ASSESSING THE ACCURACY OF LAND USE LAND COVER (LULC) MAPS USING CLASS PROPORTIONS IN THE REFERENCE DATA

C. C. Fonte [1,2,*], L. See [3], J. C. Laso-Bayas [3], M. Lesiv [3], S. Fritz [3]

[1] University of Coimbra, Department of Mathematics, Largo D. Dinis, Coimbra, Portugal - cfonte@mat.uc.pt
[2] Institute for Systems Engineering and Computers at Coimbra - INESC Coimbra, Rua Sílvio Lima, Pólo II, Coimbra, Portugal
[3] International Institute for Applied Systems Analysis, Schlossplatz 1, Laxenburg, Austria – (see, lasobaya, lesiv, fritz)@iiasa.ac.at

Commission III, WG III/IVb

KEY WORDS: Land Use Land Cover, Accuracy Assessment, Reference Data, Class Proportions, Majority Class

ABSTRACT:

Traditionally the accuracy assessment of a hard raster-based land use land cover (LULC) map uses a reference data set that contains one LULC class per pixel, which is the class that has the largest area in each pixel. However, when mixed pixels exist in the reference data, this is a simplification of reality that has implications for both the accuracy assessment and subsequent applications of LULC maps, such as area estimation. This paper demonstrates how the use of class proportions in the reference data set can be used easily within regular accuracy assessment procedures and how the use of class proportions can affect the final accuracy assessment. Using the CORINE land cover map (CLC) and the more detailed Urban Atlas (UA), two accuracy assessments of the raster version of CLC were undertaken using UA as the reference and considering for each pixel: (i) the class proportions retained from the UA; and (ii) the class with the majority area. The results show that for the study area and the classes considered here, all accuracy indices decrease when the class proportions are considered in the reference database, achieving a maximum difference of 16% between the two approaches. This demonstrates that if the UA is considered as representing reality, then the true accuracy of CLC is lower than the value obtained when using the reference data set that assigns only one class to each pixel. Arguments for and against using class proportions in reference data sets are then provided and discussed.

1. INTRODUCTION

1.1 Accuracy assessment of LULC maps

Land use land cover (LULC) maps are usually produced through the classification of satellite imagery. If a pixel-based approach is considered, the classification assigns one of the classes in the nomenclature adopted by the user to each pixel in the image. The assignment of the class to a pixel can be undertaken using many different classification approaches. Depending on the approach chosen, this may result in the assignment of different classes to some of the same pixels, even when the same training areas are used (Li et al., 2014; Mather & Tso, 2016). Moreover, for maps with a low or medium spatial resolution, many of the pixels may, in reality, include several classes, i.e., they are mixed pixels. In this case, the class that occupies the largest proportion of the pixel is usually assigned to that pixel (Li et al., 2014). Because of all these difficulties, the accuracy assessment of the per-pixel based LULC maps is a very important step in the map production, as it will define its usability in different applications, ensuring that the map is of sufficient quality for each application.

Object-based approaches may also be used to produce LULC maps. However, the accuracy assessment of object-based and per-pixel based LULC maps should be done using different approaches, as not only the thematic aspects need to be validated in the former approach, but also the validity of the generated objects (Ye et al., 2018; Tiede et al., 2010; Albrecht et al., 2010). Therefore, in this paper our focus is on the accuracy assessment of LULC maps based on a per-pixel approach.

For these, the accuracy assessment is usually done by selecting a sample of spatial units, frequently based on individual pixels, with a pre-defined sampling protocol (Stehman, 2009). The "true LULC class" is then identified, for example, by photo interpretation of very high-resolution images and/or field visits and assigned to these sample units to create a reference data set. A comparison of the map class and the reference class is then made for all sample units by creating a confusion matrix, from which accuracy indices can be extracted, such as the user's and producer's accuracy per class, and the overall accuracy (e.g. Olofsson et al., 2014; Stehman & Foody, 2019).

The quality of the reference database is, thus, crucial for accuracy assessment, and may in fact have a large influence on the accuracy indices obtained (Foody, 2010; McRoberts et al., 2018). Therefore, the definition of the procedures and methodologies used to obtain the "true LULC class", known as the reference condition, is very important.

1.2 Creation of reference data

The most common method used to assess the accuracy of a crisp raster-based LULC map is to select one class for each spatial unit of the reference database (e.g. Olofsson et al., 2014; Stehman & Foody, 2019). However, this task is not easy and there are many difficulties and sources of uncertainty in this process, including the unavailability of very high-resolution imagery in some regions for performing photo interpretation (Lesiv et al., 2018), the subjectivity of class selection (McRoberts et al., 2018) and the fact that some of the pixels are mixed pixels and include regions that correspond to different classes (Stehman & Foody, 2019).

---

* Corresponding author

Other approaches have already been suggested, such as selecting a primary and secondary class (*e.g.* Woodcock et al., 1996), the use of a linguistic scale to express the correctness of assigning each LULC class to each spatial unit and then using a fuzzy approach to generate accuracy indicators (Gopal & Woodcock, 1994; Sarmento et al., 2013, 2015; Woodcock & Gopal, 2000), or the assignment of fuzzy class memberships based on the desired or target output of the classification and computing the distance between these values and the outputs of the classification (Foody & Arora, 1996). However, these methodologies have not been applied frequently. The approach most frequently used is still the selection of the class occupying the majority of the spatial unit. When using this approach to assess the accuracy of a crisp raster LULC map, the accuracy assessment is, in fact, evaluating if the LULC map was produced according to the specifications, independent of the minimum mapping unit (MMU) (which may also be included in the assessment if an analysis of the pixel neighbourhood is made to determine the surrounding LULC), i.e., if the class assigned to each pixel is, in reality, the class occupying the majority of the pixels. However, even though it is usually stated that the reference database should report the "true LULC class", this approach requires a simplification of reality, as mixed pixels are converted to pixels corresponding to only one class. Hence, this does not assess the accuracy of the map in relation to "reality" but a simplified version of reality.

1.3 Aim of this study

The aim of this paper is to propose that the accuracy assessment of raster-based LULC maps, in particular those with low and medium spatial resolutions, should be performed with reference databases that include further information about reality, by identifying, for each pixel, the proportion of that pixel that is occupied by each of the classes in the chosen nomenclature. The use of such a reference database to assess the accuracy of crisp LULC maps requires an adaptation to the procedures used to create the confusion matrices. However, such adaptations can be easily done. In fact, they have already been proposed for assessing the accuracy of soft classifications (Pontius Jr, & Cheuk, 2006), but can be used to assess crisp LULC maps with reference data that report more than one class per spatial unit.

To illustrate the use of such an approach, the methodology is applied here to the validation of the 2012 CORINE land cover (CLC) map using the more detailed Urban Atlas (UA) LULC map for 2012, for a study area located in continental Portugal. As the reference data set (UA) is available for all of the study area in this situation, no sampling in used. This allows us to compare the differences in the outputs of the validation when using reference data in which the class is chosen that occupies the majority of the pixel as well as the proportions of several classes that exist in each pixel.

The results show that for the study area and the classes considered here, all accuracy indices decrease when the class proportions are considered in the reference database, achieving a maximum difference of 16%. This demonstrates that if you consider the UA as representing reality, then the true accuracy of CLC is lower than the value obtained when using the reference data set that assigns only one class to each pixel. These differences also influence the area estimation of the classes, which is one of the most relevant pieces of information that frequently needs to be estimated from LULC maps. This shows the relevance of this subject and the need to develop methods that are easy to use but provide more reliable accuracy estimations,

and that assess the quality of the maps in relation to "real" world conditions.

2. STUDY AREA AND DATA

2.1 Study area

The selected study area was a region around Coimbra city (Portugal) with a total area of 734.4 km^2. The region includes the city of Coimbra, with around 100 000 inhabitants, at its centre, a surrounding area that includes smaller villages, agricultural fields as well as forested regions, and a section of the Mondego River (Figure 1).

Figure 1. Study area: a region in Portugal containing the city of Coimbra

2.2 Data

The data used in this study are: 1) The 2012 version of the CORINE Land Cover (CLC) product in raster format with 100 m spatial resolution and 2) the 2012 version of Urban Atlas (UA). The CLC is a Pan-European product initiated in 1985 that has updates for 2000, 2006, 2012 and 2018. The nomenclature of CLC includes 44 classes hierarchically structured into 3 levels of detail. It has an MMU of 25 ha for areal phenomena and a minimum width of 100 m for linear phenomena. Level 1 (the most general level) includes the 5 classes described in Table 1.

Class (Code - name)	Class description
1-Artificial surfaces	Land with dominant influence and without agricultural land use - includes all artificial structures and their associated non-sealed and vegetated surfaces
2-Agricultural areas	Agricultural, semi-natural areas, including arable land, permanent crops, pastures, orchards, complex and mixed cultivation patterns
3-Forest and semi natural areas	Forests, herbaceous vegetation and open spaces with little or no vegetation
4-Wetlands	Inland and coastal wetlands
5-Water bodies	Inland and marine waters

Table 1. Level 1 classes from the CLC and UA nomenclatures, which are the classes used in this study

UA is a product available only for urban areas in Europe. Versions for 2006 and 2012 are available for 319 and 785 urban areas in Europe, respectively. UA nomenclature includes 27 classes structured into 4 hierarchical levels. The level 1 nomenclature of UA is identical to the CLC nomenclature (Table 1), which is the one used in this study. The urban classes

of UA have an MMU of 0.25 ha and the rural classes have an MMU of 1 ha.

Figures 2 and 3 show the level 1 classes of CLC (2012) and UA (2012) for the study area, respectively. Table 2 shows the area occupied in the study area by each class in the CLC and UA, in km^2 and as a percentage of the total study area.
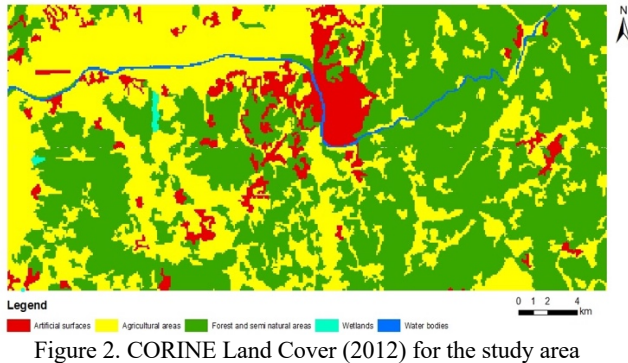

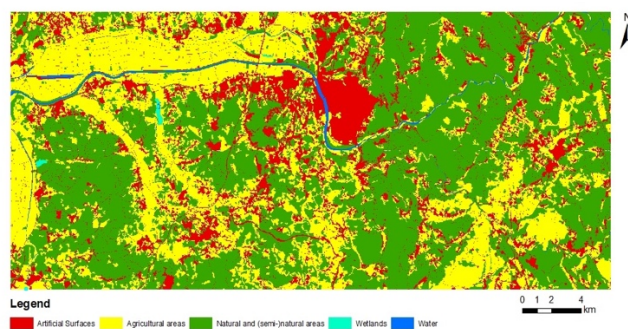Figure 2. CORINE Land Cover (2012) for the study area

Figure 3. Urban Atlas (2012) for the study area

Classes Code - Name	A$_{CLC}$		A$_{UA}$		A$_{CLC}$-A$_{UA}$	
	km^2	%	km^2	%	km^2	%
1-Artificial surfaces	54.7	7.4	114.5	15.6	-59.8	-8.2
2-Agricultural areas	284.9	38.8	245.7	33.5	39.2	5.3
3-Forest and semi natural areas	386.7	52.7	367.2	50.0	19.5	2.7
4-Wetlands	1.4	0.2	0.8	0.1	0.6	0.1
5-Water bodies	6.7	0.9	6.2	0.8	0.5	0.1
Sum	734.4	100	734.4	100	---	---

Table 2. Areas (in km^2) of the classes in the study area obtained from CLC (A$_{CLC}$) and UA (A$_{UA}$), their percentage relative to the study area, and the difference between A$_{CLC}$ and A$_{UA}$

It can be seen that there is a large difference in the areas between the two LULC products, mainly for class 1 (Artificial surfaces), where the area of this class in UA is more than double that of the area in CLC, corresponding to 7.4% and 15.6% of the whole study area for CLC and UA, respectively. This difference is probably due to the differences in the MMUs of both products, as a large proportion of urban areas in this region are not mapped in the CLC because they correspond to small villages spread along the territory, mixed with agriculture as well as natural and semi-natural areas. The area of all other classes is overestimated in the CLC when compared to the UA; this overestimation is the largest for class 2 (Agricultural areas) at 39.1 km^2 (5.3% of the study area).

3. METHODOLOGY

The differences in the areas of the classes obtained in the CLC and UA for the study area presented in the previous section (Table 2) show a bias that may be relevant for some applications if this fact is not taken into consideration. Therefore, in order to provide a good understanding of the possible applications of the products, the accuracy assessment should not be done only in relation to the map specifications, but also using reference data that represents reality as closely as possible. The methodology used in this study to illustrate this is explained in the following two subsections, addressing: 1) the approaches used to create the reference data sets; 2) the methods used to perform the accuracy assessment, both with proportions and with the dominant class per pixel.

3.1 Reference data to validate LULC maps

In this paper, the reference database is created in two different ways: 1) recording only the class occupying the majority of the pixel (referred to as Ref$_{maj}$); 2) recording the proportion of each class present in each pixel (referred to as Ref$_{prop}$). Table 3 shows an example of both reference databases for a set of 4 pixels shown in Figure 4. Note that Ref$_{maj}$ can always be extracted from Ref$_{prop}$ by choosing the class with the largest proportion.

Pixel ID (k)	Map class i (CLC)	Ref$_{prop}$		Ref$_{maj}$	
		Reference class j (UA)	Proportion in pixel - $pp_{ij}(k)$	Reference class j (UA)	Proportion in pixel - $pp_{ij}(k)$
1	1	1	$pp_{11}(1) = 0.7$	1	$pp_{11}(1) = 1$
		2	$pp_{12}(1) = 0.2$	2	$pp_{12}(1) = 0$
		3	$pp_{13}(1) = 0.1$	3	$pp_{13}(1) = 0$
2	1	1	$pp_{11}(2) = 0.8$	1	$pp_{11}(2) = 1$
		2	$pp_{12}(2) = 0.2$	2	$pp_{12}(2) = 0$
		3	$pp_{13}(2) = 0$	3	$pp_{13}(2) = 0$
3	2	1	$pp_{21}(3) = 0.4$	1	$pp_{21}(3) = 0$
		2	$pp_{22}(3) = 0.6$	2	$pp_{22}(3) = 1$
		3	$pp_{23}(3) = 0$	3	$pp_{23}(3) = 0$
4	3	1	$pp_{31}(4) = 0$	1	$pp_{31}(4) = 0$
		2	$pp_{32}(4) = 0.3$	2	$pp_{32}(4) = 0$
		3	$pp_{33}(4) = 0.7$	3	$pp_{33}(4) = 1$

Table 3. Example of reference databases Ref$_{prop}$ and Ref$_{maj}$ corresponding to the four pixels shown in Figure 4
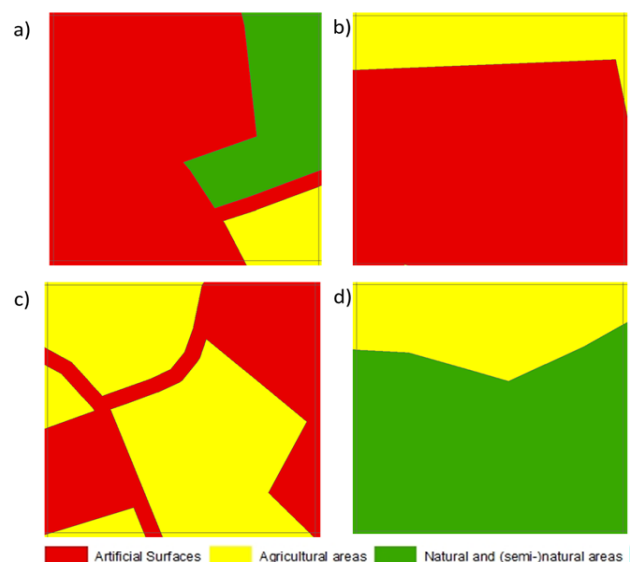

Figure 4: Example of the pixels partially occupied by the classes of interest. These pixels were used to extract the reference data shown in Table 3: a) Pixel ID=1; b) Pixel ID=2; c) Pixel ID=3; d) Pixel ID=4

3.2 Accuracy assessment

3.2.1 Creation of confusion matrices: Each cell of the confusion matrix, where the rows correspond to the classes in the map and the columns to the reference database, is computed using equation (1):

$$c_{ij} = \sum_{s=1}^{r} pp_{ij}(s) \tag{1}$$

where c_{ij} = the value of the cell in row i and column j in the confusion matrix
r = the number of spatial units in the reference database
$pp_{ij}(s)$ = the proportion of class j in the spatial unit s that is assigned in the map to class i.

This corresponds to both the Minimum operator and the Composite operator as proposed by Pontius & Cheuk (2006), which produce the same result when the map results from a hard classification (i.e., when only one class is assigned to each pixel). In this case it is the reference data that have partial membership in the classes, instead of the map pixels, as in maps resulting from soft classifications.

Table 4 shows how the example pixels in Table 3 can contribute to the confusion matrix when the Ref$_{prop}$ data set is used. Table 5 shows the traditional confusion matrix obtained with Ref$_{maj}$ for the same example pixels.

Classes	1	2	3	Sum (pixels)
1	0.7+0.8 =1.5	0.2+0.2=0.4	0.1+0=0.1	2
2	0.4	0.6	0	1
3	0	0.3	0.7	1
Sum (pixels)	1.9	1.3	0.8	4

Table 4. Confusion matrix obtained with the example pixels shown in Table 3 (Ref$_{prop}$)

Classes	1	2	3	Sum (pixels)
1	1+1 =2	0+0=0	0+0=0	2
2	0	1	0	1
3	0	0	1	1
Sum (pixels)	2	1	1	4

Table 5. Confusion matrix obtained with the example pixels shown in Table 3 (Ref$_{maj}$)

It can be seen that in both matrices, the sum of the rows will be the same, as we are using the same map for both; therefore, the class in the map is the same for all pixels in the reference database. However, the sum of the columns will be different in both matrices, because the area occupied by each class in the reference data is not the same. Hence, the sum of the columns when Ref$_{prop}$ is used will be decimal numbers instead of integers.

3.2.2 Accuracy indices: These are computed from the confusion matrices generated in section 3.2.1 using the same formulas as for traditional matrices. The users' accuracy of class i, the producer's accuracy of class j and the overall accuracy are obtained using equations (2), (3) and (4), respectively.

$$UA_i = \frac{c_{ii}}{\sum_{k=1}^{n} c_{ik}} \tag{2}$$

$$PA_j = \frac{c_{jj}}{\sum_{k=1}^{n} c_{kj}} \tag{3}$$

$$OA = \frac{\sum_{k=1}^{n} c_{kk}}{n} \tag{4}$$

where c_{ij} = the value of the cell in row i and column j in the confusion matrix
n = the number of classes in the map.

4. RESULTS

The following subsections provide the results obtained for the accuracy indices using the Ref$_{maj}$ and Ref$_{prop}$ reference databases, as well as the area estimates that can be obtained when using one or the other.

4.1 Accuracy indices

Table 6 and Table 7 show, respectively, the confusion matrices and the accuracy indices obtained with Ref$_{maj}$ and Ref$_{prop}$ using the methodologies explained in section 3.2. Table 8 shows the difference between the corresponding values obtained with both reference data sets.

Classes	1	2	3	4	5	Sum (pixels)	UA$_{maj}$ (%)
1	4522	573	319	0	58	5472	83
2	4575	18995	4834	6	86	28496	67
3	880	5414	32351	2	19	38666	84
4	0	48	12	70	5	135	52
5	12	167	126	0	366	671	55
Sum (pixels)	9989	25197	37642	78	534	73440	
Area sum (km^2)	99.9	252.0	376.4	0.8	5.3	734.4	
PA$_{maj}$ (%)	45	75	86	90	69	OA$_{maj}$:77%	

Table 6. Confusion matrix with the class majority per pixel (Ref$_{maj}$) and the accuracy indices: User's Accuracy (UA), Producer's Accuracy (PA) and Overall Accuracy (OA).

Classes	1	2	3	4	5	Sum (pixels)	UA$_{prop}$ (%)
1	4269	719	425	0	56	5469	78
2	5352	17987	4980	8	167	28494	63
3	1799	5657	31144	2	62	38664	81
4	1	49	15	65	4	134	48
5	30	153	158	0	328	669	49
Sum (pixels)	11453.2	24567.2	36723.5	76	620.1	73440	
Area sum (km^2)	114.5	245.7	367.2	0.76	6.2	734.4	
PA$_{prop}$ (%)	37	73	85	86	53	OA$_{prop}$: 73%	

Table 7. Confusion matrix with the proportion of classes per pixel (Ref$_{prop}$) and the accuracy indices: User's Accuracy (UA), Producer's Accuracy (PA) and Overall Accuracy (OA).

Class	User's Accuracy (%)			Producer's accuracy (%)		
	UA$_{maj}$	UA$_{prop}$	UA$_{maj}$-UA$_{prop}$	PA$_{maj}$	PA$_{prop}$	PA$_{maj}$-PA$_{prop}$
1	83	78	5	45	37	8
2	67	63	4	75	73	2
3	84	81	3	86	85	1
4	52	48	4	90	86	4
5	55	49	6	69	53	16
OA$_{maj}$ (%)				77		
OA$_{prop}$ (%)				73		
OA$_{maj}$- OA$_{prop}$ (%)				4		

Table 8. Comparison of the accuracy indices obtained with the reference database Ref$_{maj}$ and Ref$_{prop}$: User's Accuracy (UA$_{maj}$ and UA$_{prop}$), Producer's Accuracy (PA$_{maj}$ and PA$_{prop}$) and Overall Accuracy (OA$_{maj}$ and OA$_{prop}$).

The results in Table 8 show that all the accuracy indices obtained with the reference database containing the majority class (Ref_{maj}) have higher values than the ones obtained with the reference database where the class proportions were considered (Ref_{prop}). The larger difference was obtained for the Producer's Accuracy of class 5 (Water bodies), with a difference of 16%, and the smallest was also obtained for the Producer's Accuracy, but for class 3 (Forest and semi natural areas), with a value of 1%. The overall accuracy of the CLC evaluated with the Ref_{maj} was 4% higher than the one obtained with Ref_{prop}. This difference occurs because, when using Ref_{maj}, the pixel is considered to be 100% correctly classified (contributing to the accuracy assessment with a value of 1 in the cell of the confusion matrix corresponding to the agreement between classes) as long as at least 50% of it includes the class assigned to the pixel in the map, even if the other 50% of the pixel includes, in reality, other classes. For this same example, the accuracy assessment made with Ref_{prop} will only consider an agreement of 50% (0.5 in the cell of the confusion matrix corresponding to the agreement between classes), which in fact, corresponds to the real agreement between reality and the map, independently of the map specifications.

4.2 Accuracy of area estimates

The correct estimation of class areas in LULC maps is very important (Stehman & Foody, 2019). This estimation may be done using the LULC map, but as the map may have errors, this estimation may be incorrect. As the reference database has more reliable data, Olofsson et al. (2014) recommend that the area estimation should be based on the reference database used for accuracy assessment instead of the LULC map. When comparing the area classes estimated with the Ref_{maj} and Ref_{prop} reference databases (eighth line of Tables 6 and 7) with the true class areas shown in Table 2 (considering that UA represents ground truth), it can be seen that when considering the proportions, the true areas of the classes are correctly expressed in the reference data (Ref_{prop}), while this is not true for Ref_{maj}. In this case study, as the UA was used as the reference, and therefore no sampling approach was used, the area values obtained with the Ref_{prop} are exactly equal to the areas of the UA. When using a sampling approach, these will not be exactly the same, but they will provide a better estimation than Ref_{maj}.

5. CONCLUSIONS AND FUTURE WORK

This paper recommends that the labelling protocol used to create reference data sets for assessing the accuracy of crisp LULC maps, in particular when using photo interpretation, should be based on the assignment of not only one class to each spatial unit but the percentage of the spatial unit occupied by each class. The choice of only one class removes the existence of any other classes at that location, and therefore does not express a realistic ground truth, but rather a simplified version of it, which will introduce bias in the accuracy assessment.

However, we can identify arguments both against and in favour of considering the proportions of classes per spatial unit in the reference data set. In the next two subsections, some of these arguments are presented and discussed.

5.1 Arguments against using proportions in the reference data

1) One of the arguments that may be used against the estimation of class proportions is that it will add another source of uncertainty to the reference data, since that assessment may be subjective. Although this is true, it will be a second order uncertainty, as the uncertainty of having, for example, a class occupying 20% or 40% of a spatial unit is smaller than the uncertainty introduced by completely ignoring its potential presence, which will result in much larger errors whenever mixed pixels are present.

2) Another possible argument against this approach is that it will take more time to create a reference database with class proportions than selecting only the most representative class per spatial unit. However, mixed pixels are not present all of the time. In the situation where the LULC is relatively homogenous, class proportions are not relevant, their effect on the accuracy assessment will be small and the increase in the time taken to create the reference data set will also be very small. In contrast, in heterogeneous landscapes where many mixed pixels exist, the need to choose one class when several are present also requires interpretation time, as it is necessary to identify the class occupying the majority of the pixel. Hence in practice, identifying proportions in this situation may not require much additional time. Moreover, as the case study presented here illustrates, the implications of discarding all classes except the majority one may have a large influence, for example, on the area estimation. Finally, considering only one class per pixel may also be a factor that influences the variability of class assignment by different photo interpreters in the reference database, contributing to the presence of undesirable subjectivity in the data set. This problem could be addressed, for example, by using several photo interpreters. If it is proven that considering proportions of classes per pixels reduces subjectivity, this may result in an additional benefit and reduce the workload if fewer photo interpreters are needed. Thus, further investigation into the pros and cons of using class proportions in reference databases instead of identifying only the majority class is still needed.

3) It can also be argued that this approach may be of interest for mixed pixels, but it does not solve the problem of having difficulties in discriminating between similar classes, such as pasture and natural vegetation. In fact, this is a different problem that needs to be addressed with other procedures.

5.2 Arguments in favour of using proportions in the reference data

1) The main advantages in using class proportions in the reference data is that more realistic accuracy results are obtained and areas can be better estimated, as classes occupying many small areas are better represented in the reference database.

2) The creation of confusion matrices with class proportions is very easy and it is similar to the procedure used when pixels are considered as indivisible in terms of class occupation. Moreover, all procedures traditionally used to compute accuracy indices may also be used in this case. Both these aspects are advantageous when compared to other proposed solutions, such as the creation of additional accuracy indices such as the MAX or RIGHT proposed by Gopal and Woodcock (1994) or the fuzzy confusion matrices proposed by Sarmento et al. (2013). Although these approaches provide additional information that is potentially useful, the fact that they require additional efforts to understand and implement may contribute to the resistance of most communities in using them. The consideration of class proportions per pixel in the reference database would therefore be a simpler method, and it is very easy to implement and understand, which may be an important aspect for its effective use by the remote sensing community.

The increasing volume of remotely sensed images, collected by different sources and with higher spatial and temporal resolutions, translates into more available data to create LULC maps, and the possibility of having LULC maps updated with very high frequencies (which may reach a few days). This raises several challenges, which include the need to create LULC maps that translate the ground characteristics as faithfully as possible, so that the divergence between products with different origins decreases, and they may even be used for change detection. If the product's accuracy assessment relies less on ground truth and more on the map characteristics and specifications (such as pixel size and MMU), this will make the comparison of the accuracy of different products more difficult and less reliable, decreasing their reliability and real value. Therefore, it is important not only to have research that aims to develop methodologies to improve LULC map production, but also to develop methodologies to improve and hasten their rigorous accuracy assessment.

ACKNOWLEDGEMENTS

REFERENCES

Albrecht, F., Lang, S., & Hölbling, D. (2010). SPATIAL ACCURACY ASSESSMENT OF OBJECT BOUNDARIES FOR OBJECT-BASED IMAGE ANALYSIS. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVIII-4/C7.*

Foody, G. M. (2010). Assessing the accuracy of land cover change with imperfect ground reference data. *Remote Sensing of Environment*, *114*(10), 2271–2285. https://doi.org/10.1016/j.rse.2010.05.003

Foody, G. M., & Arora, M. K. (1996). Incorporating mixed pixels in the training, allocation and testing stages of supervised classifications. *Pattern Recognition Letters*, *17*(13), 1389–1398. https://doi.org/10.1016/S0167-8655(96)00095-5

Gopal, S., & Woodcock, C. E. (1994). Theory and Methods for Accuracy Assessment of Thematic Maps Using Fuzzy Sets. *Photogrammetric Engineering and Remote Sensing*, *60*(2), 181–188.

Lesiv, M., See, L., Laso Bayas, J. C., Sturn, T., Schepaschenko, D., Karner, M., Moorthy, I., McCallum, I., & Fritz, S. (2018). Characterizing the Spatial and Temporal Availability of Very High Resolution Satellite Imagery in Google Earth and Microsoft Bing Maps as a Source of Reference Data. *Land*, *7*(4), 118. https://doi.org/10.3390/land7040118

Li, M., Zang, S., Zhang, B., Li, S., & Wu, C. (2014). A Review of Remote Sensing Image Classification Techniques: The Role of Spatio-contextual Information. *European Journal of Remote Sensing*, *47*(1), 389–411. https://doi.org/10.5721/EuJRS20144723

Mather, P., & Tso, B. (2016). *Classification Methods for Remotely Sensed Data*. CRC Press.

McRoberts, R. E., Stehman, S. V., Liknes, G. C., Næsset, E., Sannier, C., & Walters, B. F. (2018). The effects of imperfect reference data on remote sensing-assisted estimators of land cover class proportions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *142*, 292–300. https://doi.org/10.1016/j.isprsjprs.2018.06.002

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., & Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, *148*, 42–57. https://doi.org/10.1016/j.rse.2014.02.015

Pontius Jr, R. J., & Cheuk, M. L. (2006). A generalized cross-tabulation matrix to compare soft-classified maps at multiple resolutions. *International Journal of Geographical Information Science*, *20*(1), 1–30. https://doi.org/10.1080/13658810500391024

Sarmento, P., Fonte, C. C., Caetano, M., & Stehman, S. V. (2013). Incorporating the uncertainty of linguistic-scale reference data to assess accuracy of land-cover maps using fuzzy intervals. *International Journal of Remote Sensing*, *34*(11), 4008–4024. https://doi.org/10.1080/01431161.2013.772312

Sarmento, P., Fonte, C. C., Dinis, J., Stehman, S. V., & Caetano, M. (2015). Assessing the impacts of human uncertainty in the accuracy assessment of land-cover maps using linguistic scales and fuzzy intervals. *International Journal of Remote Sensing*, *36*(10), 2524–2547. https://doi.org/10.1080/01431161.2015.1043034

Stehman, S. V. (2009). Sampling designs for accuracy assessment of land cover. *International Journal of Remote Sensing*, *30*(20), 5243–5272. https://doi.org/10.1080/01431160903131000

Stehman, S. V., & Foody, G. M. (2019). Key issues in rigorous accuracy assessment of land cover products. *Remote Sensing of Environment*, *231*, 111199. https://doi.org/10.1016/j.rse.2019.05.018

Tiede, D., Lang, S., Albrecht, F., & Hölbling, D. (2010). Object-based Class Modeling for Cadastreconstrained Delineation of Geo- objects. *Photogrammetric Engineering and Remote Sensing*, *76*(2), 193–202.

Woodcock, C. E., & Gopal, S. (2000). Fuzzy set theory and thematic maps: Accuracy assessment and area estimation. *International Journal of Geographical Information Science*, *14*(2), 153–172. https://doi.org/10.1080/136588100240895

Woodcock, C. E., Gopal, S., & Albert, W. (1996). Evaluation of the Potential for Providing Secondary Labels in Vegetation Maps. *Photogrammetric Engineering and Remote Sensing*, *62*(4), 393–399.

Ye, S., Pontius, R. G., & Rakshit, R. (2018). A review of accuracy assessment for object-based image analysis: From per-pixel to per-polygon approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, *141*, 137–147. https://doi.org/10.1016/j.isprsjprs.2018.04.002