

A multiscale multitemporal land cover classification method using a Bayesian approach

A. Robin^{a,b} and S. Mascle-Le Hégarat^a and L. Moisan^b

^aCETP, 10-12 avenue de l'Europe, Vélizy, FRANCE

^bMAP5, 45 rue des Saints-pères, Paris, FRANCE

ABSTRACT

As vegetation time evolution is one of the most relevant information to discriminate the different land cover types, land cover classification requires both temporal and spatial information. Due to the physical properties of remote sensors, this temporal information can only be derived from coarse resolution sensors such as MERIS ($300 \times 300m^2$ pixel size) or SPOT/VGT ($1km^2$ pixel size). In this paper, we propose to use jointly high and coarse spatial resolution to perform an efficient high resolution land cover classification. The method is based on Bayesian theory and on the linear mixture model permitting, through a simulated annealing algorithm, to perform a high resolution classification from a coarse resolution time series.

Keywords: multiscale classification, Bayes theory, mixture model, multitemporal series, land cover.

1. INTRODUCTION

In the last decades, the use of remote sensing data has proved to be efficient for monitoring the Earth surface. In particular, land cover maps provide essential information not only for the analysis of global or local changes but also for studies on geosphere-biosphere-atmosphere interactions that depend on reliable and unambiguous definition of the existing terrestrial vegetation. Indeed, vegetation characteristics have a significant impact on surface processes involved in water or energy exchanges. Land-cover maps are increasingly used to define policies for environmental interventions and, hence, they contribute to reduce the risks from natural disasters (floods, forest fires, etc.). Given the large size of satellite data, automatic classification techniques are required to generate land cover maps.

As time evolution is one of the most discriminating criteria for vegetation, the classification of a scene in terms of land cover types requires high temporal frequency information. Nowadays, mainly two kinds of imaging sensors are available for these applications: sensors with a high spatial resolution (e.g. SPOT/HRV, 1 pixel for $20m \times 20m$) but a two-monthly temporal acquisition frequency, and sensors with a medium or coarse spatial resolution (e.g. MERIS, 1 pixel for $300m \times 300m$, or SPOT-VGT, 1 pixel for $1km^2$) but daily or so temporal acquisition frequency. As we must take into account the cost of high resolution data and the fact the studied scene may be cloudy on the acquisition date, vegetation monitoring requires the use of medium or coarse resolution time series. Moreover, these sensors provide spectral information dedicated to vegetation applications. One of the key challenge for automatic land cover classification is hence the combination of information from different resolutions to benefit both a high discrimination between land cover types and an accurate spatial information. Therefore, we suggest to exploit jointly high and coarse spatial resolution data.

In the remote sensing literature, several methodologies have been reported for automatic classification. Many of them are based on Bayesian theory as it offers a theoretically robust foundation for the classification of remotely sensed data. It has been widely used in the image processing analysis through approaches such as Maximum A Posteriori (MAP) estimation, Iterated Conditional Mode (ICM) or Expectation-Maximization (EM). Markov and Gibbs random fields have shown to be appropriated modelling to take into account spatial information from images and consequently, they have been widely used in classification,¹⁻³ image restoration^{4,5} and segmentation.⁶ However, the constraints imposed by the data dedicated to vegetation monitoring are not directly compatible with the use of such methods. Indeed, as coarse resolution images are composed of pixels representing a surface on Earth of up to $1km^2$, two neighbour pixels are not likely to represent the same type of cover. On the opposite, data fusion and image reconstruction have been extensively investigated in the past years. The methods proposed are mainly based on statistical,^{7,8} wavelet,^{9,10} morphology,¹¹ evidence¹² or neural-networks theory.¹³ From

there, a possible approach, rather time consuming, would consist in deriving multitemporal high spatial resolution (HR) series from coarse resolution (CR) one using a data fusion method and then performing classification from this reconstructed HR series.

In this paper, we focus on the problem of land cover classification from both multitemporal coarse resolution images and one or few HR images. The approach we propose is based on Bayesian theory and assumes that the geometry of the observed scene is stationary during the considered period. This assumption enables to consider that spatial information can be extracted at any time of the period, *e.g.* processing a segmentation of at least one HR image. Such an assumption is realistic for a time period as one agricultural year, particularly if rather an over-segmentation is performed on HR images, enabling to distinguish all parcels. Notice that in case a cadastre is available, it could be used directly and dispense from the use of HR images. However, this assumption of stationarity can be restrictive if changes occur along the studied period. Parallel studies focus on the problem of land cover change detection¹⁴ and may be used jointly to classification for a complete vegetation monitoring automatic process. Assuming a segmentation image is available, the classification task refers to a labeling problem using coarse resolution radiometric information. For this purpose, the linear mixture model¹⁵ is used, first, to describe the CR time series according to the segment mixture within each pixel and, then, the Bayes rule allows to define the cost function. The minimization of this cost function leads to a HR label map, where each label stands for a specific type of land cover. The quality of this land cover map is then valued using simulated and actual data from the Danube region (Rumania). Performances are measured comparing the result obtained from one HR image and a CR time series to the one obtained from a HR times series.

The paper is organized into 5 sections. The next section provides a formulation of the considered problem and introduce the principle of the proposed approach. The labeling method we have derived is then detailed Section 3, leading to the definition of a cost function to minimize. For that purpose, a simulated annealing technique is used which algorithm is also described. The data sets used in the experiments and the results obtained are reported in Section 4 before Section 5 gathers the main lines of the paper as a conclusion.

2. PROBLEM FORMULATION

As time evolution of radiometric information is the most discriminating criteria, land cover classification mainly relies on the use of multitemporal time series. As previously mentionned, the high repetitiveness sensor constraint implies the use of coarse resolution time series. Considering that the radiometric observed value depends on the different types of land cover represented within the field of view, the classification problem consists in the estimation of a non observed realization of the random label field from the observation of noisy data. In this work, we use a Bayesian framework and the Maximum A Posteriori (MAP) criterion. Using this criterion, the cost function, which represents the importance of the error obtained while considering a given realization of the label field instead of the *accurate* one, is a 0 – 1 function: all label field realizations considered are penalized in the same way but the accurate one. Using the MAP criterion, the researched classification is the label field that maximizes the *a posteriori* probability of a label configuration knowing the time series observed. This posterior probability can be expressed, using the Bayes rule, according to the label conditional probability of observing the data and the prior probability of a label configuration. As far as the *prior* probability is concerned, a Markovian formalism is usually adopted for the label field as it enables to take spatial information into account. Modeling pixel interaction through the definition of local energies, this formalism provide a powerful tool to reduce the space of solutions and access directly to the most probable configurations.⁵ Using the Potts model for the potential definition, configurations where neighbour pixels share a same label are more probable.

The main difficulty with medium or coarse resolution images such as MERIS or SPOT/VGT is the size of the field of view relatively to the size of *objects* to classify (numerous parcels are entirely contained in one CR pixel). In the case of agricultural areas, most spatial information is the geometry of the scene that can be considered as constant during an agricultural year. Moreover, dealing with mixed CR pixels, CR time series constrains the classification solution only in terms of percentage of each class within a CR pixel, and the spatial interaction modelling using a Potts model constrains the solution only in terms of segment border length. Several configurations of the classes spatial distribution within CR pixels may hence be equivalent as these two

constraints are not sufficient to obtain a unique solution. Consequently, assuming at least one HR image is available during an agricultural year, we propose to decompose the multi-resolution classification problem into two successive problems: the segmentation of the scene and its labeling, which aim is to map to each segment a label representing the land cover type. Spatial information is taken into account during the segmentation process in addition to HR image radiometric information in order to provide a unique segment map. The segmentation problem we consider (segmentation of a HR image) has been widely investigated in the past decades, and many competitive techniques have been proposed. The main difficulty is to be able to perform good segmentation of both large and small objects. In this study, we refer to the Mumford and Shah variational method¹⁶ for segmentation as it has shown good performance on remote sensing data. Let Ω denote the HR image domain, and $u : \Omega \rightarrow \mathbb{R}$ a HR image. The segmentation is represented as a function $s : \Omega \rightarrow S$ that fragment, for a given number of segments or edges length, the image u in homogeneous segments. The set $S = \{1, 2, \dots, |S|\}$ denotes the set of obtained regions indexed from 1 to $|S|$. From there, the classification task consists in labeling each segment of S according to the type of vegetation it represents using radiometric coarse resolution information from the time series. This is the purpose of the next section. A scheme representing the outlines of the proposed approach is presented Figure 2.

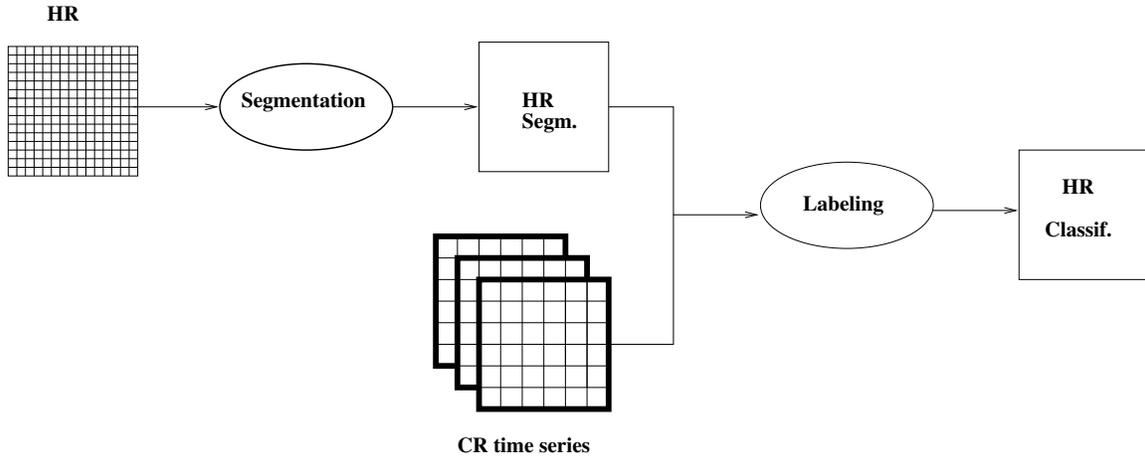


Figure 1. Outlines of the proposed approach.

3. LABELING METHOD

In this section, we describe how we use multitemporal radiometric information to discriminate land cover types and assign a label to each segment of the segmentation.

3.1. Probabilistic image model

We dispose of a high resolution segmentation and of a CR time series v . Assume the number of different land cover types represented on the studied scene is known. Each of them is described by a label and let $L = \{1, \dots, |L|\}$ be the set of labels (*e.g.* 1 for corn, 2 for forest, etc.). Let Ω and Ω' respectively denote high and coarse resolution image domains and $T = \{1, \dots, |T|\}$ the acquisition dates set. A probabilistic formalism is adopted to describe the considered images. In this context, a CR time series $v = (v^1, \dots, v^{|T|})$ is a realization of a random field $\tilde{v} = (\tilde{v}^1, \dots, \tilde{v}^{|T|})$ where, for each date t , the image observed at date t $v^t : \Omega' \rightarrow \mathbb{R}$ is a realization of a random field $\tilde{v}^t = (\tilde{v}^t(y))_{y \in \Omega'}$. Let $u = (u^1, \dots, u^{|T|})$ denote a non-observed HR time series composed, for each date t , of an image $u^t : \Omega \rightarrow \mathbb{R}$, which is a realization of a random field \tilde{u} on the domain Ω . Consider that the land cover classification $l : S \rightarrow L$ is a realization of a label random field \tilde{l} . As the radiometric observed value depends on the different types of land cover represented within the field of view, the classification problem consists in the estimation of the non observed realization l of the field \tilde{l} from the observation of noisy data v , realization of \tilde{v} . The conditional probability of observing v when the label map is l , $\mathbb{P}(\tilde{v} = v \mid \tilde{l} = l)$, describes the observing and

data acquisition process. The definition of the scale-change model Section 3.2 enables, under some hypotheses, to determine Section 3.3 the probability of CR observation conditionally to a label configuration. Finally, the cost function corresponding to the MAP is given 3.4, and the algorithm used is described Section 3.5.

3.2. Linear mixture model

Vegetation monitoring requires high temporal frequency information to discriminate vegetation types by their phenological evolutions. It explains the use of coarse resolution data such as SPOT/VGT or MERIS. As a counterpart of this temporal richness, their spatial resolution is poor and, usually, the ground area of a SPOT/VGT or MERIS pixel is a mixture of several land cover types. Observing these *mixed* pixels, the classification purpose implies to retrieve the different types of cover represented within the field of view.

The *linear mixture model* yields that the esperance of the observation measurement performed over a mixed pixel is the weighted average of the observation measurement that could have been performed over *pure pixel* representing each land cover type. It is widely used for extracting compositional information from remotely sensed images containing mainly mixed pixels, as well to estimate the relative area of agricultural crops^{17,18} as for class feature estimation.^{19,20}

This linear model is obtained assuming a CR image v^t corresponds to the average of a non-observed HR image u^t . Therefore the physical measurements used should be scale linear, even though some studies have shown that a linear combination of vegetation index such as NDVI implies only minor inaccuracies.²¹ For vegetation monitoring, foliar or vegetation index are usually used as they provide a means to extract relevant information for land cover discrimination. In this study, we use either directly luminance or cover fraction measurements as it is, by definition, scale linear. Namely, the cover fraction indicates the fraction which represents a surface covered with vegetation within a pixel. From the scale linear hypothesis, the measurement observed in a CR pixel y at a given date t can then be expressed as

$$v^t(y) = \frac{1}{N} \sum_{i \in L} \sum_{\substack{x \in W_y \\ l_x = i}} u^t(x), \quad (1)$$

where W_y is the set of the high resolution pixels $x \in \Omega$ contained in the coarse resolution pixel y , l_x denotes the label of the pixel x and $N = |\Omega|/|\Omega'|$ is the number of high resolution pixels contained in a coarse resolution pixel (it is also called the resolution ratio). For each pixel y of the CR domain Ω' , let $\alpha_i(y)$ denote the relative area of the CR pixel y covered by the land cover type labeled i . That is, writing $N_i(y)$ the number of HR pixels labeled i within the CR pixel y , the proportion of label i within y is $\alpha_i(y) = N_i(y)/N$ and, hence, $\sum_{i \in L} \alpha_i(y) = 1$. The measurement (1) observed in a CR pixel y can then be expressed as

$$v^t(y) = \sum_{i \in L} \alpha_i(y) \sum_{\substack{x \in W_y \\ l_x = i}} \frac{u^t(x)}{N_i(y)}. \quad (2)$$

3.3. Label conditional observation probability

From the linear mixture model, the observed measurement within a CR pixel depends on the occupation rate of each label and on the HR pixel value. In the following, we assume the random field \tilde{u} is Gaussian conditionally to the inherent labeling. The type of cover i is characterized by a Gaussian distribution of mean μ_i^t and variance (σ_i^t) . For a given date t and a CR pixel y , the linear mixture model yields that the CR observation $v^t(y)$ is a realization of the random variable $\tilde{v}^t(y)$ which represents a mixture of Gaussian laws. A mixture of Gaussian laws being Gaussian, CR pixel values follow a Gaussian law of mean and variance depending on the class mixture within the CR pixel. Let us determine these parameters according to class mixture and considered segmentation.

For each CR pixel y , let $\beta_k(y)$ denote the relative area of a segment k within y . The proportion of label i in a pixel y is equal to the sum of the proportions of all segments k labeled by i within y , *i.e*

$$\alpha_i(y) = \sum_{\substack{k \in S \\ l_k = i}} \beta_k(y), \quad (3)$$

where S denotes the set of segments and l_k represents the label of the segment k . For all date t and pixel y , the mean

$$\mu_{y,l}^t = \sum_{i \in L} \alpha_i(y) \mu_i^t = \sum_{i \in L} \sum_{\substack{k \in S \\ l_k = i}} \beta_k(y) \mu_i^t. \quad (4)$$

The variance is computed from (2) using the variance expression for Gaussian mixture and hence, for each date t and pixel y , it writes

$$(\sigma_{y,l}^t)^2 = \sum_{i \in L} \alpha_i^2(y) N_i(y) \left(\frac{1}{N_i(y)} \right)^2 (\sigma_i^t)^2 \quad (5)$$

$$= \frac{1}{N} \sum_{i \in L} \alpha_i(y) (\sigma_i^t)^2 \quad (6)$$

$$= \frac{1}{N} \sum_{i \in L} \sum_{\substack{k \in S \\ l_k = i}} \beta_k(y) (\sigma_i^t)^2. \quad (7)$$

Consequently, assuming label conditional independence of observations $\tilde{v}^t | \tilde{l}$ from each date, the probability of observing a series v in the CR pixel y , conditionally to the label field l ,

$$\mathbb{P}(v(y) | \tilde{l} = l) = \prod_{t \in T} \frac{1}{\sigma_{y,l}^t \sqrt{2\pi}} \exp \left(- \frac{(v^t(y) - \mu_{y,l}^t)^2}{2\sigma_{y,l}^2} \right), \quad (8)$$

with $\mu_{y,l}^t$ and $\sigma_{y,l}^t$ are respectively given by (4) and (5). The labeling dependance appears through the expression of mixed means and variances within each CR pixel. Label conditional independence has been assumed mainly for computation facilities. Even though it is inaccurate, it is widely assumed in the literature since cross-correlation between dates are very difficult to estimate.

3.4. Maximum a posteriori

The classification phase is based on the maximum *a posteriori* criterion, selecting the *optimal* labeling of the random field \tilde{l} knowing the coarse resolution observation v . Let $\tilde{l} = (\tilde{l}_1, \dots, \tilde{l}_{|S|})$ denote the optimal labeling configuration, where the k^{th} coordinate represents the label of the k^{th} segments. According to the Bayes rule, the *a posteriori* probability writes, for all date t ,

$$\mathbb{P}(\tilde{l} = l | \tilde{v}^t = v^t) = \frac{\mathbb{P}(v^t | \tilde{l} = l) \mathbb{P}(l)}{\mathbb{P}(v^t)}. \quad (9)$$

The *optimal* label field is obtained maximizing this latter probability :

$$\tilde{l} = \underset{l \in L^{|S|}}{\operatorname{argmax}} \{ \mathbb{P}(\tilde{l} = l | \tilde{v}^t = v^t) \} \quad (10)$$

Notice that $\mathbb{P}(v^t)$ is a constant for the considered problem. Moreover, in absence of more prior spatial information, all labels are assumed to have the same probability of realization. Hence, the prior $\mathbb{P}(l)$ is also a constant for the maximization problem. Therefore, for a fixed date t , the maximum *a posteriori* is the label field \tilde{l} solution of the problem

$$\min_{\{l \in L^{|S|}\}} \sum_{y \in \Omega'} \left(\frac{(v^t(y) - \mu_y^t)^2}{(\sigma_y^t)^2} + \ln(\sigma_y^t)^2 \right). \quad (11)$$

Considering time series, the mean vector writes $\mu_{y,l} = (\mu_{y,l}^t)_{t \in T}$ for all pixel y and, neglecting correlations between dates, the covariance matrix $\Sigma_{y,l}$ is diagonal with diagonal values $(\sigma_{y,l}^t)^2_{t \in T}$. Hence, (11) becomes

$$\min_{\{l \in L^{|S|}\}} \sum_{y \in \Omega'} \left((v(y) - \mu_{y,l})' \Sigma_{y,l}^{-1} (v(y) - \mu_{y,l}) + \ln(\det(\Sigma_{y,l})) \right), \quad (12)$$

where $'$ denote the transpose matrix. From this energy, we propose, on the one hand, a supervised algorithm leading to a solution of (12) and, on the other hand, an unsupervised algorithm solving a simplified problem. Indeed, if class features (mean and variance) are known *a priori*, a solution of (12) can be derived using a global optimization process. In practice, class features are generally unknown: they depend on the acquisition date, local vegetation growing process, the atmospheric conditions, etc. Therefore, unsupervised classification method seems more appropriate. However, if class means can be estimated during the minimization process (as explained in the following section), class variance estimation is much more sensitive and time-consuming. Hence, according to some robustness considerations, the unsupervised approach assumes equal class variances instead of estimating them inaccurately. This assumption leads, from (6), to constant variances for all pixel y and (12) simplifies as

$$\min_{\{l \in L^{|S|}\}} \sum_{y \in \Omega'} \sum_{t \in T} (v^t(y) - \mu_{y,l}^t)^2. \quad (13)$$

An algorithm of simulated annealing type is described in the next section as an optimisation process for both the supervised and unsupervised method.

3.5. Algorithm

Because of the size of the solution space, a systematic search of the minimum is impossible. As no heuristic seems justified for this problem, we chose a simulated annealing algorithm. It has been widely used for different optimization problems.²² As far as the unsupervised approach is concerned, the algorithm takes as inputs a HR segmentation, a CR time series of the same scene and the number of labels required. It returns the label field solution of (13) and the class features (means). For sake of simplicity, denote E_l the global energy to minimize in (13). It stands for the energy corresponding to a label field l . Each step of the algorithm changes randomly one segment label in the label field and tests whether it makes the energy decreasing or not. Denote $E_{l_{prev}}$ the energy corresponding to the previous label field. The algorithm is the following.

<p>Compute the segments proportions $(\beta_k(y))_{y,k}$ for all pixel y. Initialize randomly the label field. Estimate the label's mean (by linear regression). Initialize the temperature T to the graph's diameter. While a label is not rejected $n_r \times S$ times successively, do for $i = 0$ to S select randomly a segment k and a label c for this segment, re-estimate the label's means, compute $\Delta E = E_l - E_{l_{prev}}$ if $\Delta E \leq 0$ accept the label change, else reject it with a probability $\exp(-\Delta E/T)$. $T \leftarrow qT$.</p>
--

Theoretically, the temperature descent must be logarithmic to lead to convergence but empirically, setting $n_r = 400$ and $q = 0.999$ provided good results in our experiments.

The supervised algorithm takes as inputs a HR segmentation, a CR time series of the same scene and the class features (means and variances). It returns the label field solution of (12). The algorithm is the same as the unsupervised one except for the label's mean estimation and re-estimation (since means and variances are *a priori* known) and E_l now refers to the complete energy to minimize in (12). As the class feature estimation step is time-consuming, this version runs faster than the unsupervised one. Some results are presented next section using both simulated and actual data.

4. RESULTS

In this section, we first consider the monotemporal case in order to analyse the performance of the algorithm (supervised and unsupervised version). The application of the method to actual times series is then studied.

4.1. Validation of the proposed method

The method presented in Section 3 was, at first, assessed using simulated CR images in order to evaluate the method independently of the validity of data assumptions (*e.g.* gaussian hypothesis) and on the preprocessing quality (*e.g.* image superimposition). However, images have been simulated with actual data features. Practically, we started simulating HR images using actual SPOT/HRV resolution characteristics and radiometric means and variances. From an actual HR time series of Danubian plain, 165 images of size 300×300 have been extracted, segmented in 100 fragments and finally labeled using 5 classes. Class means are similar to the ones presented Figure 6.e and class variances are very close to each other (about 0.05) except for one class (about 0.005). Data have been simulated as a mixture of normal distributions, which means and variances were chosen to be the same as those computed from the actual data set. Secondly, CR images have been created averaging HR simulated images. The reference classification is the label map from which HR have been simulated. In this section, as we focus on the method and algorithm behavior, we run experiences using single CR images rather than CR time series. Global results are given in Table 1 for both supervised and unsupervised approaches. As far as the supervised version is concerned, results are very satisfactory since the average percentage of mislabeled pixels is less than 1%. The percentage of mislabeled segments may seem high but it is due to the presence of very small regions in the segmentation (as the percentage of pixels concerned remains very small). In fact, we made the choice not to constrain the segmentation so far. A minimum segment size may be imposed in the segmentation processing step, permitting to avoid that kind of artefacts. In the case of the unsupervised version, the results are rather convincing since the average percentage of misclassified pixels is less than 5% and let us specify that the median percentage is 0.2. Now global results have been introduced, let us analyse the different types of error we obtained in the next section.

Table 1. Results obtained for a set of 165 simulated images with a resolution ratio HR/CR of 15^2 .

	Average % of mislabeled pixels	Average % of mislabeled segments
Supervised version	0.87	23.6
Unsupervised version	4.35	31.5

4.1.1. Errors analysis

Different kinds of errors have been observed. The first one results of an ill-optimization step, showing that the algorithm has not reached the global minimum. To separate this first kind of errors from the others, we compare the energy value E_0 corresponding to the reference label map to the energy value E_k associated to a label configuration that is equal to the reference one except on the segment numbered k . This comparison is made successively for all erroneous segment k obtained by the algorithm. Such an experience has been performed in order not to disorder the analysis by the presence of errors concerning other segments. For instance, we encountered a case where three connex segments were misclassified because label correction of only one of them (leaving the two others erroneous) increased the energy and only correct labeling of the three of them simultaneously made the energy decrease. Then, in this particular case, since the optimization step should have overcome the local minima problem, the error is clearly due to a simulated annealing convergence problem. This type of error can be recognized as such considering only one erroneous segment at a time.

As a comparison, we first focus on the distribution of the energy variations that a mislabeled segment would generate. As the minimum difference value is -1567 but the main occurrences are above -8 , Figure 2.a presents the differences obtained between -8 and 0.1 . We note a peak representing almost 50% of occurrences for difference energies between -0.1 and 0 and the importance of the percentage of occurrences ranged from -2 to 0.1 . The presence of non negative values is also noticeable and will be discussed further on. They concern 10% of occurrences and the maximum difference energy value is 0.004 .

Figure 2.b represents the energy differences histogram computed for all label configurations containing exactly one mislabeled segment among all erroneous ones we obtained. Even if the attributed erroneous label corresponds

to an energy value upper than E_0 , the algorithm was unable to detect that the achieved local minimum was not the global one. If the energy difference is negative, the attribution of the considered erroneous label did not decrease the energy but was attributed to the segment. This problem of algorithm convergence represents 41.3% of occurrences. However, we note that these energy difference values are extremely low comparing to those presented Figure 2.a (about 10^{-3} lower) and hence such convergence errors are limited to the case of local minima close to the global ones. The algorithm may be improved changing the temperature descent law for a logarithmic one, but there is a compromise to do between time processing and result performance. Let us now analyse the positive difference errors that represent the major case (58.7%).

The positive difference errors result mainly from two types of phenomenon. On the one hand, we observed that little segments may be misclassified if their occupation rate within a CR pixel is too low to have an effective contribution on the energy. In some cases, a little segment may be misclassified in order to compensate for the signed error due to the contribution of the other (well classified) segments of the CR pixel. On the other hand, as gaussian distributions we consider partially overlap, an observation occurring in the distribution queue of its class may have a higher probability to belong to the adjacent class and, hence, to be misclassified. The use of multitemporal data, as performed in the actual case, may reduce this type of errors as classes become more severable with time series.

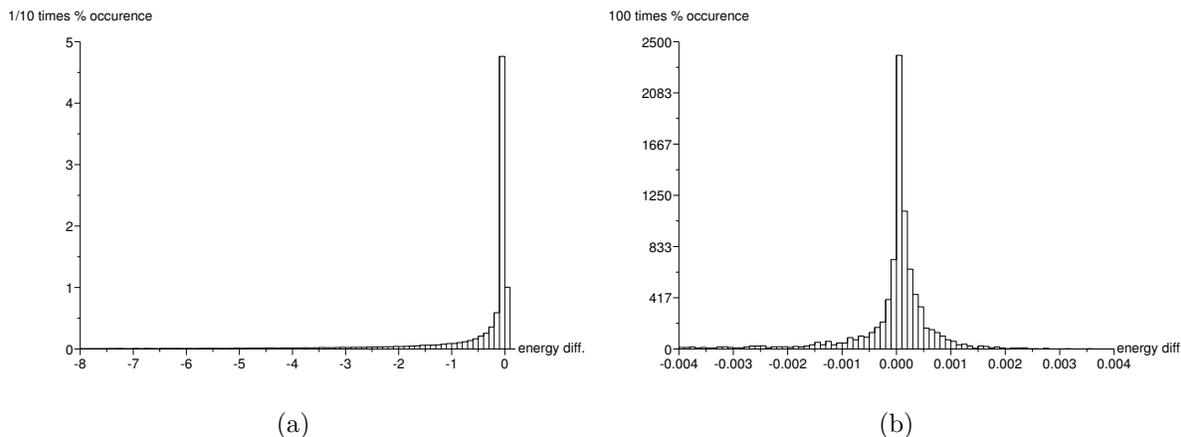


Figure 2. (a) represents the histogram of the normalized energy variation $(E_k - E_0)/E_0$ computed for all segments of S labeled successively with all inaccurate labels. (b) represents the histogram of normalized energy variation $(E_k - E_0)/E_0$ computed for all errors obtained using the supervised algorithm.

4.1.2. Sensitivity to resolution ratio

As we aim at dealing with high resolution ratio up to 15^2 *e.g.* for SPOT/HRV versus MERIS or 50^2 *e.g.* for SPOT/HRV versus SPOT/VGT, we now study the method performance when the resolution ratio varies. Moreover, as Table 1 has shown that tiny segments (containing one or few HR pixels) tended to be misclassified, this study may give some hints for the minimal acceptable segment size.

From a set of 165 simulated HR images, CR images have been created for three different resolution ratio (15^2 , 30^2 and 50^2). For each error obtained using the supervised approach, the occupation rate of the concerned segment within a CR pixel is computed. Figure 3.a-c presents the histograms of the error occupation rate for resolution ratio 15^2 , 30^2 and 50^2 . We note that these histograms look very similar, showing that the size of erroneous segments relatively to CR pixel size is independent on the resolution ratio. We observe a peak of occurrence for segments occupying less than 1% of the CR pixel. This observation could be used to constrain the segmentation processing with a minimum segment size of 1% of the CR pixel, probably leading to about half less errors with the labeling processing.

In the literature, many methods using CR data for class proportion estimation deal with parcels representing more than 5% of the CR pixel.²³ Table 2 presents some occupation rate statistics computed on the set of error segments. As we observe that about 70% of the error segments occupies less than 5% of a CR pixel, restricting to larger segments could increase strikingly the method performance.

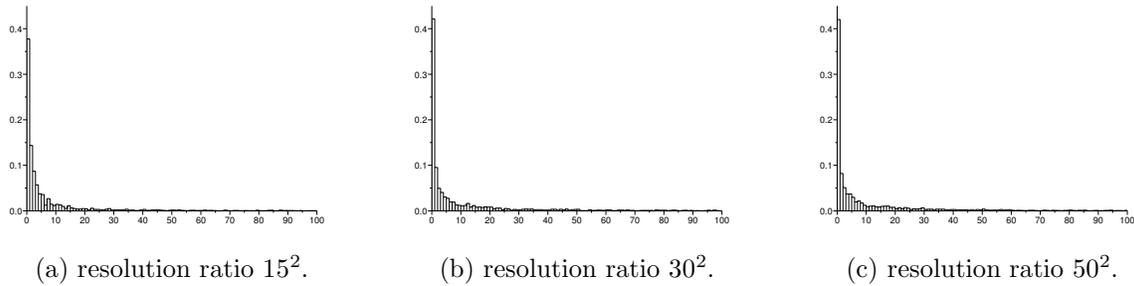


Figure 3. Histogram of error segments occupation percentage within a CR pixel for different given resolution ratio z . The x-axis represents the occupation percentage of an erroneous segment within a CR pixel. The y-axis represents the percentage of occurrence.

Table 2. Statistics on error segment occupation percentage within a CR pixel versus the considered resolution ratio z .

z	10 percentils	median	70 percentils	80 percentils	90 percentils
15^2	0.44	2	5	10	25.7
30^2	0.11	1.77	7.7	16.7	41.4
50^2	0.08	2	8.13	17.6	39.8

As far as the performance of the method versus the resolution ratio is concerned, Figure 4 shows that the error percentage values are relatively spreaded. This can be attributed to the fact these experiences use monotemporal images while some class features may be very close depending on the considered date, leading to the observed quality variability of the results. Moreover, the median curve shows very good performance for resolution ratio up to 30^2 , then performance decreases with resolution ratio 50^2 with about 20% of erroneous pixels. The dependency of the error percentage on the sensor resolution is manifest and non linear. Notice that conversely to the previous case where segment size were considered relatively to CR pixels, the error size is here *absolute*, mesured in terms of HR pixels. The coarser is the sensor resolution, the more multitemporal information is necessary for good performance. Note also that the results we obtain using the supervised algorithm are very close in performance to those obtained using the unsupervised version. Be aware that the data set used for Figure 4.b is less significant than the one used for Figure 4.a due to the more important run time of the unsupervised version.

In summary, the method has shown good performance on monotemporal simulated data using both the supervised and unsupervised version, with up to 99% of accurate classification for the supervised approach and about 95% for the unsupervised one. In the next section, an application of the method to actual time series is presented.

4.2. Application to actual data

Since the simulation results showed good performances, our method has been applied to some actual data. First, the same experience as Section 4.1.2 analyzing the behavior of the method versus to the resolution ratio has been performed using a monotemporal actual data set. Figure 5 shows, as in the case of simulated data, that the error percentage is quite variable. However, error values are significantly more important than the simulated case, *e.g.* the median error percentage value is about 4.5% for resolution ratio 15^2 while it reaches 0.3% using simulated data. This quite significant difference may be attributed to the fact that, in the case of monotemporal

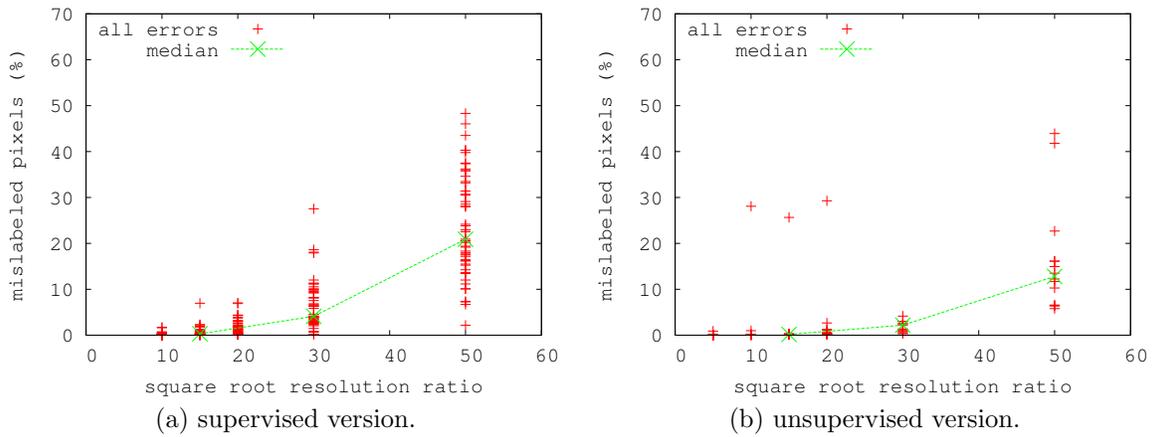


Figure 4. Each cross represents the percentage of misclassified pixels obtained for a set of images and a given resolution ratio, using the supervised version (a) and the unsupervised version (b). The superimposed line represents the median of these percentage values.

actual images, the Gaussian hypothesis is more or less verified. Moreover, the data set we used contains images acquired during winter where the response of bare soil can present a very specific pdf. Further studies will focus on the same type of experiences using time series, in order to get closer to the Gaussian assumption.

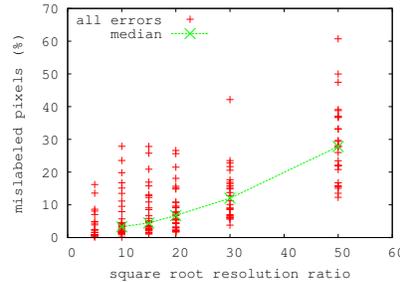


Figure 5. Percentage of mislabeled pixels obtained applying the supervised algorithm to monotemporal actual images.

In a realistic context, the proposed approach has been applied to a subpart of an actual SPOT/HRV time series of the Danubian plain (Rumania) provided by the French spatial agency (CNES) in the framework of the ADAM European project. The aim was the determination of the types of vegetation present in the scene. A time series of 8 images has been extracted from the data set, and pre-processed into land cover fraction series, as this parameter is scale linear and competitive for land cover discrimination (see Section 3.2). The Mumford and Shah segmentation method has been processed on one HR image asking for 100 segments, leading to the segmented image presented Figure 6.a. The previous analysis on simulated data concerning the sensitivity of the method to the resolution ratio showed good performance for a resolution data of 15^2 (*c.f.* Figure 4). Hence, we start applying the method for this resolution ratio. As we could not dispose of MERIS series at the time, CR MERIS data have been simulated from the actual HR time series (see Figure 6.b), by averaging with a factor 15 in each direction. Moreover, we chose to run the unsupervised version of the method as it is automatic (but the number of classes) and hence easier to handle from a user point of view. The classification we obtained for 5 classes is presented Figure 6.c. This result has been compared to the reference classification obtained from the original HR time series using the same method. This comparison leads to the error map Figure 6.d showing about 3% of erroneous pixels. Hence, the classification obtained using the unsupervised approach is up to 97% identical to the one obtained from HR time series. This error is rather satisfying even if we remark some segments of size

larger than the CR pixel size, a more thoughtful selection of the used acquisition date may improve the results. An other output of the method is the estimation of the class means. Figure 6.e presents the land cover fraction mean estimation corresponding to each label obtained from CR data and from the reference HR data set. We notice, in a satisfying way, that these estimated profiles are very close, showing unsupervised means estimation from CR data is rather accurate.

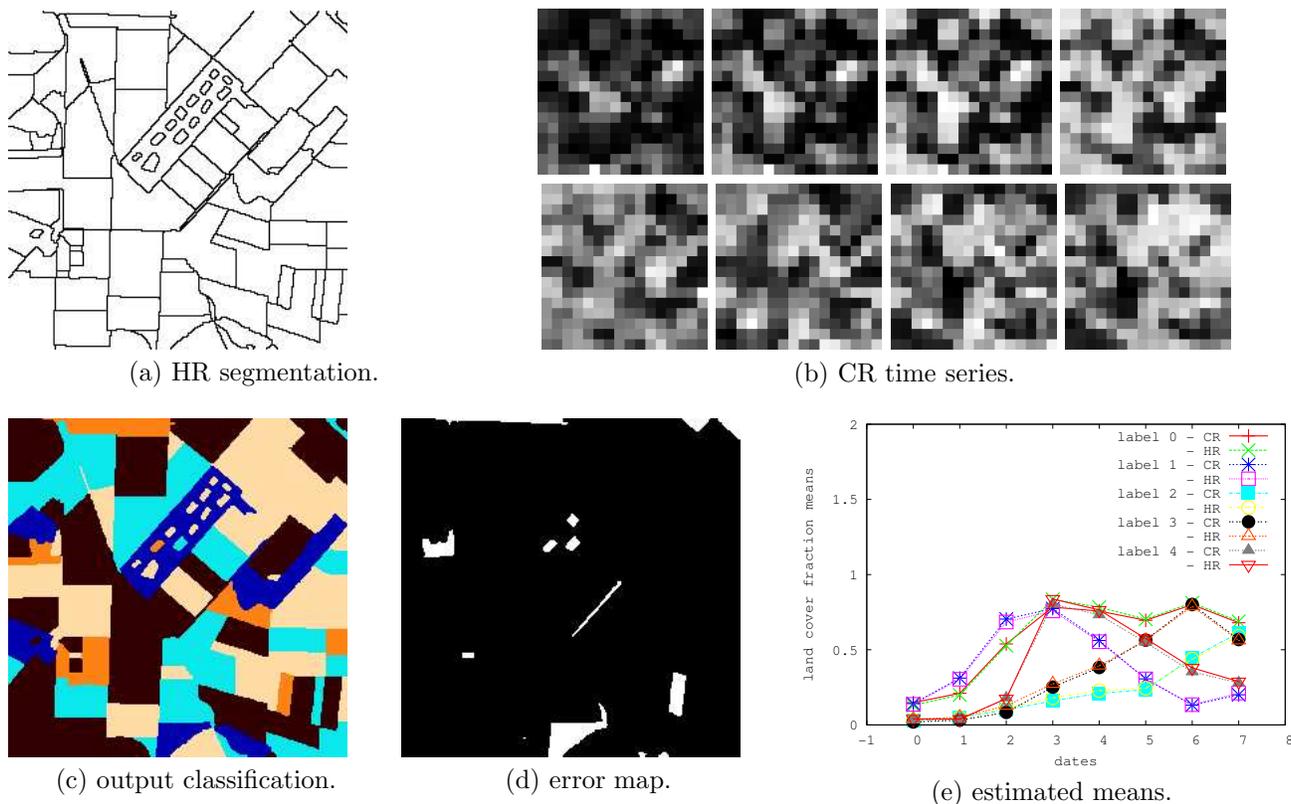


Figure 6. Results obtained with the unsupervised approach using segmentation (a) and actual time series (b): the classification (c) is 97% accurate, (d) represents in white the errors obtained (3% of HR pixels). (e) shows that land cover fraction means estimated by the algorithm from CR time series are very close to the ones performed from HR time series.

5. CONCLUSION

A new method based on Bayesian theory and linear mixture modeling has been presented, with two operational versions: the supervised version requires the *a priori* knowledge of class features while the unsupervised one is automatic except the number of class. The results obtained using simulated data showed up to 99% of accurate classification using the supervised version and more than 95% of accurate classification using the unsupervised one. The analysis of the errors has shown that less than 50% of the errors were due to the convergence of the simulated annealing algorithm into local minima very close to the global one. This local optimization results from a compromise between error rate and computation time. The analysis of the performance obtained versus the resolution ratio considered are convincing dealing with high resolution ratio. Since the requirement knowledge of class features for the supervised method may be difficult to fulfil, it favours the use of the unsupervised version. The application on actual CR time series of 8 images with a resolution ratio of 15^2 were very satisfying as the classification obtained is 97% accurate. Future works will focus on more performance analysis, particularly on the method sensitivity to the choice of acquisition dates. Moreover, the validity of the model's assumptions will be studied in the case of different types of actual data (*e.g.* luminance, land cover fraction).

5.1. Acknowledgments

We thank the french spatial agency (CNES) for providing the ADAM data base. A. Robin is funded by a grant from the CNES and EADS/Astrium.

REFERENCES

1. A. Schistad, T. Taxt, and A. Jain, "A markov random field model for classification of multi-source satellite imagery," *IEEE Trans. Geosci. Remote Sens.* **34**, pp. 100–113, Jan. 1996.
2. Y. Jhung and P. Swain, "Bayesian contextual classification based on modified m-estimates and markov random fields," *IEEE Trans. Geosci. Remote Sens.* **34**, pp. 67–75, Jan. 1996.
3. Z. Kato, J. Zerubia, and M. Berthod, "Bayesian image classification using markov random fields," pp. 375–381, Proc. 12th Int. Maxent Workshop, (Paris, Frances), July 1992.
4. D. Geman and B. Gidas, *Image analysis and computer vision*, ch. 2, pp. 9–36. Washington, DC, 1991.
5. S. Geman and D. Geman, "Stochastic relaxation, gibbs distribution and bayesian restoration of images," *IEEE Trans. on Pattern An. and Mach. Intell.* **6**(6), pp. 721–741, 1984.
6. H. Derin and H. Elliott, "Modeling and segmentation of noisy and textured images using gibbs random fields," *IEEE Trans. Pattern Anal. Machine Intell.* **PAMI-9**, pp. 39–55, Jan. 1987.
7. R. K.Sharma, T. K. Leen, and M. Pavel, "Probabilistic image sensor fusion.," in *Advances in Neural Information Processing systems*, **11**, pp. 824–830, 1999.
8. R. Molina, M. Vega, J. Abad, and A. Katsaggelos, "Parameter estimation in bayesian high-resolution image reconstruction with multisensors," *IEEE Trans. on Image Processing* **12**(12), pp. 1655–1667, 2003.
9. J. Nunez, X. Otazu, . Fors, A. Prades, V. Pala, and R. Arbiol, "Multiresolution-based image fusion with additive wavelet decomposition," *IEEE Trans. Geosci. Remote Sens.* **37**(3), pp. 1204–1211, 1999.
10. L.Wald, T. Ranchin, and M. Mangolini, "Fusion of satellite images of different spatial resolution : Assessing the quality of resulting images," *Photogramm. Eng. Remote Sensing* **63**(6), pp. 691–699, 1997.
11. G. Flouzat, O. Amram, F. Laporterie, and S. Cherchali, "Multiresolution analysis and reconstruction by a morphological pyramid in the remote sensing of terrestrial surfaces," *Signal Processing* **81**, pp. 2171–2185, October 2001.
12. S. L. Hégarat-Masclé, I. Bloch, and D. Vidal-Madjar, "Application of dempster-shafer evidence theory to unsupervised classification in multisource remote sensing," *IEEE Trans. Geosci. Remote Sens.* **35**, pp. 1018–1031, July 1997.
13. L. Bruzzone, D. Prieto, and S. Serpico, "A neural-statistical approach to multitemporal and multisource remote-sensing image classification," *IEEE Trans. Geosci. Remote Sens.* **37**, pp. 1350–1359, May 1999.
14. A. Robin, L. Moisan, and S. L. Hégarat-Masclé, "Automatic land cover change detection from coarse resolution images using an a contrario approach." MAP5 prepublication 2005-3.
15. F. V. D. Meer, "Iterative spectral unmixing," *Int. J. Rem. Sens.* **20**(17), pp. 3431–3436, 1999.
16. J.-M. Morel and S. Solimini, *Variational methods in image segmentation*, Birkhauser, 1995.
17. N. Quarmby, J. Townshend, J. Settle, K. White, M. Milnes, T. Hindle, and N. Silleos, "Linear mixture modeling applied to avhrr data for crop area estimation," *Int. J. Remote Sens.* **13**(3), pp. 415–425, 1992.
18. H. Cardot, R. Faivre, and M. Goulard, "Functional approaches for predicting land use with the temporal evolution of coarse resolution remote sensing data.," *J. of Applied Stat.* **30**, pp. 1185–1199, 2003.
19. R. Faivre and A. Fischer, "Predicting crop reflectances using satellite data observing mixed pixels," *J. of Agr. Bio. and Env. Stat.* **2**, pp. 87–107, 1997.
20. R. Defries, M. Hansen, and J. Townshend, "Global continuous fields of vegetation characteristics : a linear mixture model applied to multi-year 8 km avhrr data.," *Int. J. Remote Sens.* **21**(6), pp. 1389–1414, 2000.
21. H. Kerdiles and M. Grondona, "Narvi decomposition and subpixel classification using linear mixing in the argentinean pampa," *Int. J. Remote Sens.* **16**(7), pp. 2231–2240, 1995.
22. P. V. Laarhoven and E. Aarts, *Simulated Annealing: Theory and Applications*, Springer, 1987.
23. S. L. Hégarat-Masclé, C. Ottlé, and C. Guérin, "Land cover change detection at coarse spatial scales based on iterative estimation and previous state information," *Remote Sensing of Environ.* **95**(4), pp. 464–479, 2005.