Remote Sensing of Environment 142 (2014) 9–19

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Accounting for the area of polygon sampling units for the prediction of primary accuracy assessment indices

Julien Radoux *, Patrick Bogaert

Earth and Life Institute, Université catholique de Louvain, Croix du Sud, L7.05.16, B-1348, Belgium

ARTICLE INFO

Article history:
Received 3 February 2012
Received in revised form 25 November 2012
Accepted 31 October 2013
Available online 4 December 2013

Keywords:
GEOBIA
Object-based
Image-segment
User accuracy
Producer accuracy
Overall accuracy
Probabilistic sample
Land cover map

ABSTRACT

GEographic Object-Based Image Analysis (GEOBIA) has become a popular alternative for land cover and land use classification. In this case, polygons can be selected as sampling units to match the conceptual model of the map. However, little attention has been paid to the use of polygons for the validation of those maps. In this paper, we quantitatively assess the prediction of the primary thematic accuracy indices when the sampling unit is a polygon. The variable size of the sample polygons is a major concern for the prediction of the accuracy indices. Indeed, the classification accuracy, in addition to being class-dependent, depends on the polygon area. A practical solution supported by a theoretical framework that is conditional to the sample dataset is proposed in this study. This new predictor takes advantage of the known classification results for an improved efficiency. Empirical results based on synthetic maps show that the new predictor outperforms alternative methods for overall accuracy. The RMSE of the area weighted predictor was achieved with 50% less sample polygons thanks to our new predictor.

1. Introduction

Land cover/land use maps are of paramount importance in various applications such as land monitoring, land use planning, hydrological modelling or natural resource management. Consequently, map users need reliable quality information about those products for using them in an appropriate way. Previous works on accuracy assessment have designed standard quality indices and methods which are now widely accepted by the remote sensing community. The core of the accuracy assessment typically relies on a confusion matrix based on a validation sample, which matches the mapped land cover to some reference information (Congalton, 1991). The confusion matrix is often accompanied by global indices such as the overall, the user's and the producer's accuracy indices (Congalton, 1991; Foody, 2002; Stehman, 1997), which provide a useful summary of the map's quality. According to Liu, Frazier, and Kumar (2007), these are the three primary thematic accuracy indices.

Standard accuracy assessment methods rely also on the definition of a sampling unit used in the response design. Congalton and Green (2009) identified 3 types of sampling units: points, pixel clusters and polygons. A universally best spatial unit does not exist, so it is critical to recognize how the choice of a sampling unit affects the accuracy assessment process (Stehman & Wickham, 2011). This choice also depends on the conceptual model of the map (Fig. 1), i.e. spatial object, field or spatial regions according to definitions of Bian (2007):

- A spatial object is used as a conceptual model for spatially discrete information. The spatial extent of these objects is limited in space and their boundaries are defined by a set of rules. At least one categorical variable, the object type, is associated with those objects after a chosen typology. In a response design, spatial objects are most of the time unambiguously validated, either by photo-interpretation or from the field, because they can be embraced by the validation crew. Furthermore, their integrity is often assessed as a whole, using reference polygons and resulting in class-specific metrics where the geometric component plays a major role (Persello & Bruzzone, 2010). In this case Zhan, Molenaar, Tempfli, and Shi (2005) therefore concluded that polygon-based sampling units provide additional information compared with point-based approaches.
- A field consists in a spatially continuous quantitative variable that can be measured in any point of space. A typical example is the elevation above a reference surface, which is an important variable in, e.g., hydrological modelling. Although it is possible, and in some cases recommended (Lambin, 1999), to describe the land cover using continuous fields, classification is more popular (Huang, Davis, & Townshend, 2002), especially for large scale mapping. In any case, the validation of spatial fields primarily relies on point-based sampling (Hansen, DeFries, Townshend, Marufu, & Sohlberg, 2002) because polygons would introduce abrupt changes in the field values.

* Corresponding author. Fax: +32 10 47 88 98.
E-mail address: julien.radoux@uclouvain.be (J. Radoux).

0034-4257/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.rse.2013.10.030

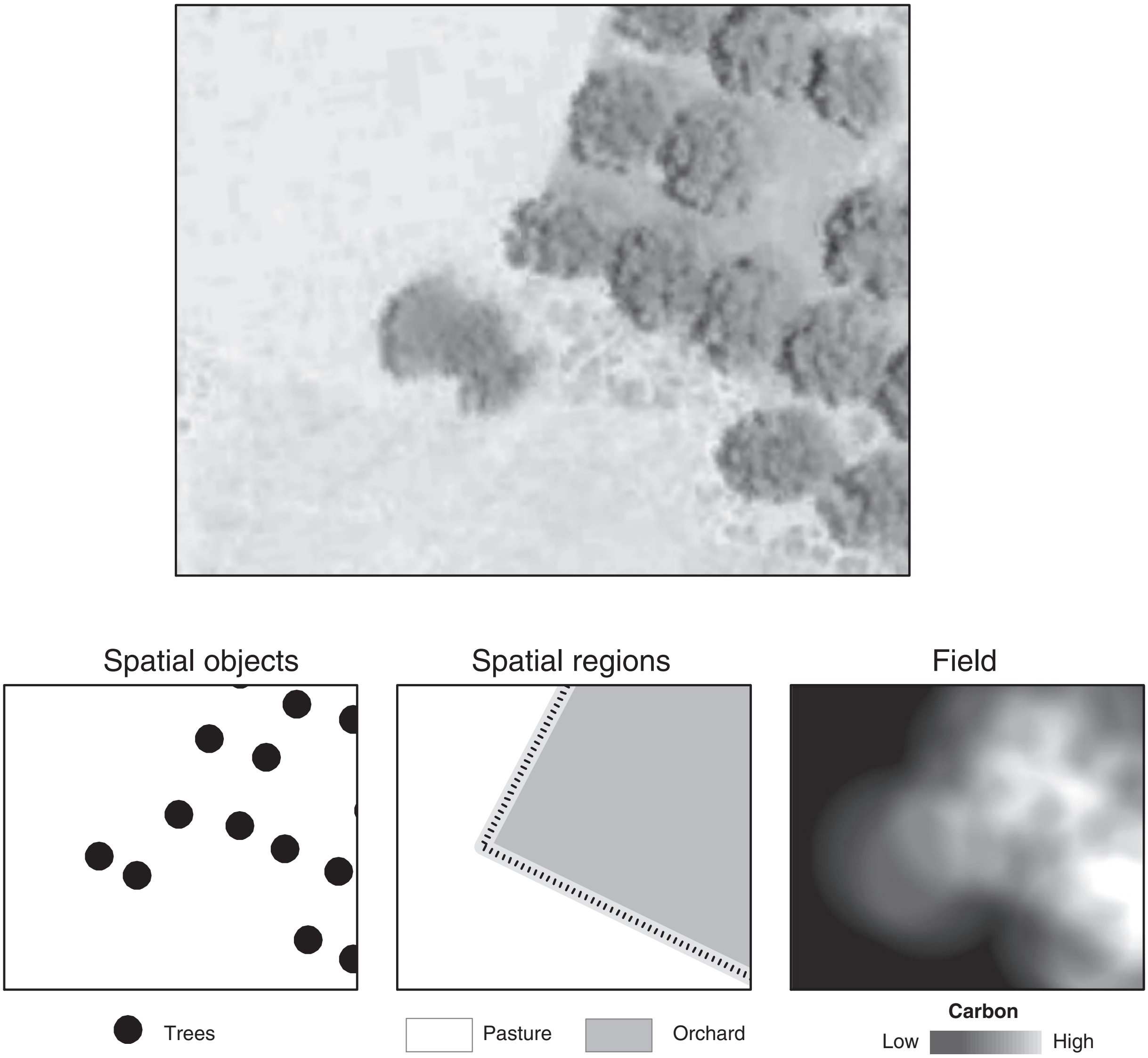

Fig. 1. Representations of the same site using three different conceptual models.

- A spatial region represents a mass of individuals that can be conceptualised both as a continuous field and as discrete spatial objects, which is often the case of land cover. This duality is also found at the level of the logical model: they can be discretised as vector polygons with consensual (fuzzy) boundaries or represented as a grid with the proportion of each individual and no defined boundaries. Spatial regions are delimited with an arbitrary boundary that is difficult to define with a set of rules (e.g. ecotones), so that their position is often uncertain and the sources of geometric errors are diverse (Radoux & Defourny, 2007). Concerning the labels, the use of unambiguous classification systems, such as the UN Land Cover Classification System (Di Gregorio & Jansen, 2000), is recommended in order to avoid overlapping class definitions.

When spatial objects or spatial regions are identified on a map as polygons, Congalton and Green (2009) recommend the use of sample polygons to assess the thematic accuracy. GEographic Object Based Image Analysis (GEOBIA) is a typical case where the resulting map is partitioned in a set of polygons. GEOBIA is increasingly used to process remote sensing data (Blaschke, 2010) and has been successfully applied in image classification and change detection (Bontemps, Bogaert, Titeux, & Defourny, 2008; Radoux & Defourny, 2010). A rationale of this approach is that the interpretation of a group of spatially adjacent pixels with similar properties is closer to human interpretation of spatial regions than independent pixel interpretation. Intrinsically, the polygons used in GEOBIA are thus considered as homogeneous in terms of land cover (Hay & Castilla, 2008). Those polygons are built based on an image (so called image-segments or image-objects) or obtained from an ancillary data source.

Various methods were developed to evaluate image segmentation goodness based on supervised and unsupervised indices (Clinton, Holt, Scarborough, Yan, & Gong, 2010; Neubert, Herold, & Meine, 2008; Zhang, Fritts, & Goldman, 2008). These indices are most of the time related to the four criteria proposed by Haralick and Shapiro (1985): i) regions should be homogeneous with respect to some characteristics, ii) adjacent regions should exhibit marked differences with respect to these characteristics, iii) region interiors should be free of holes, and iv) boundaries should be spatially accurate and precise. In the frame of GEOBIA, the two first criteria are directly related to over- and under-segmentation concerns for an image, respectively when an image-segment is only a part of a spatial region or a spatial object, and when more than one spatial object or region are included in the same image-segment (Carleer, Debeir, & Wolff, 2005). After classification, over-segmentation and holes are potentially removed while under-segmentation may lead to artificial class associations that often reduce the semantic map quality.