmixing-based data fusion approach was selected because it takes the spatial details from the high resolution image and the spectral information from the low resolution image.

Table 1. ERGAS values

nc / *k*	10	20	40	60	80
5	0.687	0.556	--	--	--
9	0.844	0.780	0.681	0.612	0.530
13	0.909	0.858	0.797	0.742	0.698
17	0.942	0.902	0.854	0.816	0.787

k is the size of the neighbourhood; *nc* is the number of classes used to classify the Landsat TM image and '--' means that this combination could not be fused.

The best ERGAS value was obtained for combination *nc*=80 and *k*=5. Figure 2 shows an RGB color composite of a 25 x 25 subset of this fused image. The original Landsat TM and MERIS images are also provided in Figure 2.

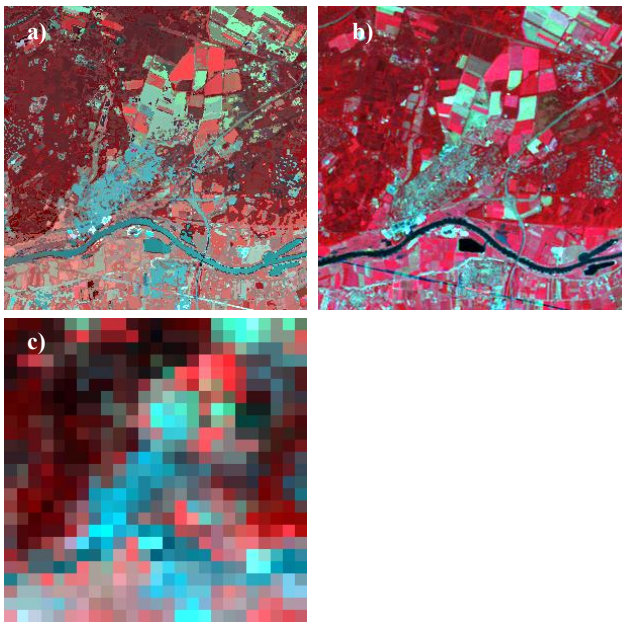

Figure 2. RGB color composite of the Fused (a), Landsat TM (b) and MERIS images over the study area

3.2 Classification

A supervised maximum likelihood classification rule was applied to: i) all the fused images, ii) the Landsat TM images (all bands and 4 bands cases) and iii) the original MERIS image resampled to 25m using cubic convolution. Table 2 summarises the best classification results.

Table 2. Classification results

	O.A [%]	K
Best fused image (*nc*=20; *k*=17)	57.54	0.476
Landsat TM 6 bands	63.32	0.550
Landsat TM bands 1 till 4	57.98	0.484
MERIS 25m cubic convolution	40.57	0.295

O.A. is the overall classification accuracy; K is the kappa coefficient

In all cases, the overall classification accuracies and the kappa coefficients of the Landsat TM images were better than for the fused images. However, the classification results of our second experiment, where the mid-infrared bands of Landsat TM were omitted, are very similar to the ones obtained for the best fused image (Table 2). This indicates that the mid-infrared bands, which are missing in MERIS, might play an important role in the final classification accuracy. Nevertheless, the temporal resolution of MERIS could help to overcome this issue.

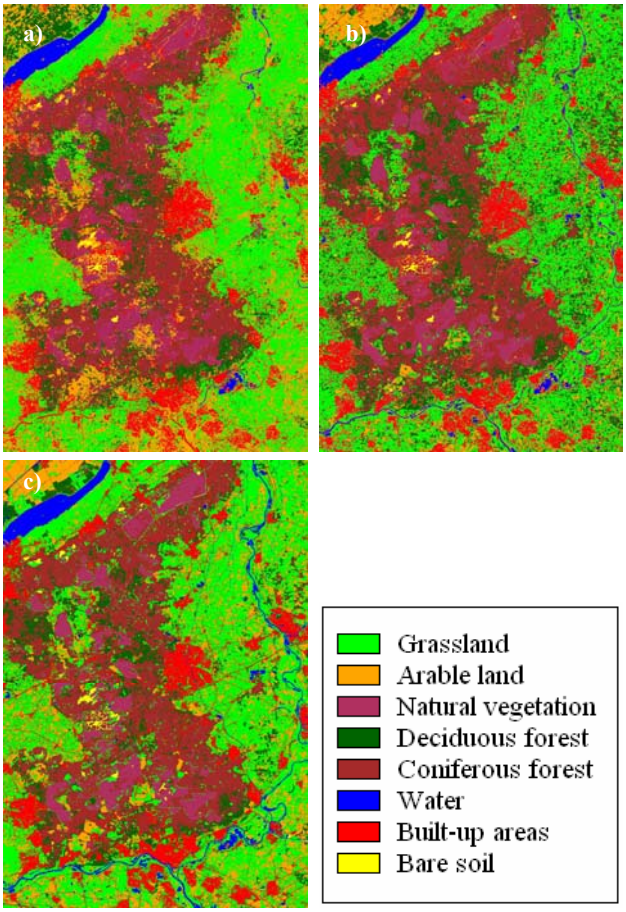

Figure 3. Classification results for the Fused (a), and the Landsat TM (b) images and the LGN reference dataset (c).

The best fused image performed much better than the cubic convolution resampling of the original MERIS image. This shows that the selected data fusion approach is very useful to downscale MERIS imagery. Figure 3 shows the best classification results obtained with the fused images together with the Landsat TM (all bands) and the LGN5.

The land cover map produced using the best fused image offers a good representation of the main landscape features despite its lower classification accuracy. Furthermore, the classification results obtained in this study suggest that using larger *nc* and *k* values might increase the final classification accuracy to exceed that of the Landsat TM.

In addition to this, the co-registration of the input images has to be as good as possible because potential errors in the geometry of the images have a great impact of the performance of this method. This impact can be minimised by using a large neighbourhood size. However, that would result in spatially averaged endmembers that would reduce the "dynamism" of the fused image.

Finally, special attention should be paid to the pixels that reach the upper or lower boundaries of the constrained least-squares method used to solve Eq. 1. This could be done by using regularisation methods (Golub et al., 2000).

4. CONCLUSIONS

In this paper we have used the linear mixing model to fuse a Landsat TM (high spatial resolution) image with a MERIS full resolution level 1b (high temporal and spectral resolutions) image. This is particularly interesting to produce land cover maps over spatially heterogeneous and cloudy areas.

The unmixing-based data fusion approach was evaluated using a quantitative indicator of data fusion quality: the ERGAS index. All the fused images had an ERGAS below 1, which is considered to be a very good fusion. This indicates that the selected approach succeeded in preserving the spectral information of the MERIS image.

The classification results of the Landsat images were better than ones obtained for the fused images. Nevertheless, the classification accuracy of the best fused image was very similar to the one of the Landsat TM 4 bands image. This could indicate the mid-infrared region might play an important role on the final classification accuracy. However, the fused image yielded better classification results that the cubic convolution resampling of the original MERIS image. This shows the potential of this fusion approach. Furthermore, the classification results obtained in this study suggest that using larger *nc* and *k* values might increase the final classification accuracy. Additional research should be devoted to explore the temporal dimension of MERIS because producing temporal series of fused images might also improve the classification accuracy.

Finally, the findings presented here should support the use of data fusion to downscale medium and low resolution imagery since the temporal resolution of provided by these sensors offers new possibilities to monitor vegetation status (e.g. in terms of FAPAR, LAI or chlorophyll content).

ACKNOWLEDGEMENTS

The contribution of R. Zurita-Milla is granted through the Dutch SRON GO programme (EO-061).

REFERENCES

Adams, J. B., D. E. Sabol, V. Kapos, R. Almeida, D. A. Roberts, M. O. Smith and A. R. Gillespie (1995). Classification of Multispectral Images Based on Fractions of Endmembers - Application to Land-Cover Change in the Brazilian Amazon. *Remote Sensing of Environment* 52(2): 137-154.

Christensen, N. L., A. M. Bartuska, J. H. Brown, S. Carpenter, C. D'Antonio, R. Francis, J. F. Franklin, J. A. MacMahon, R. F. Noss, D. J. Parsons, C. H. Peterson, M. G. Turner and R. G. Woodmansee (1996). The report of the ecological society of america committee on the scientific basis for ecosystem management. *Ecological Applications* 6(3): 665-691.

Foley, J. A., R. DeFries, G. P. Asner, C. Barford, G. Bonan, S. R. Carpenter, F. S. Chapin, M. T. Coe, G. C. Daily, H. K. Gibbs, J. H. Helkowski, T. Holloway, E. A. Howard, C. J. Kucharik, C. Monfreda, J. A. Patz, I. C. Prentice, N. Ramankutty and P. K. Snyder (2005). Global consequences of land use. *Science* 309(5734): 570-574.

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment* 80(1): 185-201.

Golub, G. H., P. C. Hansen and D. P. O'Leary (2000). Tikhonov Regularization and Total Least Squares. *SIAM Journal of Matrix Analysis and Applications* 21: 185-194.

Gutman, G., A. C. Janetos, C. O. Justice, E. F. Moran, J. F. Mustard, R. R. Rindfuss, D. Skole, B. L. Turner II and M. A. Cochrane, Eds. (2004). *Land Change Science*. Remote Sensing and Digital Image Processing. Dordrecht (The Netherlands), Kluwer Academic Publishers.

Hazeu, G. (2005). The Dutch Land Use Database LGN. [web page] http://www.lgn.nl/ [accessed 22nd February 2006].

Minghelli-Roman, A., M. Mangolini, M. Petit and L. Polidori (2001). Spatial resolution improvement of MeRIS images by fusion with TM images. *IEEE Transactions on Geoscience and Remote Sensing* 39(7): 1533-1536.

Park, J. H. and M. G. Kang (2004). Spatially adaptive multi-resolution multispectral image fusion. *International Journal of Remote Sensing* 25(23): 5491-5508.

Pielke Sr., R. A. (2005). Land use and climate change. *Science* 310(5754): 1625-1626.

Pohl, C. and J. L. Van Genderen (1998). Multisensor image fusion in remote sensing: Concepts, methods and applications. *International Journal of Remote Sensing* 19(5): 823-854.

Settle, J. J. and N. A. Drake (1993). Linear mixing and the estimation of ground cover proportions. *International Journal of Remote Sensing* 14(6): 1159-1177.

Thomas, C. and L. Wald (2004). Assessment of the quality of fused products. 24th EARSeL Symposium, New Strategies for European Remote Sensing, 25–27 May 2004, Dubrovnik, Croatia.

Treitz, P. and J. Rogan (2004). Remote sensing for mapping and monitoring land-cover and land-use change. *Progress in Planning* 61: 269-279.

Ustin, S. L., M. O. Smith and J. B. Adams (1993). Remote sensing of ecological processes: a strategy for developing and testing ecological models using spectral mixture analysis. *Scaling physicological processes: leaf to globe*. J. R. Ehleringer and C. B. Field. San Diego, CA, USA, Academic Press**:** 388.

Wald, L. (2002). *Data Fusion Definitions and Architectures: Fusion of Images of Different Spatial Resolutions*, Ecole des Mines Pres.

Woodcock, C. E. and M. Ozdogan (2004). Trends in land cover mapping and monitoring. *Land Change Science*. G. Gutman, A. C. Janetos, C. O. Justiceet al. Dordrecht, Kluwer Academic Publishers. 6**:** 367-377.

www1 MERIS handbook [web page] http://envisat.esa.int/dataproducts/meris/ [accessed 22nd February 2006].

Zhang, Y. (2002). Problems in the fusion of commercial high-resolution satelitte images as well as Landsat 7 images and initial solutions. *International Archives of Photogrammetry and Remote Sensing* 34.

Zhang, Y. (2004). Understanding image fusion. *Photogrammetric Engineering and Remote Sensing* 70(6): 657-661.

Zhukov, B., D. Oertel, F. Lanzl and G. Reinhackel (1999). Unmixing-based multisensor multiresolution image fusion. *Ieee Transactions on Geoscience and Remote Sensing* 37(3): 1212-1226.

Zurita-Milla, R., J. G. P. W. Clevers, M. E. Schaepman and M. Kneubuchler (2006). Effects of MERIS L1b radiometric calibration on regional land cover mapping and land products. *International Journal of Remote Sensing (In press)*.