Land use classification at meso-scale using remotely sensed data

Sonia Bouzidi, Fabien Lahoche, Isabelle Herlin, Volker Hochschild, Helmut Staudenrausch

LAND USE CLASSIFICATION AT MESO-SCALE USING REMOTELY SENSED DATA

S. Bouzidi[1], F. Lahoche[1], I. Herlin[1], V. Hochschild[2] and H. Staudenrausch[2]
[1] INRIA - Domaine de Voluceau - B.P. 105 - Rocquencourt - 78153 Le Chesnay Cedex, France
Sonia.Bouzidi@inria.fr, Fabien.Lahoche@inria.fr, Isabelle.Herlin@inria.fr
[2] Institute for Geography, University of Jena, D-07743 Jena, Germany
c5voho@geogr.uni-jena.de, c5sthe@geogr.uni-jena.de

ABSTRACT

In this paper we present a framework to generate a land cover classification from coarse spatial resolution remotely sensed data acquired by NOAA-AVHRR sensor. We define a model for the pixels' content and a process allowing to compute the individual proportions of the different land cover types for each pixel. The method is based on a linear mixture model of reflectances and exploits the good temporal frequency of NOAA acquisitions. The result provides a description in terms of land covers percentage within each NOAA pixel. A quality evaluation is performed on a test area for which high spatial resolution and temporal NOAA data are simultaneously available.

1 INTRODUCTION

Image classification is a widely used method for extracting information on surface land cover from remotely sensed images. The resulting cartography is helping decision makers in different research fields such as hydrology, agriculture and deforestation.
At local scale a classification of images at high spatial resolution (as Landsat TM or SPOT data) improved with ground truthing gives a good information about land cover. However, high spatial resolution data or ground truthing are not always available at least for important surfaces. On these large areas, the NOAA/AVHRR sensors provide daily acquisitions with an important coverage; but due to the size of a pixel ($1.1 \times 1.1km$), which is generally larger than a single field, each pixel includes several covers.
In order to extract information from such daily data, we propose a method to access the sub-pixel content. We define a process having as input a low spatial resolution temporal sequence of images (a NOAA sequence). It exploits the high temporal frequency of these acquisitions in order to estimate within each pixel the surface percentage of each land cover type. The approach is based on a mixture modeling of the pixel's response. This process will be called in this paper "NOAA unmixing". The result of this analysis is not a map of labels corresponding to different land covers (thematic classification) but a set of images (with the same size as the original NOAA image), giving the percentage of each land cover type for each pixel.
The mixture model is described in section 2. The inversion of this model to estimate land cover proportions is explained in section 3. Lastly results are presented and evaluated in section 4.

2 THE MIXTURE MODEL

The basic physical assumption underlying the linear mixture model is that the different occupations present in a pixel contribute independently to its reflectance. The reflectance of that pixel can be considered as a sum of the individual land covers reflectances weighted by their area proportions. This model was used in different studies to un-mix pixels (Quarmby et al., 1992, Settle and Drake, 1993). The quoted authors consider the pixel's response in different spectral bands of satellite images and resolve a system of nb linear equations where nb corresponds to the number of bands. This approach restricts the number of studied land cover types. In fact, only the case for which we have fewer interesting components than the number of acquisition bands can be evaluated. In other words, if we are dealing with NOAA/AVHRR data, only three land cover types can be studied because of the only three available channels measuring the reflectance in the visible and near infrared bands. To avoid this restriction, we propose to exploit the temporal information offered by the NOAA/AVHRR sensor. We make use of the whole sequence with a daily frequency and covering the vegetative cycle. The number of linear equations is then important and we can study all land covers of a scene. At each date t we consider the linear relation between the pixel's reflectance and the individual reflectances of its components in the visible (channel 1) and near infrared (channel 2) channels:

$$R_i^k(t) = \sum_{j=1}^{N_c} \rho_{ij} \mathcal{R}_j^k(t) \qquad (k = 1, 2) \tag{1}$$

where:

- N_c is the number of land covers,
- $R_i^k(t)$ the reflectance of a NOAA pixel i in channel k (k = 1 or 2), at date t,
- $\mathcal{R}_j^k(t)$ the individual reflectance of the land cover type j in channel k (k = 1 or 2), at date t,
- ρ_{ij} the proportion of the land cover j within a NOAA pixel i.

We first use this model within a learning process to estimate the individual temporal reflectances $\mathcal{R}_j^k(t)$ for each land cover type in the visible and near infrared channels. For this purpose, we consider a small learning area for which ground truthing and high spatial resolution remote sensing data are simultaneously available. A precise land cover thematic classification image is obtained for this area from the remote sensing and the ground truthing (Flügel, 1998). Then, after geometrical registration, we superpose the Landsat classification image and the NOAA images. Consequently, for each NOAA pixel, its composition in terms of percentage of land cover can be obtained by directly counting pixels on the Landsat classification image. We then obtain the ρ_{ij} values for this learning area. At each date t, we consider the linear model (1) for several pixels i; knowing the ρ_{ij}, the inversion of the obtained linear system allows the estimation of individual values $\mathcal{R}_j^k$ for this date. By repeating this process for all the dates, the temporal profiles $\mathcal{R}_j^k(t)$ are computed. This method and its results are detailed in (Bouzidi et al., 1997a, Bouzidi et al., 1997b).
In this paper, we use the hypothesis that these temporal profiles are spatially constant. At each pixel i, we consider the linear model (1) for several dates t; knowing the $R_i^k(t)$ from the NOAA image and the individual temporal profiles $\mathcal{R}_j^k(t)$ from the learning process, the inversion of the obtained linear system allows the estimation of the proportions ρ_{ij} of each land cover j within the pixel. This point is described in the next section.

3 LAND COVER PROPORTIONS ESTIMATION

Our task is to estimate for each pixel i in the NOAA image the proportion ρ_{ij} of land cover j. We then consider, for this given pixel i, its reflectance in the visible and near infra-red channels for a number of dates. These dates are selected to avoid bad quality acquisitions and only images containing at least 80% of free clouds pixels are considered. Moreover, the acquisitions are constraint to cover an important part of the vegetative cycle in order to discriminate at best the different land covers. An over-determined system ($\mathcal{S}$) of M linear equations (M denotes the number of dates) and N_c unknown variables ($\rho_{ij}, j = 1, \ldots, N_c$) is then obtained:

$$(\mathcal{S}) \qquad R_i^k(t) = \sum_{j=1}^{N_c} \rho_{ij} \mathcal{R}_j^k(t) \quad (k = 1, 2) \quad (t = 1, \ \ldots, M) \tag{2}$$

with the following constraints:

$$\sum_{j=1}^{N_c} \rho_{ij} \quad = \quad 1 \tag{3}$$

$$0 \leq \quad \rho_{ij} \quad \leq 1 \qquad (j = 1, \ \ldots, N_c) \tag{4}$$

To solve these equations, we formulate the problem as an optimization problem. We aim to minimize, for a given pixel i, the difference between the different observed reflectance and that estimated by the linear model. This leads to the minimization of a set of functions for each pixel i:

$$|R_i^k(t) - \sum_{j=1}^{N_c} \rho_{ij} \mathcal{R}_j^k(t)| \qquad (k = 1, 2) \quad (t = 1, \ \ldots, M) \tag{5}$$

under the constraints (3) and (4). This minimization is performed using an algorithm based on sequential quadratic programming described in (Zhou and Tits, 1993). This process, that we call “NOAA unmixing process” outputs, for each pixel, N_c values corresponding to the proportions: percentage of each land cover type. This process allows the generation of N_c images, each one having the size of the original NOAA image and giving for the studied land cover type its proportion within each pixel.

4 RESULTS AND VALIDATION

The test area chosen to evaluate the results is the Mkomazi catchment in Kwazulu-Natal (South Africa). The data available for this site are obtained within the context of the European INCO-PED IWRMS project (*Integrated Water Resources Management System*). The available images are the following:

- a temporal sequence of NOAA images covering an important part of southern Africa,
- a Landsat TM image covering the main part of the catchment,
- a land cover classification computed from the Landsat TM image and enhanced by ground truthing. It is described in table 1.

Name of land cover	Relative surface of land cover in the thematic classification image
indeterminate	8.69%
water	18.36%
urban	6.91%
field	12.70%
bare soil	4.47%
grass	26.79%
bushland	11.40%
woodland	1.16%
forest	9.52%

Table 1: Characteristics of the Mkomazi Landsat thematic classification image.

The process of validation is explained in figure 1. We consider a test zone covered simultaneously by the Landsat image and the NOAA acquisitions. We first establish the pixels' composition using the classification image obtained from Landsat data (only ground pixels are considered, and a mask is applied on pixels containing water). In fact, each NOAA pixel contains approximately $37 \times\ 37$ Landsat pixels which can be localized by the geometrical registration. From this set of Landsat pixels and from the classification image we compute the composition of the NOAA pixels in terms of the different land cover percentages. The obtained compositions serve as reference ones and are compared to the pixels proportions obtained after unmixing process of the NOAA data. This can be done for each land cover type. We present in figure 2 and figure 3 results obtained respectively for bare soil and grass. The lighter a pixel is, the greater the value of proportion or composition (in%) is.
In order to evaluate the difference between the pixels' compositions from the Landsat classification and the proportions obtained after the unmixing process of the NOAA data, we compute an euclidian distance between both results. So, for each land cover type j, a distance value is computed using the equation (6):

$$d_j = \sqrt{\frac{\sum_{i=1}^{n}(P_i^j - C_i^j)^2}{n}} \tag{6}$$

where :

- P_i^j denotes the proportion of land cover j in pixel i obtained after the unmixing process of the NOAA pixels,
- C_i^j denotes the composition of NOAA pixel i in land cover j obtained after the counting process on the thematic Landsat classification image,
- n is the number of NOAA pixels used for the evaluation.

The table 2 gives the euclidian distance values obtained for the different land cover types. The highest error values

Land cover	Euclidian Distance
urban	0.07
field	0.16
bare soil	0.05
grass	0.16
bushland	0.09
forest	0.11

Table 2: Euclidian distances for each land cover.

(measured by the euclidian distance) are observed for land covers occupying important surfaces of the test area (grass

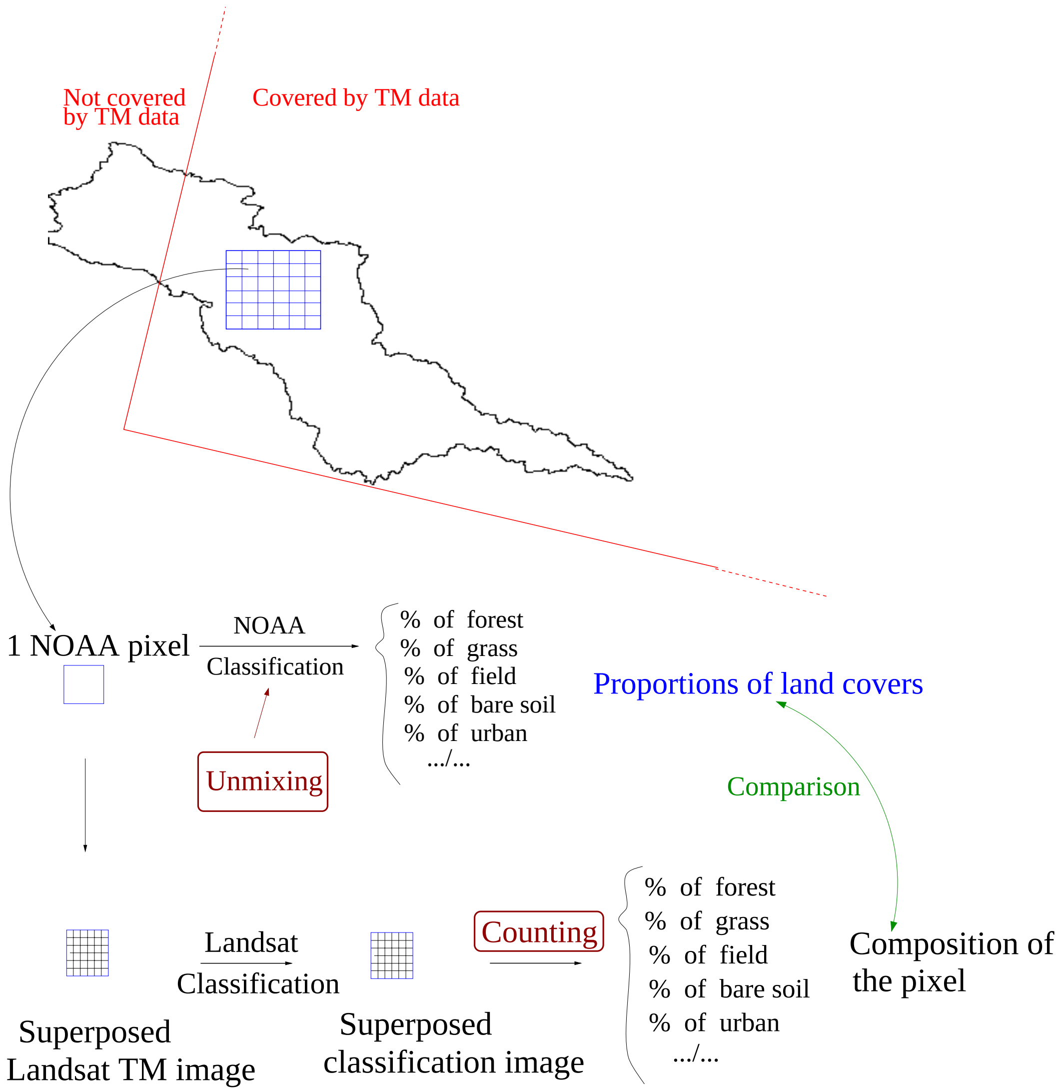

Figure 1: Results validation scheme.

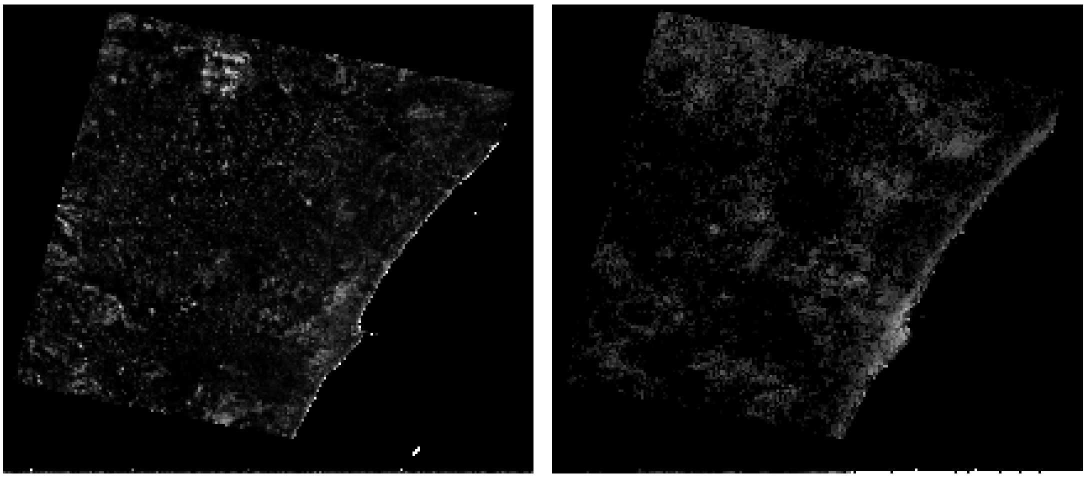

Figure 2: Bare soil. On the left, composition computed from the classification image; on the right, proportion obtained after the unmixing process of the NOAA pixels.

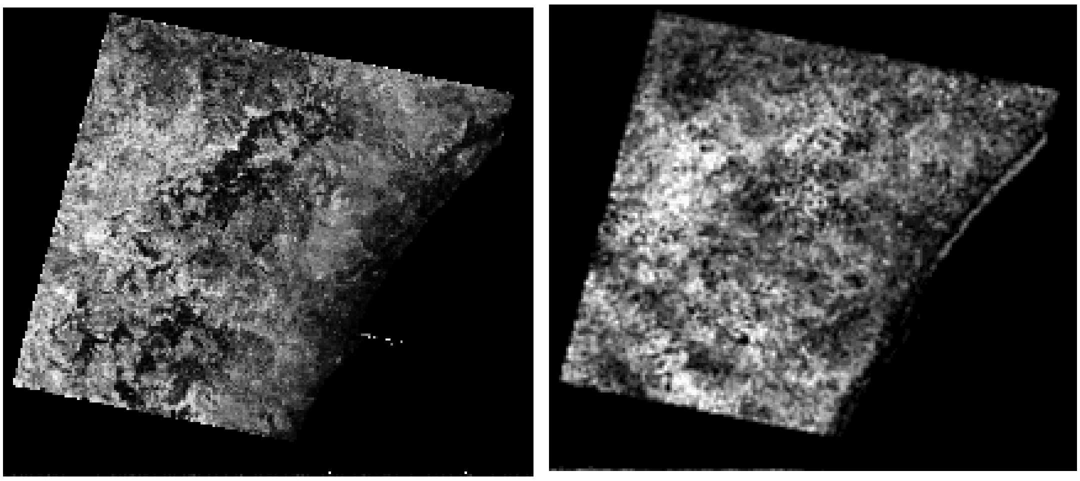

Figure 3: Grass. On the left, composition computed from the classification image; on the right, proportion obtained after the unmixing process of the NOAA pixels.

26.79%, field 12.7%). These euclidian distances are computed from exact numerical values of composition and proportion and are not so easy to interpret even if we can see that fields and grass seem to be the land cover types difficult to analyse. We then propose to consider a qualitative criterion for comparison of results. So, for each land cover type, for each process (counting process to compute composition, unmixing process to compute proportion), and for each pixel, we define a label according to table 3. Consequently, for each land cover, we generate two new images of labels values (1...5)

Label	Range of composition	Range of proportion
1	$0\% \leq$ composition $< 20\%$	$0\% \leq$ proportion $< 20\%$
2	$20\% \leq$ composition $< 40\%$	$20\% \leq$ proportion $< 40\%$
3	$40\% \leq$ composition $< 60\%$	$40\% \leq$ proportion $< 60\%$
4	$60\% \leq$ composition $< 80\%$	$60\% \leq$ proportion $< 80\%$
5	$80\% \leq$ composition $\leq 100\%$	$80\% \leq$ proportion $\leq 100\%$

Table 3: Labels and their corresponding values for proportion and composition.

corresponding to the two process (counting for composition, unmixing for proportion) as it is illustrated by figures 4 et 5, respectively for bare soil and grass. Then, we compute for each land cover, a (5×5) matrix L, where each value $l(c, u)$ $(c, u = (1...5))$ represents the quantity of pixels belonging to label c for composition (computed after the counting process of the Landsat classification image) and retrieved with a label u for the proportion (computed after the unmixing process of the NOAA data). The normalized trace of this matrix is then a global measure of accuracy for the unmixing process, compared with the classification process of the Landsat images. The table 4 displays the results obtained for the different land covers. High rates of label recognition are obtained for the land covers urban, forest and bare soil. However, the

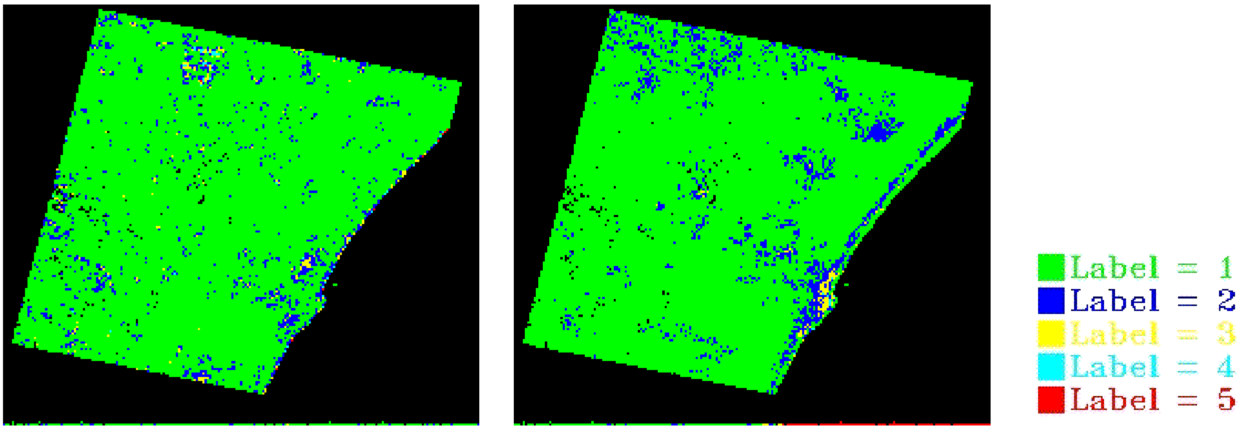

Figure 4: Bare soil. On the left, labels of composition; on the right, labels of proportion.

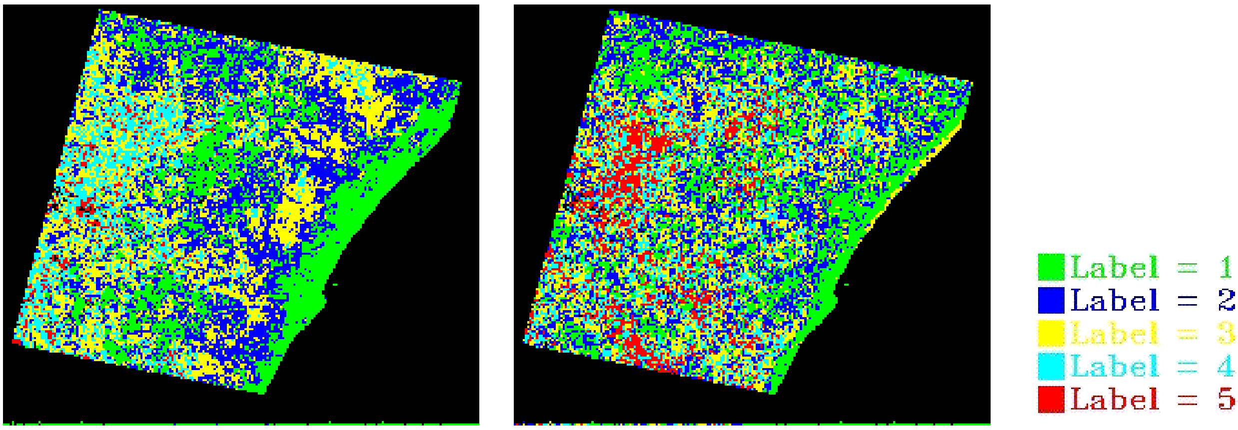

Figure 5: Grass. On the left, labels of composition; on the right, labels of proportion.

Land cover	Percentage of good retrieved labels
urban	78%
field	43%
bare soil	86%
grass	32%
bushland	62%
forest	71%

Table 4: Percentage of identical labels for each land cover type.

errors observed for the land covers grass and field are important. We are currently wondering about the kind of qualitative labels to create. We also suppose a confusion between these two classes because of their similar visible and near infra-red behaviors which have been used for the unmixing process of NOAA data. This is illustrated on figure 6 by displaying the NDVI (Normalized Derived Vegetation Index) curves estimated from the visible and near infra-red NOAA/AVHRR channels ($NDVI = (nir - vis)/(nir + vis)$): the curves are quite similar and that surely explains the difficulty to discriminate grass and fields in the visible and near infra-red channels. A future task is then to compute a differentiation

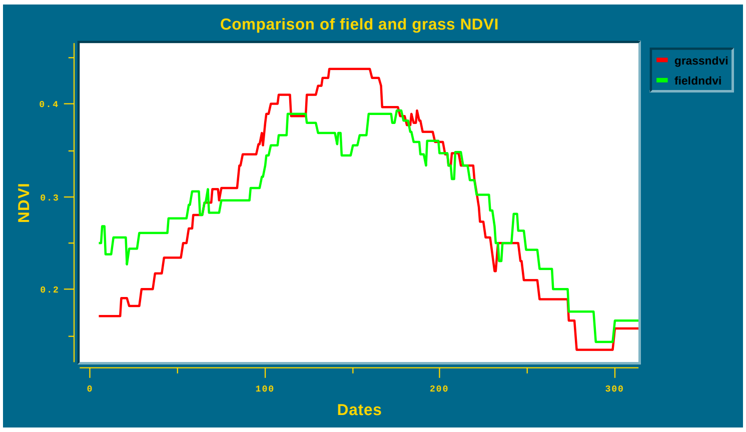

Figure 6: Temporal profiles of NDVI, for field and grass.

criterion between these two land covers, in order to minimize this confusion. To do this, we propose to make use of the thermal channels and to compute proportions by considering the physical based mixture model that has been discussed in (Lahoche et al., 2000).

5 CONCLUSION

We describe a method for land use classification at meso-scale using coarse spatial resolution remotely sensed data acquired by the NOAA/AVHRR sensors. We define a model for the pixels' content and a process allowing to compute the individual proportions of the land covers for each pixel. This process exploits the temporal information of the NOAA data and is based on inverting a linear mixture model of reflectance, we call it the "unmixing process". The result provides a description in terms of land covers percentage within each NOAA pixel. Results evaluation is carried out on a test site and show that the method gives a global idea about land cover proportion inside the NOAA pixels. If this method seems to be promising, it must still be improved: define which types of land cover can be discriminated, and with which precision level. On another hand, for the inversion of the linear mixture model, we have considered a determinist resolution. A new formulation based on a probabilistic framework is under investigation, in order to estimate proportion values and also the associated likelihood.

6 ACKNOWLEDGMENTS

This research is made within the context of the European INCO-PED IWRMS project (*Integrated Water Resources Management System* http://www.iwrms.uni-jena.de/) funded by European Commission under the contract ERBIC18CT97044.

REFERENCES

Bouzidi, S., Berroir, J.-P. and Herlin, I., 1997a. Simultaneous use of SPOT and NOAA/AVHRR data for vegetation monitoring. In: proceedings of the 10th Scandinavian Conference on Image Analysis.

Bouzidi, S., Berroir, J.-P. and Herlin, I., 1997b. Subpixel mixture modeling applied for vegetation monitoring. In: proceedings of the International Symposium on Environmental Software Systems.

Flügel, W., 1998. Classification of the Mkomazi river catchment. Technical report, Institute of Geography, University of Jena, Germany.

Lahoche, F., Berroir, J., Bouzidi, S. and Herlin, I., 2000. Fusion of Landsat TM and NOAA/AVHRR data for generating evapotranspiration maps in a semi-arid catchment. In: Proceedings of 3^{rd} International Conference "Fusion of Earth Data", Sophia-Antipolis, France.

Quarmby, N., Townshend, J., Settle, J., White, K., Milnes, M., Hindle, T. and Silleos, N., 1992. Linear mixture modelling applied to AVHRR data for crop area estimation. International journal of remote sensing.

Settle, J. J. and Drake, N. A., 1993. Linear mixing and the estimation of ground cover proportions. International journal of Remote Sensing pp. 1159–1177.

Zhou, J. and Tits, A., 1993. Nonmonotone line search for minimax problems. Journal of optimization theory and application 76, pp. 455–476.