

# A Comparison of Accuracy Measures for Remote Sensing Image Classification: Case Study In An Amazonian Region Using Support Vector Machine

**G. B. Scofield**

*Center for Natural Disaster Monitoring and Alert - CEMADEN  
São José dos Campos, São Paulo, 12630-000, Brazil*

*graziela.scofield@cemaden.gov.br*

**E. Pantaleão**

*Faculty of Computer, Federal University of Uberlândia - UFU  
Patos de Minas, Minas Gerais, 38700-126, Brazil*

*epantaleao@ufu.br*

**R. G. Negri**

*Environmental Engineering Department,  
São Paulo State University - UNESP  
São José dos Campos, São Paulo, 12247-004, Brazil*

*rogerio.negri@ict.unesp.br*

---

## Abstract

This work investigated the consistency of both the category-level and the map-level accuracy measures for different scenarios and features using Support Vector Machine. It was verified that the classification scenario and the features adopted have not influenced the accuracy measure consistency and all accuracy measures are highly positively correlated.

**Keywords:** Image Classification, Accuracy Measures, Category-level, Map-level, Comparison.

---

## 1. INTRODUCTION

In geoscience, Pattern Recognition methods have been shown useful for detecting targets in images obtained by satellites for many different purposes. Monitoring natural resources such as forests, rivers and glaciers [1; 2; 3], study areas affected by natural disasters [4] and urban planning and road monitoring for governmental purposes [5] are some examples. The Pattern Recognition techniques used to identify targets on images are called Image Classification.

Several accuracy measures are developed to quantify category-level and map-level accuracy measures based on the *confusion matrix*, which is computed from the Image Classification results. Each accuracy measure may be more relevant than others for a particular objective, since different measures incorporate different information about the confusion matrix.

In [6], the consistency of both category-level and map-level measures was investigated through the statistical correlation analysis of the confusion matrices using the probability of concordance. These confusion matrices were collected from previous studies presented in literature.

The objective of this work is to investigate the consistency of category-level and map-level accuracy measures, for image classification results with different numbers of classes. The image classification process was done using the Support Vector Machine (SVM) method and LISS-3 multi-spectral optical images, onboard of the Indian Remote Sensing Satellite (IRS).

## 2. ACCURACY MEASURES

Confusion matrix is one of the most effective ways to represent the Image Classification accuracy, being able to describe inclusion and exclusion errors. A generic error matrix is represented in Table 1, where for a given class  $\omega_i$ ,  $p_{ij}$  is the number of pixels from  $\omega_i$  sample that

was classified as  $\omega_j$ ,  $p_{i+}$  and  $p_{+i}$  denotes the marginal producer and user, respectively, to a given class  $\omega_i$ , and  $m$  is the total of validation pixels.

		User				
		$\omega_1$	$\omega_2$	...	$\omega_k$	
Producer	$\omega_1$	$p_{11}$	$p_{12}$	...	$p_{1k}$	$p_{1+}$
	$\omega_2$	$p_{21}$	$p_{22}$	...	$p_{2k}$	$p_{2+}$
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	$\omega_k$	$p_{k1}$	$p_{k2}$	...	$p_{kk}$	$p_{k+}$
		$p_{+1}$	$p_{+2}$	...	$p_{+k}$	$m$

TABLE 1: A Generic Error Matrix and Its Elements.

Several measures of thematic map accuracy were developed using the confusion matrix [7; 8; 9; 10]. These measures are organized into two groups: category-level and map-level accuracy measures. The category and map-level measures shown in Table 2 were analyzed.

	Measure name	Formula	Range
Category-level	User accuracy [11]	$ua_i = p_{ii}/p_{i+}$	[0,1]
	Producer accuracy [11]	$pa_i = p_{ii}/p_{+i}$	[0,1]
	User conditional kappa [9]	$uck_i = (ua_i - p_{+i})/(1 - p_{+i})$	]-∞, 0]
	Producer conditional kappa [9]	$pck_i = (pa_i - p_{i+})/(1 - p_{i+})$	]-∞, 0]
	Modified user conditional kappa [10]	$muck_i = (ua_i - \frac{1}{m})/(1 - \frac{1}{m})$	[-1,1]
	Modified producer conditional kappa [10]	$mpck_i = (pa_i - \frac{1}{m})/(1 - \frac{1}{m})$	[-1,1]
Map-level	Overall accuracy [8]	$oa = \sum_{i=1}^m p_{ii}$	[0,1]
	Average accuracy user perspective [11]	$aa_u = \frac{1}{m} \sum_{i=1}^m ua_i$	[0,1]
	Average accuracy producer perspective [11]	$aa_p = \frac{1}{m} \sum_{i=1}^m pa_i$	[0,1]
	Double average user and producer accuracy [12]	$daup = \frac{aa_u - aa_p}{2}$	[0,1]
	Average of Hellden mean accuracy index [13]	$amah = \frac{1}{m} \sum_{i=1}^m \frac{2p_{ii}}{p_{i+} + p_{+i}}$	[0,1]
	Average of Short mapping accuracy index [13]	$ams = \frac{1}{m} \sum_{i=1}^m \frac{p_{ii}}{p_{i+} + p_{+i} - p_{ii}}$	[0,1]
	Combining accuracy from both user and producer perspectives [11]	$caup = \frac{aa_u - amah}{2}$	[0,1]
	Kappa [7]	$kap = \frac{oa - \sum_{i=1}^m (p_{i+} + p_{+i})}{oa - \sum_{i=1}^m (p_{i+} + p_{+i})}$	]-∞, 1]
	Modified kappa [12]	$mkap = (oa - \frac{1}{m})/(1 - \frac{1}{m})$	]-∞, 1]
	Average mutual information [14]	$ami = \frac{1}{m} \sum_{i,j=1}^m p_{ij} \frac{p_{ii}}{p_{i+} + p_{+i} - p_{ii}}$	[-∞, 1[

TABLE 2: Category-level and Map-level Accuracy Measures [6].

The overall accuracy is the sum of the total correct pixels divided by the total number of pixels, since the accuracies of individual classes are calculated by the producer accuracy. The producer accuracy gives the probability of a pixel being correctly classified. On the other hand, the user accuracy indicates the probability that a pixel on the map represents the same class on the ground [11]. In [15], the averages of the user and producer accuracy are proposed.

The Kappa coefficient, introduced by [7], is a widely used accuracy measure for classification assessment. Kappa may be used as a measure of agreement between prediction and reality, or to determine if the values in a confusion matrix represent a significantly better than a randomly obtained one [16]. Several modifications on Kappa coefficient, for category and map-level assessment, have been suggested. The Conditional Kappa [9] and the Modified Conditional Kappa [10] have been presented as the additional means of incorporating an adjustment for hypothetical chance agreement into the category-level accuracy assessment. Among many other map-level measures, the Average Accuracy of the user, the producer, and the combination of both perspectives [11], Average of Short Mapping Accuracy Index [12], Hellden's Mean Accuracy Index [13] and the Average Mutual Information [14] may be mentioned.

In [6], it is conducted a consistent analysis of accuracy measures using Kendall's Tau ( $K_\tau$ ), which is the difference between percentage of concordance ( $PC$ ) and the percentage of discordance ( $PD$ ) of two accuracy measures computed on a collection of confusion matrices. Formally, given  $n$  confusion matrices ( $M_1, M_2, \dots, M_n$ ) and two accuracy measures ( $a$  and  $b$ ), a series of pairs is determined from these  $n$  confusion matrices, that is:  $(a_1, b_1), (a_2, b_2), \dots, (a_n, b_n)$ . For two specific confusion matrices  $M_i$  and  $M_j$ , if either  $a_i > a_j$  and  $b_i > b_j$  or  $a_i < a_j$  and  $b_i < b_j$ , there is a *concordant pair*; if either  $a_i > a_j$  and  $b_i < b_j$  or  $a_i < a_j$  and  $b_i > b_j$ , there is a *discordant pair*. For  $n$  confusion matrices,  $T = n(n - 1)/2$ , for  $1 \leq i, j \leq n$ , is the number of pairs of possible matrix combinations. Denoting  $C$  and  $D$  as the number of concordance and discordance pairs,  $PC$  and  $PD$  are the concordance and discordance percentage, written respectively, as  $C/T$  and  $D/T$ .

### 3. METHODOLOGY

The adopted methodology consists of four steps: definition of a classification problem; attributes extraction from an adopted study image; classification of the selected image; calculation and comparison of accuracy measures. The image and the adopted study area are shown in Subsection 3.1, Subsection 3.2 presents the adopted image classification method and the attributes extracted from the study image to process the classifications are mentioned in Subsection 3.3. For each classification result, the different accuracy measures presented in Table 2 are computed with basis on validation samples (Table 3). It is worth mentioning that the aim of this study is to investigate the assessment measures previously discussed and not the obtained classification results.

It is worth mention that in this study, different from [6] where confusion matrices were observed from previously studies through diverse images and classification methods, the confusion matrices were obtained using a particular image classification method, various features extracted from a selected image and considering distinct classes (classification problems).

#### 3.1 Data and Study Area

The study area corresponds to a region near to Tapajós National Forest (FLONA), Brazil. The FLONA is a federal conservation unit having approximately 544 thousand hectares. This unit is bounded by the urban area of Belterra at North, by the BR-163 highway at East, by the Tapajós River at West and by the Tinga and Cupuari rivers at South. It is also characterized by wood extraction that occurred over the past two decades, provoking significant formation of secondary forest areas, concentrated mainly along the BR-163 highway. Additionally, this area has received governmental incentive for mechanized grain production, aiming to export in the first decade of 2000 [17].

Examples of land cover samples collected on the study area are shown in Figure 1. These samples were verified in a field work performed during September 2009. Fortunately, the IRS LISS-3 multispectral sensor (23.5 meters of resolution) was able to acquire images over the FLONA area with low cloud coverage in September 10<sup>th</sup>, 2009. The collected samples were used to train the classification method and to validate the results. For each class, two sets of samples were randomly defined, where the first one is responsible for training and another for validation. The training set is approximately twice larger than the validation one. Table 3 summarizes the quantity of pixels adopted to train the classification method and to validate its results.

Based on the field work information, a hierarchical class tree was organized, as shown in Figure 2. This arrangement allows to define four scenarios, that is, four ways to partitionate the set of classes. The first scenario is the most specific, since it uses all the tree leaves. The complete tree presents 15 classes, which are: cloud, cloud shadow, dirty pasture, clean pasture, agriculture with high grass, abandoned agriculture, primary forest, secondary forest with 1 to 5 years, secondary forest with 6 to 12 years, secondary forest with 13 to 30 years, urban area, soil with rest of agriculture, prepared soil for agricultural use, bare soil and water (river). Superclasses were used to generate the second scenario, which are pasture, agriculture use and secondary forest, together with cloud, cloud shadow, bare soil, primary forest, urban area, agriculture with high grass and water, completing 10 classes. The term “superclass” determines that a class is composed by the union of similar classes. The third scenario is composed by 8 superclasses which are cloud, cloud shadow, urban area, water, soil, pasture, agriculture and forest. The fourth scenario gathers superclasses cloud, cloud shadow, urban area, water, soil and vegetation.

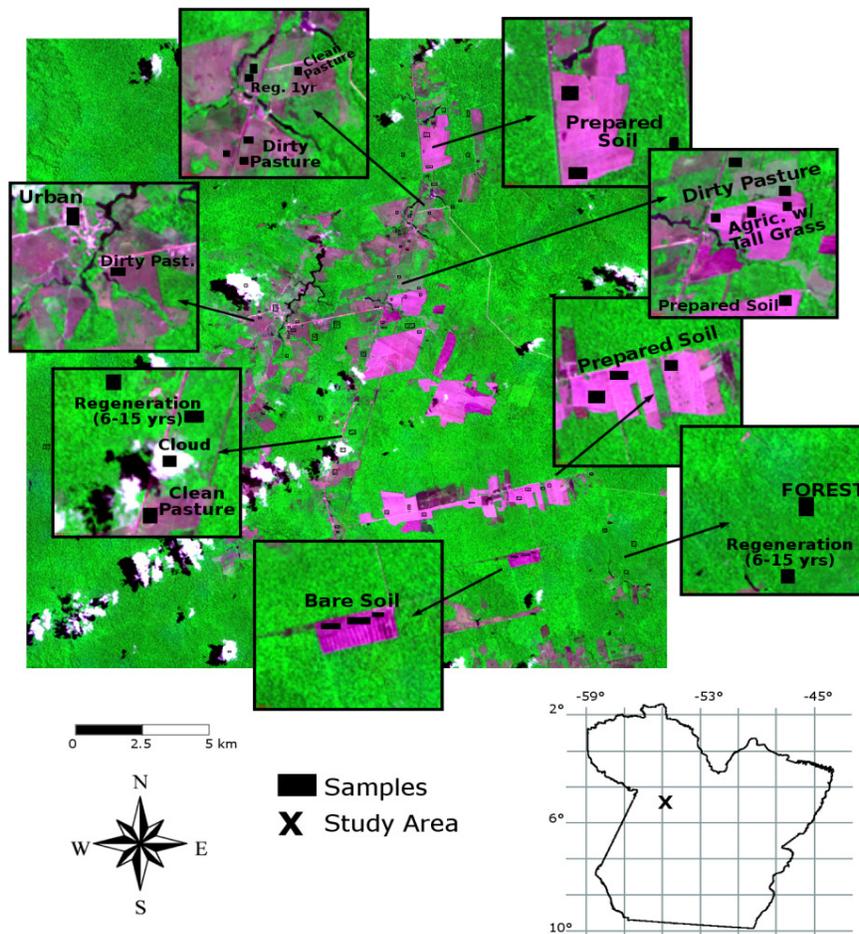
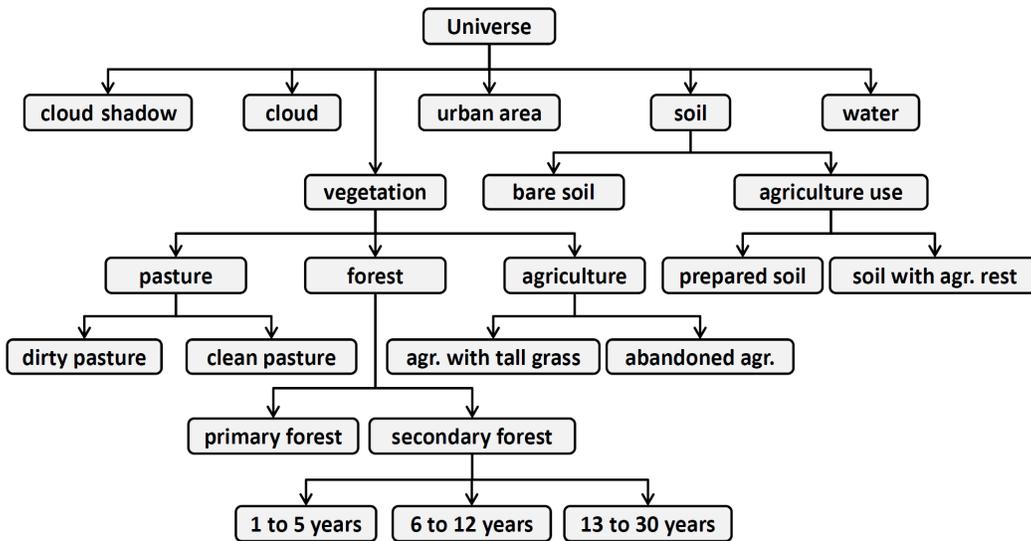


FIGURE 1: Study Area and Land Cover Samples.

Classes	Ground Truth Pixels	
	Training	Validation
Cloud	205	96
Cloud Shadow	117	53
Dirty Pasture	246	126
Clean Pasture	397	185
Agriculture with High Grass	34	14
Abandoned Agriculture	60	28
Primary Forest	115	61
Secondary Forest (1 to 5 years)	80	37
Secondary Forest (6 to 12 years)	191	99
Secondary Forest (13 to 30 years)	23	9
Urban Area	81	40
Soil with Rest of Agriculture	59	30
Prepared Soil for Agriculture	37	16
Bare Soil	26	12
Water (River)	28	12

**TABLE 3:** A summary about the quantity of ground truth pixels used to train the classification method and to validate the classification results.



**FIGURE 2:** Hierarchical Class Tree Used In This Study.

### 3.2 Support Vector Machine

The Support Vector Machine (SVM) method was adopted in this study. SVM is a recent technique which has received great attention in recent years due to its excellent generalization ability, data distribution independence and its robustness on the Hughes phenomena.

This method consists of finding a separation hyperplane between the training samples with the larger margins. The separating hyperplane is the geometric place where the following linear function is zero:

$$f(\mathbf{x}) = \langle \mathbf{x}, \mathbf{w} \rangle + b \tag{1}$$

where  $\mathbf{w}$  represents the orthogonal vector to the hyperplane  $f(\mathbf{x}) = 0$ ;  $b/\|\mathbf{w}\|$  is the distance from the hyperplane to the origin and  $\langle \cdot, \cdot \rangle$  denotes the inner product. The parameters of (1) are obtained from the following quadratic optimization problem [18]:

$$\sum_{i=1}^m \lambda_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \lambda_i \lambda_j y_i y_j \langle \varphi(\mathbf{x}_i), \varphi(\mathbf{x}_j) \rangle$$

$$\text{subject to: } \begin{cases} 0 \leq \lambda_i \leq C; \quad i = 1, \dots, m \\ \sum_{i=1}^m \lambda_i y_i = 0 \end{cases} \quad (2)$$

where  $\lambda_i$  are Lagrange multipliers,  $y_i \in \{-1, +1\}$  define the class of  $\mathbf{x}_i$ , since SVM is a binary classifier,  $C$  acts as an upper bound of  $\lambda$  values and  $\varphi(\mathbf{x})$  is a function adopted to remap the input vectors into a higher dimensionality space. The inner product  $\langle \varphi(\mathbf{x}_i), \varphi(\mathbf{x}_j) \rangle$  is known as Kernel function. A popular example of Kernel is the Radial Basis Function (RBF), expressed by  $\langle \varphi(\mathbf{x}_i), \varphi(\mathbf{x}_j) \rangle = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / 2\sigma^2)$ ,  $\sigma \in \mathbb{R}_+$ , which was adopted in this study. The parameters  $C$  and  $\sigma$  tuning was performed according to a Grid Search procedure considering as values  $C = \{1, 10, 100, 1000\}$  and  $\sigma = \{0.5, 1.0, 1.5, 2.0, 2.5\}$ .

The optimization problem (2) is solved considering a training set  $\mathcal{D} = \{(\mathbf{x}_i, y_i) : i = 1, \dots, l\}$ , where  $\mathbf{x}_i \in \mathbb{R}^d$ . Let  $SV = \{\mathbf{x}_i : \lambda_i \neq 0; i = 1, \dots, l\}$ , known as support vector set. The parameters  $\mathbf{w}$  and  $b$  are computed by:

$$\mathbf{w} = \sum_{\mathbf{x}_i \in SV} \lambda_i y_i \varphi(\mathbf{x}_i) \quad (3)$$

$$b = \frac{1}{\#SV} \sum_{\mathbf{x}_i \in SV} y_i + \sum_{i=1}^l \sum_{j=1}^l \lambda_i \lambda_j y_i y_j \langle \varphi(\mathbf{x}_i), \varphi(\mathbf{x}_j) \rangle \quad (4)$$

To apply this method to a multiclass problem (problems with more than two classes), the adoption of a multiclass strategy is necessary, such as “one-against-all” or “one-against-one”. Here, the “one-against-all” strategy was adopted. For more details about multiclass strategies, please refer to [19]. The ENVI 4.7 software was used to perform the SVM classifications.

### 3.3 Feature Extraction and Classification

Several features were extracted from the multispectral IRS LISS-3 image using the ENVI software. These features were obtained using the LISS-3 Red (R), Green (G), Blue (B) and Near Infrared (IRNear) spectral bands, the Normalized Difference Vegetation Index (NDVI) [20] and the Haralick’s Texture Features [21] such as mean (M), variance (Var), homogeneity (H), contrast (C), dissimilarity (D), entropy (E), second moment (SM), and correlation (CR).

The Principal Component Analysis (PCA) [22] was applied to extract features optimally uncorrelated, and the Minimum Noise Fraction technique (MNF) [23] was applied to reduce the data dimensionality and isolate the noise. Combinations of these features were used on SVM to classify each set of classes mentioned in the Subsection 3.1. Finally, the consistency of the accuracy measures were computed and then their consistency were investigated.

The next step was to investigate the consistency of category-level and map-level accuracy measures through statistical correlation analysis. Figure 3 shows the activities flow chart.

Initially, the SVM was applied on the Red, Green, Blue and Near Infrared (R+G+B+IRNear) bands. The NDVI information was added to Red, Green, Blue and Near Infrared bands (NDVI+R+G+B+IRNear) for the second classification.

The PCA was applied in the Red, Green, Blue, Near Infrared and NDVI features and the first three principal components (PCA 1, 2 and 3 respectively) were used on the third classification. The fourth classification uses all features together, that is, the spectral, NDVI and texture (denoted by NDVI+R+G+B+IRNear+M+Var+H+C+E+D+SM+CR). The PCA was applied on the eight texture bands (i.e., M+Var+H+C+D+E+SM+CR) and the first three PCA components were used for classification. The three first PCAs for NDVI+R+G+B+IRNear+E+D+SM were also classified.

The features entropy (E), dissimilarity (D) and second moment (SM) were chosen through visual analysis, since they presented more information in the contrast between different classes. The NDVI+R+IRNear bands are selected as a color composition to generate a HSV color transformation [22]. This HSV transformation was used for classification. Finally, the MNF transform was applied on the NDVI, R, G, B and IRNear features and the first three produced features (denoted by MNF1, 2 and 3) were used for classification.

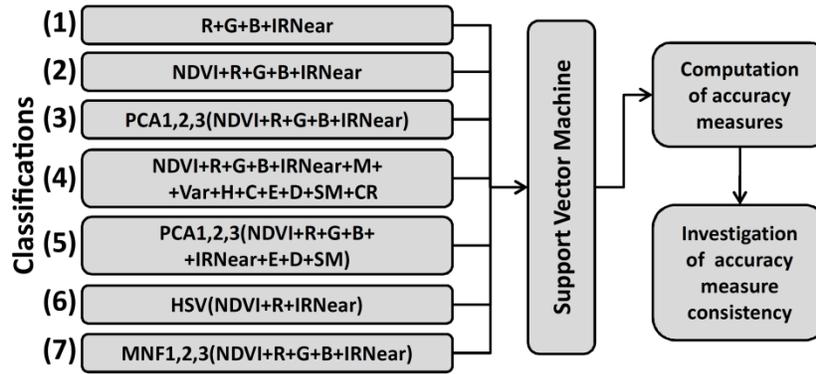


FIGURE 3: Activities Flow Chart.

Each combination of features, illustrated in Figure 3, was used to classify each of the four set of classes. As a result, 28 confusion matrices were obtained. Afterwards, these matrices were used to compute the accuracy measures.

#### 4. RESULTS

For each of the four scenarios defined on Subsection 3.1, SVM classifications using each one of the seven feature combinations determined in Subsection 3.3 were produced. So, 28 classification results and its respective confusion matrices were obtained. From the confusion matrices, the accuracy measures presented in Table 2 were calculated (6 category-level and 10 map-level measures). To summarize a comparison among category-level measures, the  $PC$  and  $PD$  (percentage of concordance and discordance) were calculated over all classes of each scenario, and then  $K_{\tau}$  was calculated. These results are shown in Table 4. Table 5 presents the  $K_{\tau}$  between each pair of map-level accuracy measure.

Analyzing the results, the  $K_{\tau}$  varied between 60% to 100% on Scenario 1, 34% to 100% on Scenario 2, and 48% to 100% on Scenarios 3 and 4. For the pairs  $ua - uck$ ,  $ua - muck$  and  $uck - muck$ ,  $K_{\tau}$  were kept equal to 100% in all Scenarios, which means that every user measure may be used equivalently. For the producer measures the same result was observed only for  $pa - mpck$ . The observed  $K_{\tau}$  for the other producer measures was above 82%.

As expected,  $K_{\tau}$  was not so different when users and producers measures are compared. Observing the results on different scenarios, most of the category-level measures showed less correlation for the second scenario (10 classes).

As can be seen in the map-level measures (Table 5),  $ami$  presents higher percentage of discordance compared with other measures. The  $ami$  uses all the elements of confusion matrix,

differently from other measures that use just the class ( $p_{ii}$ ), producer ( $p_{i+}$ ) and user ( $p_{+i}$ ) accuracies, which explains this behavior. It may also be observed that *oa*, *cap*, *kap* and *mkap* are perfectly concordant. Except for *ami*, the measures present more than 90% of concordance.

The map-level measures, computed from each one of the 28 confusion matrices, are represented on the line graphic of Figure 4. The confusion matrices are organized as follows: 1, 8, 15 and 22 refers for classification 1 of the scenarios 1, 2, 3 and 4, respectively; 2, 9, 16 and 23 for classification 2; 3, 10, 17 and 24 for classification 3; 4, 11, 18 and 25 for classification 4; 5, 12, 19 and 26 for classification 5; 6, 13, 20 and 27 for classification 6; and 7, 14, 21 and 28 for classification 7.

Based on the accuracy levels, it was possible to verify that the analyzed map-level measures have very similar results. The fourth set of classes (22<sup>nd</sup> to 28<sup>th</sup> confusion matrices) presented an increase in the accuracy measures in comparison with the other sets. This is explained by the class configuration, since a less specific scenario tends to provide less complexity on classification problems.

Scenario 1	<i>pa</i>	<i>uck</i>	<i>pck</i>	<i>muck</i>	<i>mpck</i>	Scenario 2	<i>pa</i>	<i>uck</i>	<i>pck</i>	<i>muck</i>	<i>mpck</i>
<i>ua</i>	.68	1.0	.60	1.0	.68	<i>ua</i>	.41	1.0	.34	1.0	.41
<i>pa</i>		.68	.82	.68	1.0	<i>pa</i>		.41	.84	.41	1.0
<i>uck</i>			.60	1.0	.68	<i>uck</i>			.34	1.0	.41
<i>pck</i>				.60	.82	<i>pck</i>				.34	.84
<i>muck</i>					.68	<i>muck</i>					.41

Scenario 3	<i>pa</i>	<i>uck</i>	<i>pck</i>	<i>muck</i>	<i>mpck</i>	Scenario 4	<i>pa</i>	<i>uck</i>	<i>pck</i>	<i>muck</i>	<i>mpck</i>
<i>ua</i>	.48	1.0	.49	1.0	.48	<i>ua</i>	.48	1.0	.56	1.0	.48
<i>pa</i>		.48	.94	.48	1.0	<i>pa</i>		.48	.92	.48	1.0
<i>uck</i>			.49	1.0	.48	<i>uck</i>			.56	1.0	.48
<i>pck</i>				.49	.94	<i>pck</i>				.56	.92
<i>muck</i>					.48	<i>muck</i>					.48

**TABLE 4:**  $K_{\tau}$  between each category-level accuracy measure pair: Scenario 1 (15 classes), Scenario 2 (10 classes), Scenario 3 (8 classes) and Scenario 4 (6 classes).

	<i>aau</i>	<i>aap</i>	<i>daup</i>	<i>amah</i>	<i>ams</i>	<i>caup</i>	<i>kap</i>	<i>mkap</i>	<i>ami</i>
<i>oa</i>	.889	.778	.836	.815	.825	.915	.968	1.00	-.640
<i>aau</i>		.825	.894	.884	.873	.921	.857	.889	-.593
<i>aap</i>			.931	.942	.931	.862	.767	.778	-.450
<i>daup</i>				.979	.968	.921	.804	.836	-.508
<i>amah</i>					.989	.899	.794	.815	-.497
<i>ams</i>						.910	.804	.825	-.508
<i>caup</i>							.884	.915	-.556
<i>kap</i>								.968	-.619
<i>mkap</i>									-.640

**TABLE 5:** Results of  $K_{\tau}$  between each map-level accuracy measure pair.

## 5. CONCLUSIONS

The objective of this study was to evaluate the consistency of measures used to quantify the accuracy of image classification results. The SVM method was applied in different classification scenarios using different features extracted from the IRS multi-spectral images. Results showed that measures, for category or map-level accuracy, are consistent with every classification scenario and the adopted features. This consistency is expected since all analyzed measures are computed from the same information source, which is the confusion matrix.

User measures are equivalent for category-level assessment, since they present 100% of concordance. However, the producer measures are not equivalent, but show a high level of concordance (at least 82%). In general, map-level accuracies are highly correlated. Lower levels of concordance are verified between the Average Mutual Information measure which every other measures.

Unlike most of the measures based on class accuracy and the marginal user and producer, Average Mutual Information uses all confusion matrix information.

According to the results, the analyzed accuracy measures reflect a common behavior for most situations. None of them can be used to mask or highlight certain results, and they do not lead to divergent conclusions.

As future perspectives, this investigation must be reproduced using others methodologies for scenario definition, instead use an intuitive scenario definition, as performed in this study. Additionally, different features and other image classification methods, more/less robust than SVM, must be considered too.

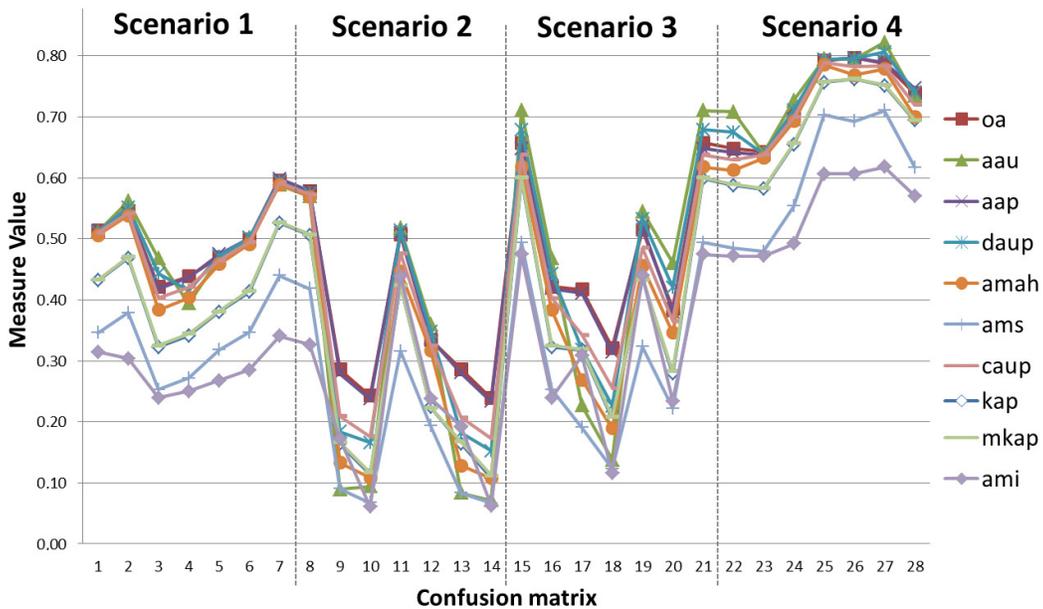


FIGURE 4: Map-level measures computed for each confusion matrix.

## 6. Acknowledgments

The authors would like to thanks CAPES, CNPQ (3284469/2006-0) and FAPESP (2014/14830-8) for the financial support.

## 7. REFERENCES

- [1] Y. E. Shimabukuro, R. Almeida-Filho, T. M. Kuplich, R. M. Freitas, "Mapping and monitoring land cover in Corumbiara area, Brazilian Amazonia, using JERS-1 SAR multitemporal data", IEEE International Geoscience and Remote Sensing Symposium, Barcelona, 2007.
- [2] A. Niedermeier, S. Lehner, J. Sander, "Monitoring big river estuaries using SAR images", IEEE International Geoscience and Remote Sensing Symposium, Sydney, 2001.
- [3] K. Whitehead, B. Moorman, P. Wainstein, "Determination of variations in glacier surface movements through high resolution interferometry: Bylot Island, Canada", IEEE International Geoscience and Remote Sensing Symposium, Cape Town, 2009.
- [4] A. Chesnel, R. Binet, L. Wald, "Object oriented assessment of damage due to natural disaster using very high resolution images", IEEE International Geoscience and Remote Sensing Symposium, Barcelona, 2007.
- [5] M. Kasanko, V. Sagris, C. Laval, J. I. Barredo, L. Petrov, K. Steinnocher, W. Loibl, C. Hoffmann, "GEOLAND spatial planning observatory: How remote sensing data can serve the needs of urban and regional planning", Urban Remote Sensing Joint Event, Paris, 2007.
- [6] C. Liu, P. Fraizer, L. Kuman, "Comparative assessment of the measures of thematic classification accuracy," Remote Sensing of Environment. vol. 107, pp. 606-616, 2007.
- [7] J. A. Cohen, "Coefficient of agreement of nominal scales," Educational and Psychological Measurement. vol. 20, pp. 37-46, 1960.
- [8] M. Story, R. G. Congalton. "Accuracy assessment: a user's perspective," Photogrammetric Engineering and Remote Sensing. vol. 52, pp. 397-399, 1986.
- [9] G. H. Rosenfield, K. A. Fitzpatrick-Lins, "Coefficient of agreement as a measure of thematic classification accuracy," Photogrammetric Engineering and Remote Sensing. vol. 52, pp. 223-227, 1986.
- [10] S. V. Stehman, "Selecting and interpreting measures of thematic classification accuracy," Remote Sensing of Environment. vol. 62, pp. 77-89, 1997.
- [11] T. Fung, E. Ledrew, "The determination of optimal threshold levels for change detection using various accuracy indices," Photogrammetric Engineering and Remote Sensing. vol. 54, pp. 1449-1454, 1988.
- [12] N. M. Short, "The LANDSAT tutorial workbook: basics of satellite Remote Sensing". National Aeronautics and Space Administration, Scientific and Technical Information Branch, 1982.
- [13] U. A. Hellden, "Test of LANDSAT-2 imagery and digital data for thematic mapping illustrated by an environmental study in northern Kenya", Lund University Natural Geography Institute Report, 1980.
- [14] J. T. Finn, "Use of the average mutual information index in evaluating classification error and consistency," International Journal of Geographical Information Systems. vol. 7, pp. 349-366, 1993.
- [15] R. G. Congalton, "A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data," Remote Sensing of Environment. vol. 37, pp. 35-46, 1991.

- [16] J. R. Jansen, *Introductory Digital Image Processing: A Remote Sensing Perspective*, Pearson Prentice Hall, Upper Saddle River, 2005.
- [17] ZEE BR-163, 2011. Zoneamento ecológico-econômico da rodovia BR-163, Access in December 2011. <http://zeebr163.cpatu.embrapa.br>.
- [18] S. Theodoridis, K. Koutrombas, *Pattern Recognition*, Academic Press, San Diego, 2006.
- [19] A. R. Webb. *Statistical Pattern Recognition*, Jhon Wiley and Sons, Chichester, 2002.
- [20] J. G. P. W. Clevers, "The derivation of a simplified reflectance model for the estimation of leaf area index," *Remote Sensing of Environment*. vol. 35, pp. 53-70, 1988.
- [21] R. M. Haralick, K. Shanmugam, I. Dinsten, "Texture features for image classification," *IEEE Transactions on Systems, Manchine and Cybernetics*. vol. 3, pp. 610-622, 1973.
- [22] R. C. Gonzales, R. E. Woods. *Digital Image Processing*, Prentice Hall, California, 2007.
- [23] A. A. Green, M. Berman, P. Switzer, M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Transactions on Geoscience and Remote Sensing*. vol. 26, pp. 65-74, 1998.