Simultaneous multi-source and multi-temporal land cover classification using a Compound Maximum Likelihood classifier

Mariane Souza Reis, Luciano Vieira Dutra*, Maria Isabel Sobral Escada

Brazilian National Institute for Space Research
Postal Code 515–12227–010 – São José dos Campos – SP – Brazil

{reis, dutra, isabel}@dpi.inpe.br

Abstract. *The most widely used change detection method is to classify remote sensing images independently for each date, and stack them to form a class sequence vector. However, impossible transitions within the sequences might occur and errors might be accumulated. To solve this, we propose a novel algorithm called Compound Maximum Likelihood (CML), based on the Maximum Likelihood classifier (ML). In CML information from all images is used jointly by considering the a priori probability of each class sequence. The algorithm was tested for Synthetic Aperture Radar and optical images classification in a study area in Pará state, within the Brazilian Amazon. CML presented either similar or very improved accuracy index values over ML land cover classifications.*

Resumo. *O método de detecção de mudanças mais comumumente utilizado é comparar imagens classificadas independentemente para obter vetores de sequências de classes no tempo. No entanto, transições impossíveis podem ser classificadas e erros são acumulados. Para solucionar esses problemas, propõe-se o algoritmo de Máxima Verossimilhança Composta (MVC), como uma extensão do classificador de Máxima Verossimilhança (MaxVer). No MVC, todas as imagens são usadas em conjunto, dada a probabilidade a priori de cada sequência de classes. Testou-se o MVC para classificar imagens ópticas e de Radar de Abertura Sintética de uma área do estado do Pará, na Amazônia. O MVC apresentou resultados ou similares ou consideravelmente melhores que MaxVer.*

1. Introduction

The understanding of ecosystems functioning over time and the effects of natural phenomena and human activities over the environment require information about the dynamics of Land Use and Land Cover (LULC). This information is usually obtained by change detection, defined as the process of identifying differences in the state of an object or phenomenon at distinct times [Singh 1989]. A common input data for change detection in environmental studies is remote sensing data. Different change detection methods have been proposed over time and organized in various ways [Lu et al. 2004, Kennedy et al. 2009, Tewkesbury et al. 2015, Blaschke 2005]. Among those, we can cite

*Corresponding author.

four categories:'*layer arithmetic*', '*post-classification comparison*', '*direct classification*' and '*hybrid methods*'.

Layer arithmetic refers to methods that calculate change indicators directly over pixel values or derived features of two or more images. The changes themselves are usually detected if the indicator values are lower or higher than a given threshold [Lu et al. 2004]. These methods are usually easy to implement and previous knowledge of land cover and changes are practically unnecessary. However, the use of data from the same sensor and different dates may require careful calibration. The use of different sensors within this method is still incipient [Chatelain et al. 2008, Prendes 2015], and only possible if some feature depicting common characteristics of targets can be calculated. *Layer arithmetic* also usually provides only binary (change and no-change) maps. According to [Lu et al. 2004], adequate studies in change detection should provide, besides the occurrence or not of a change, information about change area, rate, spatial distribution and type (from some LULC class to another). Therefore, other change detection methods are often necessary.

Post classification comparison is the direct comparison of separately classified images. According to Tewkesbury et al. (2015), this is "*one of the most established and widely used change detection methods*". This method has some clear advantages over *layer arithmetic*. Firstly, differences among data are automatically diminished, since each image is classified separately, so radiometric transformation/calibration is usually not necessary. It also provides a complete matrix of change information (from-to), while only the *a priori* knowledge about land cover is previously necessary. Nonetheless, this method is often criticized because the final change map depends on the quality of individual classified images [Lu et al. 2004, Tewkesbury et al. 2015, Fuller et al. 2003] and it is not possible to measure changes occurring in a lower scale than the combined errors of the individual maps [Fuller et al. 2003]. Additionally, it is possible to map changes that could never happen in the field, because of errors in the individual classified images [Anjos et al. 2015, Reis et al. 2017].

Direct classification consists in classifying a multi-temporal set of images directly into multi-temporal classes. Like *post classification comparison* methods, radiometric calibration is usually not necessary and multi-sensor data can be used. Change detection by *direct classification* can be done using supervised or unsupervised classification algorithms [Tewkesbury et al. 2015]. Since the use of unsupervised algorithms seems to offer very complex results, it may be interesting to focus on the use of supervised algorithms instead, for which labeled samples of classes are needed. In change detection cases, these samples must be collected over a set of multi-temporal images. On one hand, the analyst is able to set which changes he wishes to detect, so that unimportant or impossible changes would not be mapped [Anjos et al. 2015]. On the other hand, the types of changes must be identified *a priori* and it is considerably more difficult to collect samples of change classes than those of land cover classes. Additionally, unidentified types of change that occurred in the set of images would be incorrectly identified, or not mapped as changes at all.

Lastly, *hybrid methods* are those in which two or more types of change detection methods previously described are used together. The clear advantage of this methods is that change detection can be improved. At the same time, it may be time consuming and

the appointed problems of other methods have the potential to persist.

Using *post classification comparison* based methods, it is possible to derive a class sequence for an object in time from the classification of this object at each time. Similarly, if labeled samples are collected over the class sequences, it is possible to derive the classification of each time from change maps generated using *direct classification* approaches. However, given the limitations of each methodology, derived land cover or change maps could present accuracy problems. Based on these two change detection methodologies, this work presents a novel algorithm named Compound Maximum Likelihood (CML), derived from the widely known Maximum Likelihood (ML) classifier. CML was conceived to jointly classify two or more images from the same area and it is based on the fact that knowledge of land cover dynamics in a given study area and time interval can be used to both refine time-series classifications and to restrict impossible or improbable land cover transitions [Gómez et al. 2016]. The proposed algorithm is described in Section 2. The data and methodology employed for a case study in Tapajós region, within the Brazilian Amazon, is presented in Section 3. Two uses of CML were presented: 1) CML was used to improve both change detection and land cover classification; 2) CML was adapted to use multi-sensor information jointly, to classify the land cover of only one date. Classification results are presented and analyzed in Section 4. Conclusion and considerations for future work are drawn in Section 5.

2. Compound Maximum Likelihood

Consider $s_l = \{\omega^1_{k_1}, ..., \omega^t_{k_t}, ..., \omega^F_{k_F}\}$, $s_l \in S = \{s_1, ..., s_M\}$, the l_{th} temporal class sequence s in the set of M possible sequences. F is the length of the time sequence. $\omega^t_{k_t} \in \Omega_t = \{\omega^t_1, ..., \omega^t_{k_t}, ...\omega^t_{N_t}\}$, where ω^t is the actual class at time position t of s_l. k_t is an indicator of class ω^t in the set Ω_t, in which Ω_t is the set of N_t possible classes on time t. A given observation vector $\vec{X} = \{\vec{x_1}, ..., \vec{x_t}, ..., \vec{x_F}\}$ contain the F temporal observations that can indicate the class of an object (like digital numbers in image pixels). The rule to determine a particular nature of a sequence $\hat{S}$ (which classes are attributed to the object at each analyzed time) is proposed as

$$\hat{S} = \arg_s \max(P(s \in S|\vec{X})) \tag{1}$$

in which

$$P(s_l|\vec{X}) = \frac{P(\vec{X}|s_l) \times P(s_l)}{P(\vec{X})}, \tag{2}$$

as defined by Bayesian rule. Since $P(\vec{X})$ is inconsequential to the maximum calculation

$$\hat{S} = \arg_s \max(P(\vec{X}|s) \times P(s)). \tag{3}$$

Supposing that the observations are time independent and that each one depends only on the observed object, we have

$$P(\vec{X}|s_l) = P(\vec{x_1}|\omega^1_{k_1}) \times ... \times P(\vec{x_t}|\omega^t_{k_t}) \times ... \times P(\vec{x_F}|\omega^F_{k_F}) \times P(s_l) \tag{4}$$

in which $P(s_l)$ is the *a priori* probability of a sequence s_l.

For a Gaussian distribution, $P(\vec{x_t}|\omega_{k_t}^t)$ is given by:

$$P(\vec{x_t}|\omega_{k_t}^t) = (2\pi)^{B/2}|\Sigma_{k_t}|^{-1/2}\exp\{-1/2(\vec{x_t} - m_{k_t})^T\Sigma_{k_t}^{-1}(\vec{x_t} - m_{k_t})\} \quad (5)$$

in which B is the number of channels of the image for time t, m_{k_t} and Σ_{k_t} are respectively the mean vector and the covariance matrix of the class $\omega_{k_t}^t$ and $|\cdot|$ is the determinant function.

Observe that the rule expressed in equations (3) and (4) would return the same sequence as concatenating the Maximum Likelihood classifications in each point of time if the presence of the *a priori* probability of a particular sequence is inconsequential. The expression (4), excluding the *a priori* term, is called **compound likelihood**.

3. Methodology

To test the proposed algorithm, we selected an area of approximately 412 km^2 in Belterra, Pará state, within the Brazilian Amazon, as illustrated in Figure 1. It is a relatively plane area of humid tropical climate [IBAMA 2004]. Originally, the region presents dense forest vegetation, in which woody lianas, palms and epiphytes are found. Due to the occupation process, the study area also presents patches of secondary vegetation, pasture and agriculture within the forest matrix. Through field work and remote sensing data, groups of classes were identified in the study area for three analyzed dates, as described in Table 1.

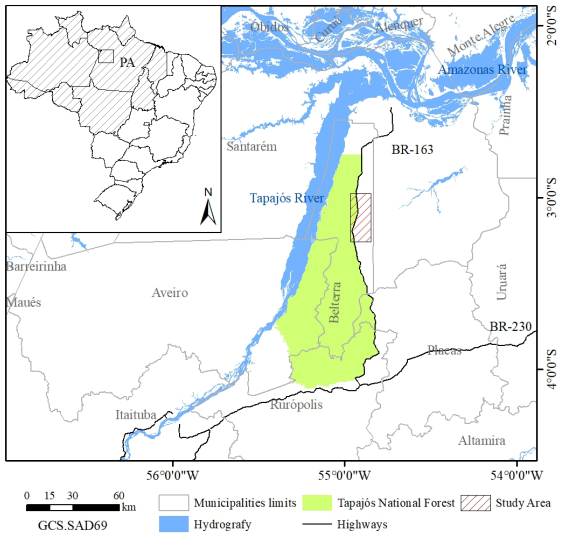

Figure 1. Study area

Two types of analysis were conducted in this work. In both of them, we considered CML applied for per pixel classification of two or more images. The first analysis aims to detect and typify changes in land cover occurred over three different dates (2008, 2010 and 2013), based on the classification of three remote sensing images. These are:

one Advanced Land Observing System (ALOS)/Phase Array L-Band Synthetic Aperture Radar sensor (PALSAR) image from June 15^{th} 2008, acquired in Fine Beam Dual (FBD) mode, 1.1 processing level; one LANDSAT5/Thematic Mapper (TM) image from June 29^{th} 2010 and one Earth Observer 1 (EO-1)/ Advanced Land Imager (ALI) image from October 05^{th} 2013.

Table 1. Land cover classes and legends definitions for each analyzed date.

2008/2010a	2013	2010b	Class name	Description
AG	BS	BS	Bare Ag. Soil	Agricultural areas presenting bare soil.
	IA	IA	Idle Ag. Area	Fallow annual agriculture areas.
	CA	CA	Cultivated Area	Cultivated crops
PA	PA	CP	Clean Pasture	Pasture areas with less than 15% of invasive plants.
		OP	Overgrown Pasture	Pasture areas with more than 15% of invasive plants.
ISV	SV1	SV1	Initial S.V.	Secondary vegetation formed by herbaceous vegetation and shrubs.
	SV2	SV2	Intermediate S.V.	Secondary vegetation composed mainly by shrubs and small trees.
F	SV3	SV3	Advanced S.V.	Secondary vegetation formed mainly by trees.
	MF	MF	Modified Forest	Forested areas modified by logging and/or fire.
	MA	MA	Mature Forest	Climax forests, with small to no evidence of alteration.

Note:AG = Agriculture, PA = Pasture, ISV = Initial stages of secondary vegetation, F = developed forests.

On the second analysis, it is shown how to use CML to classify two images from different sensors and approximately the same date for land cover classification, provided they meet the initial hypothesis of CML derivation. The same LANDSAT5/TM image from June 29^{th} 2010 was used, along a ALOS/PALSAR FBD 1.1 image from June 21^{th} 2010. Radar images are independent from optical images, even if they are close in time. This experiment can be thought of jointly classifying contemporaneous optical and radar imagery without stack layering or executing a fusion process. Additionally, different set of classes can be used for optical and radar in this context, which is not applicable in layer stacking or data fusion.

All images were orthorrectified, projected to UTM (fuse 21S) WGS84 and resampled to 15 m of pixel size, in order to enable comparisons. Additionally, ALOS/PALSAR images were speckle filtered using the Stochastic Distances Nonlocal Means filter [Torres et al. 2014] with the parameters: filtering window equal to 5x5 pixels, patch equal to 3x3 pixels and confidence level equal to 90%. For LANDSAT5/TM and EO-1/ALI images, feature selection processes were executed, based on the Jeffries-Matusita (JM) distance [Schowengerdt 2006] between the pairs of classes from their respective legends. The selected bands for each data, as well as the main characteristics of each image and the respective legends utilized are presented in Table 2. For each image, one land cover legend was adopted and labeled samples were collected. These are presented in Figure 2.

These images were classified using both the traditional ML classifier and the proposed CML. Figure 3 shows the general methodology for CML classification. In this figure, data and processes regarding time 1 are detached from the other times, for clarity purposes. Training for the classifier in each input data is done separately as a standard ML classifier. The classes and reference sets for each site does not need to be the same for all input data. The definition of *a priori* probabilities for class sequences develops the relation among meaningful classes for each image. All classifications are done at the same time by the CML classifier. Note that if the *a priori* probability of sequences is not set, this methodology would return the same results as ML classifier.

Table 2. Remote sensing images characteristics.

	ALOS/PALSAR	LANDSAT5/TM	EO-1/ALI
Type of sensor	Synthetic Aperture Radar	Optical	Optical
Channels	L band (23 cm): polarizations HH and HV	7 spectral bands	9 spectral bands
Acquisition date	June 15 2008 June 21 2010	June 29 2010	October 05 2013
Spatial resolution	10 m in range, 4.5 m in azimuth	30 m	30 m [1]
Channels selected for classification	Filtered HH and HV polarizations	2(0,52-0,60 μm) 4(0.76-0.90 μm) and 5(1.55-1.75 μm)	4'(0.84-0.89 μm), 5'(1,2-1.3 μm) and 7(2.08-2.35 μm)

[1] With exception of panchromatic band, which has 10 m of nominal spatial resolution.

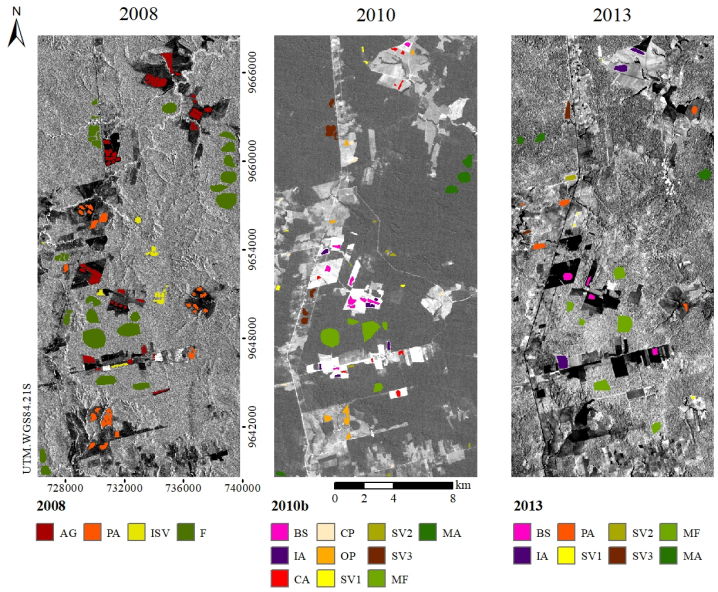

Figure 2. Labeled samples for each legend and date, presented over a band from the respective image: polarization HV from ALOS/PALSAR image (2008), band 5 from LANDSAT5/TM image (2010) and band 7 from EO-1/ALI image (2013). Legend 2010a results from the grouping of legend 2010b.

Exclusively for the case in which different data for 2010 is used jointly, we also classified an image obtained by the fusion of LANDSAT5/TM and ALOS/PALSAR images (both from 2010), for comparison purposes. Selective Principal Components (SPC-SAR) fusion method was selected because it presented better results than using these images separately, stacked or fused by other methods, considering the same study area and similar land cover classes [Pereira 2012]. This step was done so we could further analyze the vantages and disadvantages of using the information of both images jointly.

The definition of the *a priori* probability of class sequences was done differently for each analysis. For the first studied case, in which images from 2008, 2010 and 2013 were classified, we defined which transitions could be acquired between 2008 and 2010 and between 2010 and 2013, by tabulating the legends. These transitions were firstly classified in two classes: *possible* (something that can actually happen in the study area, like a forested class being converted to an agricultural class) and *impossible* (something that

could never happen given the time interval and study area, like a deforested area completely regenerating in only three years). We then weighted these transitions to reflect what we believe is the potential of each transition happening. Weights vary among 0.0 (impossible), 0.3 (possible but not expected in the study area and time interval), 0.5 (possible in specific conditions), 0.7 (transitions that are possible if classes are near transitional states in succession processes) and 1.0 (expected transitions), as presented in Table 3. Both analysis were based on the transitions proposed by Reis et al. (2017). The possibility of transitions from 2008-2010-2013 was calculated by the product of 2008-2010 and 2010-2013 weights. We used the *a priori* probability of class sequences in both the binary (0 for *impossible* classes and 1 for all the *possibles* ones) and weighted way.

For the second study case, in which two images from 2010 are classified, we considered that a given pixel in a image must be classified as corresponding classes in both images. Therefore, the transitions between correspondent classes (formed by the same detailed classes) received a weight equal to 1, while the others received a weight 0.

Classified images were evaluated by the comparison of confusion matrices and Kappa index values. For each class in an used legend, 100 labeled samples were randomly selected and used to calculate a confusion matrix between reference samples and a classified image, from which the value of Kappa index was calculated. This process was repeated 1000 times for each classified image. The mean Kappa index values and the mean confusion matrix were then analyzed.

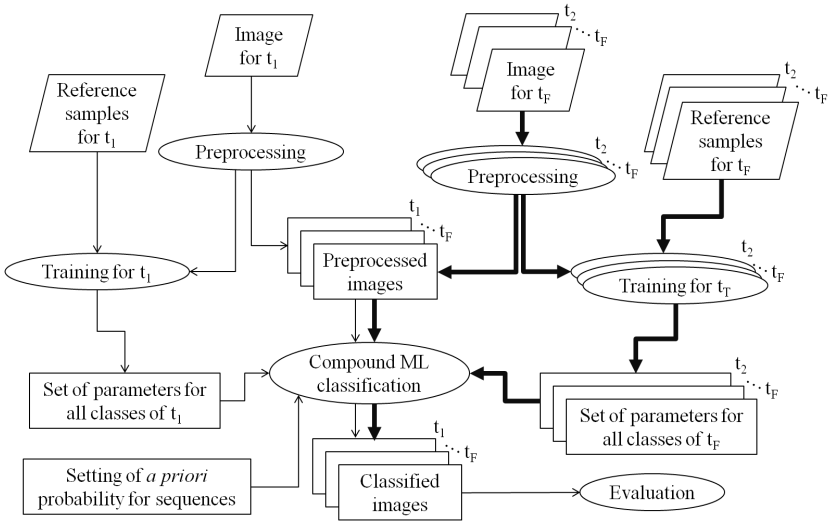

Figure 3. General methodology for Compound Maximum Likelihood classification. Thicker connection lines indicate a high number of outputs.

Table 3. Weights for 2008-2010 and 2010-2013 transitions. For binary case, the weight is 1.0 for each value different than 0.0.

		2010									
		BS	IA	CA	CP	OP	SV1	SV2	SV3	MF	MA
2008	AG	1.0	1.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0
	PA	1.0	1.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0
	ISV	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.7	0.0	0.0
	F	1.0	1.0	1.0	1.0	1.0	0.5	0.0	1.0	1.0	1.0

		2013								
		SE	AP	CA	PL+PS	VS1	VS2	VS3	FD	FP
2010	SE	1.0	1.0	1.0	1.0	0.7	0.0	0.0	0.0	0.0
	AP	1.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0
	AC	1.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0
	PL	1.0	1.0	1.0	1.0	0.7	0.0	0.0	0.0	0.0
	PS	1.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0
	VS1	1.0	1.0	1.0	1.0	1.0	0.7	0.0	0.0	0.0
	VS2	1.0	1.0	1.0	1.0	0.5	1.0	0.7	0.0	0.0
	VS3	1.0	1.0	1.0	1.0	0.5	0.0	1.0	0.0	0.0
	FD	1.0	1.0	1.0	1.0	0.5	0.0	0.0	1.0	0.0
	FP	1.0	1.0	1.0	1.0	0.5	0.0	0.0	0.3	1.0

4. Results and discussion

The mean Kappa index and standard variation values of the classifications from 2008, 2010 and 2013 obtained using CML with binary or weighted *a priori* probabilities and the ones obtained using ML are presented in Table 4. A hypothesis t-test showed that means of classifications from the same date are statistically different in 0.01 significant level. When using the CML, the classifications are done simultaneously. The likelihoods values for each year are multiplied by the *a priori* probabilities for each class sequence. The CML classifications are those whose class conditional probabilities give the maximum when including the *a priori* value. As the *a priori* value is set as 0.0 for sequences with impossible transitions, the class which gives the next viable sequence with maximum compound likelihood forms the CML classifications.

All classifications were improved by the use of CML, using either binary or weighted *a priori* probability of sequences. Although classifications from CML with weighted probability of sequences showed the highest mean Kappa index values, this difference is small compared to the one using binary probabilities of sequences. The gain (in %) in mean Kappa index values for CML classifications over ML ones is also shown in Table 4. As can be seen, ALOS/PALSAR data from 2008 were the most improved classification in CML approaches, with gains over 20% in Kappa index values.

Table 4. Mean Kappa index and standard variation values of the multi-temporal classifications obtained using CML with binary *a priori* probabilities, ML and gain in Kappa value of CML over ML.

Data	Year	# of classes	Kappa index			Gain in Kappa over ML (%)	
			ML	CML		CML	
				Binary	Weighted	Binary	Weighted
ALOS/PALSAR	2008	4	0.49±0.02	0.59±0.03	0.60±0.03	20.5	21.9
LANDSAT5/TM	2010	10	0.70±0.01	0.72±0.01	0.73±0.01	3.0	3.3
EO-1/ALI	2013	8	0.70±0.02	0.74±0.01	0.76±0.01	6.8	9.1

Note: The class CA was not found in the study area in 2013. Therefore, only 8 classes were used.

Given the improvement in mean kappa index values from CML with weighted *a priori* probabilities over ML classifications of ALOS/PALSAR data from 2008, the

mean confusion matrices of these two classified images are presented in Table 5. As can be observed, the increase in mean Kappa values are due to the improvement in the classification of the classes ISV and F. This result was expected, since most impossible transitions between 2008 and 2010 involve secondary vegetation and forest classes, which are also classes with high confusion between them in the classification of ALOS/PALSAR data, as previously observed by Pereira (2012). A subset of images classified using ML and CML with weighted *a priori* probabilities is shown in Figure 4. Visually, it is possible to note that while pixels are similarly classified as pasture or agricultural classes in both methods, many pixels that would be misclassified as secondary vegetation classes are changed to either MF, MA or mostly F in CML classifications.

Table 5. Mean confusion matrices (%) for ALOS/PALSAR classifications obtained with ML and CML (with Weighted *a priori* probabilities).

		Classified image							
		ML				CML (weighted)			
		AG	PA	ISV	F	AG	PA	ISV	F
Reference samples	AG	90.9	40.7	0.6	0.0	90.9	40.5	0.6	0.0
	PA	8.7	49.7	4.6	1.4	8.7	49.7	1.9	0.1
	ISV	0.3	8.9	39.1	31.9	0.3	9.3	49.1	11.2
	F	0.1	0.8	55.7	66.7	0.1	0.6	48.3	88.7

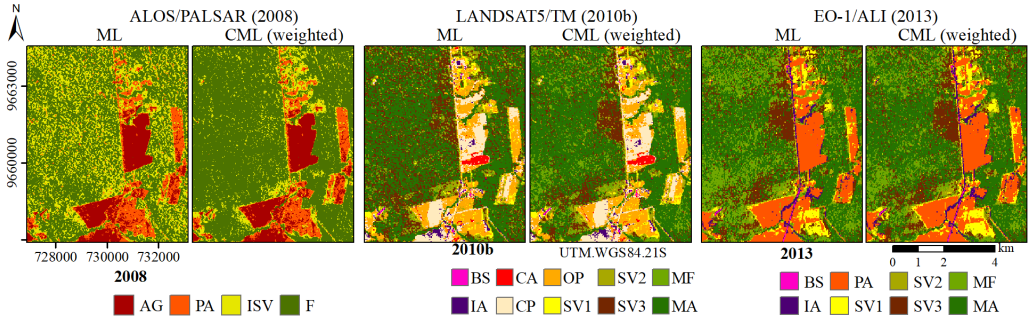

Figure 4. Subset of images classified using ML and CML with weighted *a priori* probabilities.

The mean Kappa index and standard variation values of the classifications from 2010 multi-source data are presented in Table 6. A hypothesis t-test showed that all mean Kappa index values are statistically different at a 0.01 significance level. The use of both CML with original data or ML with fused data similarly improved land cover classification for ALOS/PALSAR data and 4 classes. However, results for LANDSAT5/TM data and 10 classes are slightly less accurate when using CML to integrate the information of this data and ALOS/PALSAR one. This behavior was also noted in fused data classification, denoting that the additional use of ALOS/PALSAR data has the potential to decrease accuracy of LANDSA5/TM image classification using either method. Nonetheless, CML returned better values than those of ML classification of fused data in each legend. Even though ML classification of LANDSAT5/TM presented more accurate results, differences in mean Kappa index values for this classified image and CML using the same input data are small (less than 0.01 in mean Kappa Value and a p value = 0.002). A subset of the ALOS/PALSAR and LANDSAT5/TM images classified using CML and the fused images

classified with ML are presented in Figure 5. Note that in this subset, CML was capable of providing similar results than those of ML over fused data, without the necessity of a fusion process. It also allows for the use of different legends in the same classification process.

Table 6. Mean Kappa index and standard variation values of classifications of the same year and gain in Kappa value of CML over ML of original and fused data.

Data	Year	# of classes	Kappa index		Gain in Kappa from CML[a] (%)
			ML	CML	
ALOS/PALSAR	2010	4	0.52±0.03	0.79±0.02	51.1
LANDSAT5/TM	2010	10	0.70±0.01	0.70±0.01	-0.3
Fused data	2010	4	0.76±0.02	-	3.7
Fused data	2010	10	0.66±0.01	-	5.9

[a] In fused data line, this value refers to the gain in Kappa value for either ALOS/PALSAR (legend L1) or LANDSAT5/TM (legend L3) over the classification of the fused image using ML in the respective legend level.

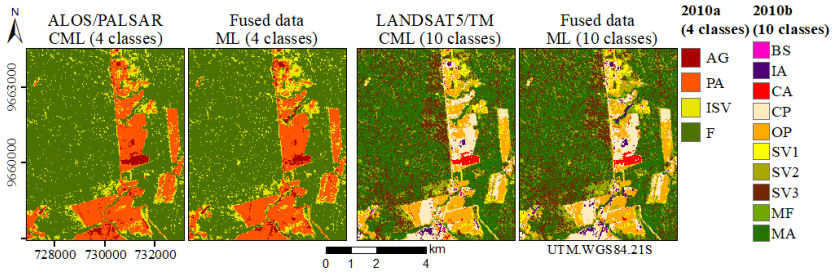

Figure 5. ALOS/PALSAR and LANDSAT5/TM images from 2010 classified using CML and the fused images classified with ML

5. Conclusions

A novel algorithm called Compound Maximum Likelihood (CML) was proposed in this work. This algorithm expands the traditional Maximum Likelihood (ML) classifier, by adding a multi-temporal *a priori* probability of class sequences. Therefore, information from different images and legends can be used jointly in image classification, that occurs in only one process. In CML, although errors occur, impossible transitions within class sequences are eliminated from analysis. For each image being analyzed, it means that if the class with the highest likelihood would return in a impossible transition, the next one with the highest likelihood resulting in a probable transition would be assigned to the pixel (or other analyzed object), which is not viable in hard algorithms like the traditional ML.

Additionally to solving some of the problems of traditional *post-classification comparison* change detection method, like eliminating impossible transitions from analysis, CML also incorporates the capacity of using information from multi-temporal set of images from *direct classification* methods without the need to acquire labeled samples from all class sequences in the data set. Since CML uses principles from both change detections methods but is not a properly *hybrid* method, it pertains to a new change detection category methods.

CML was used to classify three remote sensing images from three different years and in the same study area in Pará state, within the Brazilian Amazon. These were compared to images classified independently by ML. Different legends and types of images were used for each year. In this study case, all CML classified images presented higher accuracy index values than ML ones, with the more increased values pertaining to the classification of an ALOS/PALSAR image from 2008.

CML was first thought as a change detection algorithm, and as so is fully capable to classify a set of multi-temporal images and return class sequence vectors that can be then analyzed as change or no-change classes. It also has the potential to present the probability of the class sequence. Compared to ML applied independently to two or more images, the only additional feature necessary to use CML is the *a priori* probability of the class sequences occurring. As presented in this work, the simple indication of possible or impossible transitions is enough to improve land cover classification (improvement of 20% of the mean Kappa value for our test using ALOS/PALSAR data, for example). The possibility of transitions is dependent on the studied area, time being analyzed and class definition, but should be easily derived by the analyst who is familiar with the problem being studied.

Nonetheless, a methodology to use CML to classify images from proximate dates was presented in this work. Results were compared to the ones obtained by ML classification of the same data set and of fused data. In this case, CML presented either very similar or very improved results than the original data set, while all results for CML were better than those obtained by the classification of fused data. Besides the better classification results, in CML the fusion process is not necessary. It is also possible to use different legends for each data in the same classification process. Although there is the need to define the *a priori* probabilities of sequences, this process is relatively simple when considering data from the same date and legends with correspondent classes or group of classes.

A small data set was used in this work, considering a pixel based approach, to show this approach potential. The use of a much larger multi-temporal data sets has the potential to return even more improved results. Future work also include the application of CML in contextual and region based approaches.

Acknowledgments

Funded by CAPES and CNPq (grants #401528/2012-0 and #309135/2015-0). Authors are also thankful to the Monitoramento Ambiental por Satélite no Bioma Amazônia project, process #1022114003005-MSA-BNDES.

References

Anjos, D., Lu, D., Dutra, L., and Sant'Anna, S. (2015). Change detection techniques using multisensor data. In *Remotely Sensed Data Characterization, Classification, and Accuracies*, volume 1, pages 375–395. crc press, London.

Blaschke, T. (2005). Towards a framework for change detection based on image objects. *Göttinger Geographische Abhandlungen*, 113:1–9.

Chatelain, F., Tourneret, J. Y., and Inglada, J. (2008). Change detection in multisensor SAR images using bivariate gamma distributions. *IEEE Transactions on Image Processing*, 17(3):249–258.

Fuller, R., Smith, G., and Devereux, B. (2003). The characterisation and measurement of land cover change through remote sensing: problems in operational applications? *International Journal of Applied Earth Observation and Geoinformation*, 4(3):243 – 253.

Gómez, C., White, J. C., and Wulder, M. A. (2016). Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 116:55 – 72.

IBAMA (2004). Floresta Nacional do Tapajós plano de manejo: volume I - informações gerais. Technical report, Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis, Brazil. 76p.

Kennedy, R. E., Townsend, P. A., Gross, J. E., Cohen, W. B., Bolstad, P., Wang, Y., and Adams, P. (2009). Remote sensing change detection tools for natural resource managers: Understanding concepts and tradeoffs in the design of landscape monitoring projects. *Remote Sensing of Environment*, 113(7):1382 – 1396. Monitoring Protected Areas.

Lu, D., Mausel, P., Brondízio, E., and Moran, E. (2004). Change detection techniques. *International Journal of Remote Sensing*, 25(12):2365–2401.

Pereira, L. O. (2012). Avaliação de métodos de integração de imagens ópticas e de radar para a classificação do uso e cobertura da terra na região amazônica. Master's thesis, Instituto Nacional de Pesquisas Espaciais, São José dos Campos.

Prendes, J. (2015). *New statistical modeling of multi-sensor images with application to change detection*. PhD thesis, Universit'e Paris-Saclay, Toulouse.

Reis, M. S., Dutra, L. V., Sant'Anna, S. J. S., and Escada, M. I. S. (2017). Examining multi-legend change detection in amazon with pixel and region based methods. *Remote Sensing*, 9(1):Article 77.

Schowengerdt, R. (2006). *Remote sensing*. Academic Press, USA, 3 edition.

Singh, A. (1989). Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing*, 10(6):37–41.

Tewkesbury, A. P., Comber, A. J., Tate, N. J., Lamb, A., and Fisher, P. F. (2015). A critical synthesis of remotely sensed optical image change detection techniques. *Remote Sensing of Environment*, 160:1 – 14.

Torres, L., Sant'Anna, S. J., Freitas, C. C., and Frery, A. C. (2014). Speckle reduction in polarimetric SAR imagery with stochastic distances and nonlocal means. *Pattern Recognition*, 47(1):141–157.