Incorporating uncertainty in the accuracy assessment of land cover maps using fuzzy numbers and fuzzy arithmetic

Pedro Sarmento, Hugo Carrão and Mario Caetano*
Centro de Estatística e Gestão de Informação (CEGI)
Instituto Superior de Estatística e Gestão de Informação (ISEGI)
Lisbon, Portugal
Remote Sensing Unit (RSU)
Portuguese Geographic Institute (IGP)
Lisbon, Portugal
*mario.caetano@igeo.pt

Cidália C. Fonte** and Nuno Cortês
Institute for Systems and Computer Engineering at Coimbra
Department of Mathematics, University of Coimbra
Coimbra, Portugal
**cfonte@mat.uc.pt

Abstract—This paper proposes an effort to include uncertainty in reference databases used to assess the accuracy of land cover maps. Five linguistic levels of confidence in land cover labelling are assigned to each sample observation and converted into fuzzy numbers. This information is introduced in a fuzzy confusion matrix and fuzzy accuracy measures, similar to the global, user’s and producer’s accuracy, are then derived from the fuzzy confusion matrix using fuzzy arithmetic. These measures consist of fuzzy numbers that incorporate the uncertainty in identifying the reference land cover class of the sample data. Fuzzy accuracy measures can be defuzzified to generate real numbers, enabling the conversion into crisp measures, which allow the comparison with the accuracy results obtained with traditional confusion matrices. The proposed methodology is tested on a case study. The quality of a map for Continental Portugal, derived from the automatic classification of MERIS images, is evaluated using a reference database generated with the proposed methodology.

Keywords: land cover maps; accuracy assessment; reference database uncertainty; fuzzy numbers; fuzzy arithmetic

1. INTRODUCTION

Land cover maps are essential to understand several geographical phenomena, such as climate change, loss of biodiversity, land cover change and vegetation distribution. Even though the production of land cover maps is of great importance, the evaluation of their quality is also fundamental. If decisions are based on these maps, their better or worst quality will inevitably affect the quality of the decisions.

The methodology traditionally applied to assess the accuracy of land cover maps requires the comparison of the produced maps with a reference sample database (Foody, 2002), that represents the ‘true’ land cover. The sample database is composed by the observations collected in several geographical sites, which are inspected with field visits and/or using high resolution satellite/aerial images. The comparison between the two datasets is represented in a confusion matrix, where generally the reference data and the map data are inserted in the columns and rows, respectively. This approach assumes that only one land cover class exists at each geographical site. However, more than one land cover class may exist in several locations, and the technician has often difficulties in choosing the most adequate class. This uncertainty arises because: 1) land cover classes rarely presents abrupt transitions between them; 2) landscape fragmentation; or 3) the natural continuum between land cover types, which makes difficult the process of discriminating between the land cover classes defined in the nomenclature.

Several authors use fuzzy set theory to deal with this uncertainty (e.g., Woodcock and Gopal, 2000; Lein, 2003). Gopal and Woodcock (1994) developed one of the pioneer’s studies that introduced fuzzy sets in reference databases, to assess the accuracy of crisp land cover maps. In their work, the photo-interpretation uncertainty in elaborating a reference database was introduced using a rating system considering a linguistic scale. This scale is based on the premise that experts most often use linguistic constructs to describe map accuracy (Woodcock and Gopal, 2000). The linguistic scale used by the authors is composed by five levels: 1) Absolutely wrong; 2) Understandable but wrong; 3) Reasonable or acceptable answer; 4) Good answer; 5) Absolutely right. At each sample observation, the technician uses this linguistic scale to express their perception when identifying reference land cover classes. Using this linguistic scale, the authors developed several methods based on fuzzy functions that provide more information about the accuracy of a map than a confusion matrix. These functions are presented in the form of four tables, that provide information about the frequency of errors (MAX and RIGHT operators); magnitude of errors (DIFFERENCE operator); source of errors (MEMBERSHIP operator) and nature of errors (CONFUSION and AMBIGUITY operators). Gopal and Woodcock (1994) refer the need to develop a method to express the results in a single

- 25 -
hard accuracy measure, standardize the accuracy results from
different experts and the construction of the traditional
confusion matrix with the fuzzy data. This is precisely the
research we carried out and that is reported in this paper.

This paper deals with human uncertainty in the
elementation of reference databases. In accordance with Gopal
and Woodcock (1994), we consider also five linguistic levels
to assign each sample observation (in the form of artificial
pixels), to each land cover class. These linguistic levels are
then converted into fuzzy numbers according to the photo-
interpreters interpretation of each linguistic level. This
information is then introduced in a fuzzy confusion matrix,
which is adapted to incorporate the collected fuzzy numbers.
Accuracy measures in the form of fuzzy numbers, equivalent
to global, user's and producer's accuracy are then derived
from the confusion matrix using fuzzy arithmetic. Fuzzy
numbers are then defuzzified to generate real numbers,
allowing the conversion of the fuzzy accuracy measures into
 crisp values, that allow a comparison with the accuracy
results obtained with traditional confusion matrices.

The proposed methodology is tested on a case study. The
accuracy of a map for Continental Portugal, derived from
MERIS (MEdium Resolution Imaging Spectrometer) images,
is evaluated with the proposed methodology. A fuzzy
confusion matrix is generated and the fuzzy global, user's and
producer's accuracy measures are obtained. In addition, from
the reference database with uncertainty, a traditional
reference database is generated and the traditional accuracy
measures are computed and compared with the fuzzy ones.

II. METHODOLOGY

A. Linguistic scale

The proposed approach for the accuracy assessment of
land cover maps starts with the definition of the linguistic
scale used to assign the membership degree of each land
cover class to each sample observation. Like the linguistic
scale proposed by Gopal and Woodcock (1994), also five
linguistic values were considered, namely: Absolutely wrong
(W); Understandable but wrong (U); Reasonable or
acceptable (A); Good (G); Absolutely right (R). Each one of
these linguistic values (fuzzy set) is then converted into a
membership function in the form of a trapezoidal fuzzy
number (Klir and Yuan, 1995). The core of a membership
function for some fuzzy set is characterized by complete and
full membership in a set, and his support is defined has the
region of the universe that is characterized by nonzero
membership in a set (Klir and Yuan, 1995). Each linguistic
level considers the percentage of each land cover class
existing in each sample observation, expressing this way the
heterogeneity of the landscape, like in the approach proposed
by Gopal and Woodcock (1994). The only exceptions are the
extremes of the linguistic scale (Absolutely right and
Absolutely wrong). In these cases, it was considered that
there was no uncertainty and a real number was used. An
observation completely occupied by a land cover class was
considered as "Absolutely right" for that class, and the total
absence of a land cover class is considered as "Absolutely
wrong" for that class. The fuzzy numbers corresponding to
each values of the linguistic scale are shown in Table I and
Fig. 1.

<table>
<thead>
<tr>
<th>Linguistic value</th>
<th>Support of the fuzzy number</th>
<th>Core of the fuzzy number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolutely wrong (W)</td>
<td>[0, 0]</td>
<td>[0, 0]</td>
</tr>
<tr>
<td>Understandable but wrong (U)</td>
<td>[0, 40]</td>
<td>[10, 30]</td>
</tr>
<tr>
<td>Reasonable or acceptable (A)</td>
<td>[30, 70]</td>
<td>[40, 60]</td>
</tr>
<tr>
<td>Good (G)</td>
<td>[60, 100]</td>
<td>[70, 90]</td>
</tr>
<tr>
<td>Absolutely right (R)</td>
<td>[100, 100]</td>
<td>[100, 100]</td>
</tr>
</tbody>
</table>

Figure 1. Uncertainty functions corresponding to the linguistic values
presented in Table I.

Table II shows an example of the information collected in
seven sample observations.

<table>
<thead>
<tr>
<th>n</th>
<th>Map class</th>
<th>Reference class</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>U</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>G</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>G</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>G</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>W</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>W</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>W</td>
</tr>
</tbody>
</table>

B. Fuzzy confusion matrix

The most common method used to compute and analyze
the accuracy assessment of land cover maps is through a
confusion matrix (Story and Congalton, 1986). A confusion
matrix allows the comparison between the produced maps
and a reference database. Sample observations that fall in
the major diagonal, represent matches between the two datasets.
Sample observations outside the major diagonal are
classification errors. The overall accuracy (OA) is obtained
summing the sample units in the main diagonal and dividing
them by the sample size. The producer's accuracy (PA) is
obtained dividing the number of correctly classified samples
of a certain land cover class, by the column total of samples
of the same category in the reference data. This evaluates
the omission errors. The user's accuracy (UA), which measures
the commission errors, is obtained using a similar approach
for the matrix rows.

The confusion matrix proposed in this paper allows the
computation of accuracy measures similar to the traditional
ones, but instead of using crisp numbers, uses fuzzy numbers.
Table III presents the fuzzy confusion matrix obtained with the data of Table II. W, U, A, G and R are the fuzzy numbers shown in Fig. 1, and the integer coefficients represent the number of times the linguistic values are found for each combination between the classification results and the reference data for the considered sample observations. The resulting value for each cell of the matrix is computed with fuzzy arithmetic, and a fuzzy number is obtained for each. The UA and PA are computed as in a traditional confusion matrix, but using now the fuzzy numbers existing in each cell, that is, the fuzzy number in the diagonal is divided by the sum of the rows or column values, respectively for the UA and the PA. The OA is obtained dividing the sum of the values in the main diagonal by the sum of the values in all cells.

<table>
<thead>
<tr>
<th>Reference Data</th>
<th>UA</th>
</tr>
</thead>
<tbody>
<tr>
<td>3G</td>
<td>2U+2A</td>
</tr>
<tr>
<td>2G</td>
<td>W</td>
</tr>
<tr>
<td>1U</td>
<td>R</td>
</tr>
<tr>
<td>5</td>
<td>W</td>
</tr>
</tbody>
</table>

In order to transform the fuzzy accuracy measures, into real numbers to allow a comparison with the traditional accuracy measures, three defuzzification methods were used (Ross, 1995): 1) Centroid method; 2) First of maxima (fom); 3) Last of maxima (lom).

III. Case Study

To test the proposed methodology of accuracy assessment, a land cover map of 2005 for Continental Portugal was used (Carrão, Araújo, Gonçalves and Caetano, in press). The land cover map is composed by five land cover classes (Urban Areas; Agriculture; Natural Vegetation; Forest and Water and Wetlands) derived from the generalization of the LANDEO nomenclature (Carrão, Araújo, Cerdeira, Sarmiento, Capão and Caetano, 2007) originally with 16 land cover classes. To obtain the reference data to assess the accuracy of the land cover map, a reference sample with 250 observations was collected trough a simple random sampling design. For each sample observation a linguistic value was assigned for each class, according to Table I. This information was collected through the photo-interpretation of a set of aerial images from 2004/2005/2006. These images have a spatial resolution of 0.5 m and four spectral bands (three bands within the visible and one in the near-infrared wavelengths).

A. Overall accuracy

The fuzzy number obtained for the overall accuracy is shown in Fig. 2. Its support is the interval [0.43, 0.78], which corresponds to all possible values the OA can take. The core of the fuzzy number are the values with full membership of occurrence, and in this case vary between 0.50 and 0.66, corresponding to the defuzzified values fom and lom respectively (Table IV). The OA was computed considering the MAX (that considers the land cover class with the highest linguistic value) and the RIGHT (that considers the two land cover classes with the highest linguistic values) operators, corresponding the first one to the traditional approach and the second to a more optimistic approach. The obtained values are shown in Table IV, along with the results obtained with the defuzzification methods considered. Comparing the results, it can be seen that MAX is similar to lom (0.65 and 0.66, respectively). An important outcome that should be referred, is that MAX is a point estimation and for a certain confidence level, this point is the center of a symmetric interval with equal probability of occurrence to left and right. In the proposed approach, the maximum possibility of the accuracy occurs only on the left side of lom (0.66). At right of this value, the possibility of the accuracy starts decreasing due to uncertainty in photo-interpretation. RIGHT has the highest value of accuracy (0.80), very similar to the maximum of the support of the fuzzy overall accuracy (0.78). In this sense, RIGHT gives an optimistic perspective of accuracy. A similar conclusion, considering a different approach to incorporate uncertainty in the accuracy assessment of land cover maps, was obtained by Sarmiento, Carrão, Caetano and Stelman (2009).

Table IV: Defuzzified overall accuracy of the considered land cover map, where FOM stands for first of maxima and LOM for last of maxima.

<table>
<thead>
<tr>
<th>Defuzzification methods</th>
<th>Traditional methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>centroid</td>
<td>fom</td>
</tr>
<tr>
<td>0.58</td>
<td>0.50</td>
</tr>
</tbody>
</table>

B. User's accuracy

The UA and PA were computed for all classes, but only the results obtained for UA are presented. Fig. 3 and Fig. 4 show respectively, the results obtained for the user's accuracy of Urban Areas and Agriculture classes. Table V shows the defuzzified values obtained for all land cover classes. The range (interval between the values of the support of a fuzzy number) of the fuzzy user's accuracy for Urban Areas (Fig. 3) is larger than the one for Agriculture (Fig. 4) (0.49 and 0.26 respectively), and the fuzzy accuracy is higher for Agriculture (Table V). For the Agriculture class, the high accuracy and the small uncertainty mean that a high proportion of samples that are Agriculture in the map have in the reference database high values of the linguistic scale for this class and low values for the others. The higher uncertainty for Urban Areas reflects that several of the samples are associated to this class.
in the reference database with considerable uncertainty. Like for the OA, RIGHT continues to present the highest value of accuracy for all land cover classes, exceeding slightly the maximum of the supports of fuzzy accuracies. Another issue that should be discussed is the fact that with the proposed approach, MAX is very similar to the highest value of accuracy with full membership (lom), for all land cover classes. In this sense, we suppose that MAX is the highest value of accuracy for whom we could be certain of the accuracy, and also MAX could be considered an optimistic thematic accuracy measure.

**TABLE V. USER’S ACCURACY FOR THE LAND COVER CLASSES OF THE NOMENCLATURE, WHERE FOM STANDS FOR FIRST OF MAXIMA AND LOM FOR LAST OF MAXIMA.**

<table>
<thead>
<tr>
<th>Land cover class</th>
<th>Defuzzification methods</th>
<th>Traditional methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>centroid</td>
<td>lom</td>
</tr>
<tr>
<td>Urban Areas</td>
<td>0.49</td>
<td>0.35</td>
</tr>
<tr>
<td>Agriculture</td>
<td>0.82</td>
<td>0.72</td>
</tr>
<tr>
<td>Natural Vegetation</td>
<td>0.42</td>
<td>0.33</td>
</tr>
<tr>
<td>Forest</td>
<td>0.51</td>
<td>0.44</td>
</tr>
<tr>
<td>Water and Wetlands</td>
<td>0.76</td>
<td>0.71</td>
</tr>
</tbody>
</table>

**IV. CONCLUSIONS**

The proposed methodology allows the incorporation of uncertainty in the elaboration of reference databases for the accuracy assessment of land cover maps. The representation of the uncertainty in the reference database with fuzzy numbers enables the computation of thematic accuracy indices with fuzzy arithmetic, giving a better understanding of the classification accuracy and the estimation of the uncertainty present in the accuracy indices. Defuzzification methods may be used to generate crisp accuracy measures from the fuzzy ones, enabling a direct comparison with the traditional measures and giving different perspectives of accuracy that may complement the traditional ones.

In the example presented in this study, trapezoidal fuzzy numbers were used to define the photo-interpreter uncertainty functions. This approach allows the aggregation of photo-interpretation uncertainty functions of each photo-interpreter and use these final functions to compute the fuzzy confusion matrix, incorporating the uncertainty of a team of photo-interpreters in the elaboration of reference databases. However, a more exhaustive analysis of the behavior of the photo-interpreter in the discrimination of land cover thresholds in each sample observation is necessary, allowing the elaboration of more accurate uncertainty functions. This issue is of extreme importance, since uncertainty functions are the core to compute the fuzzy confusion matrix and derive the fuzzy thematic accuracy measures.

**REFERENCES**


