International Journal of Applied Earth Observation and Geoinformation 28 (2014) 117–130

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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

Area-based and location-based validation of classified image objects

Timothy G. Whiteside[a,*], Stefan W. Maier[b], Guy S. Boggs[c]

[a] Environmental Research Institute of the Supervising Scientist, Darwin, NT 0820, Australia
[b] Research Institute of the Environment and Livelihoods, Charles Darwin University, Darwin, NT 0909, Australia
[c] Wheatbelt NRM Inc, P.O. Box 311, Northam, WA 6401, Australia

ARTICLE INFO

Article history:
Received 19 June 2013
Accepted 21 November 2013

Keywords:
Geographic object-based image analysis
Validation
Accuracy assessment

ABSTRACT

Geographic object-based image analysis (GEOBIA) produces results that have both thematic and geometric properties. Classified objects not only belong to particular classes but also have spatial properties such as location and shape. Therefore, any accuracy assessment where quantification of area is required must (but often does not) take into account both thematic and geometric properties of the classified objects. By using location-based and area-based measures to compare classified objects to corresponding reference objects, accuracy information for both thematic and geometric assessment is available. Our methods provide location-based and area-based measures with application to both a single-class feature detection and a multi-class object-based land cover analysis. In each case the classification was compared to a GIS layer of associated reference data using randomly selected sample areas. Error is able to be pin-pointed spatially on per-object, per class and per-sample area bases although there is no indication whether the errors exist in the classification product or the reference data. This work showcases the utility of the methods for assessing the accuracy of GEOBIA derived classifications provided the reference data is accurate and of comparable scale.

1. Introduction

Site-specific accuracy assessment methods typically associated with per-pixel classifications (Congalton, 1991; Congalton and Green, 2009) have obvious limitations when applied within the geographic object-based image analysis (GEOBIA) paradigm (Clinton et al., 2010; Schöpfer and Lang, 2006). While these methods do provide information on the quality or accuracy of a classification at particular locations (x,y) across the image (Zhan et al., 2005), when applied to the output of a GEOBIA, there is uncertainty about the extent of the reference class beyond that location. The assumption that the thematic value of that reference point is consistent over the entire area of the object is therefore debatable, even if the reference is large enough to be representative of the preferred block of pixels (Stehman and Wickham, 2011). In short, single pixel- and block-based approaches for accuracy assessment do not answer the following question: How well does the classified object typify, both thematically and geometrically, the real world object it is meant to represent?

Methods of assessing image segmentation accuracy are well documented (Clinton et al., 2010; Delves et al., 1992; Hoover et al., 1996; Lucieer, 2004; Möller et al., 2007; Prieto and Allen, 2003), and generally compare the output of a segmentation algorithm to manually delineated features and their outlines in the imagery. Although segmentation accuracy can influence thematic accuracy, the methods do not assess thematic accuracy. Therefore, a method of assessing both the thematic and geometric accuracy of classified objects is needed (Schöpfer et al., 2008).

The accuracy assessment of GEOBIA outputs has been identified as an area of emerging research (Blaschke, 2010). One advantage of GEOBIA is an output (classified objects) that is claimed to be ready for GIS implementation (Benz et al., 2004). While it is important to know how well an initial segmentation provides objects suitable for classification (Clinton et al., 2010; Möller et al., 2007), the end result of objects also needs to be assessed particularly if they are used as input into a GIS model or used in decision making processes. For such an output to be valuable in GIS analysis, the output would require an assessment of the geometric accuracy (location and shape) of its classified objects (Schöpfer et al., 2008). Traditional site-specific accuracy assessment methods based upon site-specific reference data, such as confusion matrices (Congalton and Green, 2009; Story and Congalton, 1986), do not provide this type of information, and there has been very little work undertaken on determining suitable spatial accuracy measures for object-based image analysis (Schöpfer et al., 2008; Weidner, 2008; Winter, 2000; Zhan et al., 2005). Much of the work has focussed on the assessment of building detection where spatial accuracy is a requirement (Weidner, 2008; Winter, 2000). Very little research has been undertaken into the application of spatial accuracy measures for object-based multi-class analysis (Lang et al., 2009;

* Corresponding author. Tel.: +61 8 89201161; fax: +61 8 89201195.
E-mail address: tim.whiteside@environment.gov.au (T.G. Whiteside).

0303-2434/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jag.2013.11.009

Lang and Tiede, 2008; Schöpfer et al., 2008), particularly in spatially and spectrally variable land cover such as tropical savanna. Land cover in such landscapes has been difficult to map due to the gradual transitions and inherent heterogeneity of the landscape components (Hayder, 2001; Whiteside et al., 2011a).

The objectives of this paper are to implement a number of area-based accuracy measures and assess the measures' efficacy in providing accuracy information about classified objects derived from imagery over a spatially and spectrally variable landscape. The measures will be applied to two different sets of GEOBIA derived classified objects: (1) a single class (or feature detection) tree crown delineation and (2) a multi-class land cover layer. The remainder of the Introduction (Sections 1.1–1.3) provides background information on the spatial accuracy of objects and accuracy measures relevant to GEOBIA.

1.1. The problem with confusion matrices

While assessing whether an object has been assigned to the correct class can be determined using a simple confusion matrix (as described by Congalton and Green, 2009), there are issues with using confusion matrices. While per-class and overall classification accuracies are highlighted and confusion between classes can be identified (Foody, 2002), confusion matrices do not show (spatially) where agreement or confusion may occur. In addition, accuracy values derived from confusion matrices for single category (target vs. non-target) classifications such as feature detection can be dubious. Traditional confusion matrix metrics such as user's and producer's accuracies for the non-target class do not contain accuracy information relevant to the intent of classification (Zhan et al., 2005). Due to the non-target class invariably consisting of a number of land cover types, the information from a confusion matrix including a non-target class does not enable reliable calculation of the Kappa statistic (Zhan et al., 2005). This, however, may be a moot point as there are strong arguments that the calculation of a Kappa statistic provides no new information (to the overall accuracy measure) for classification accuracy and is therefore unnecessary (Foody, 2011; Pontius and Millones, 2011). Accuracy measures that include spatial information as well as thematic should overcome some of these problems.

1.2. Object accuracy

In determining the classification accuracy of post-classification objects it is important to consider both (a) the classification (also known as categorical or thematic) accuracy of the objects and (b) the spatial accuracy (the shape and location) of the objects. Spatial accuracy measures do require a layer of reference objects prior to implementation. In some cases, that layer maybe either of inappropriate scale, of dubious accuracy, or may not be available at all.

1.3. Spatial accuracy

Spatial accuracy refers to how well a classified object (C) spatially matches (location and shape wise) the real world object (R, represented by reference data) it represents. Location accuracy refers to the position in space of a classified object in relation to a corresponding reference object. Shape accuracy refers to the degree of similarity of the two objects based on a number of shape-based criteria (including area, perimeter, length, and width). Similarity as described here is based on Tversky's (1977) feature contrast model (Eq. (1)):

$$s(a,b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A), \text{ for some } \theta, \alpha, \beta \geq 0; \quad (1)$$

where $s(a,b)$ is the similarity between sets a and b and is a function (f) of three arguments: $f(A \cap B)$ are features common to both a and b, $f(A - B)$ features of a but not b, $f(B - A)$ features of b not a, and α, β and θ are the respective weightings for the three relationships. This model assumes that the similarity between two items or sets is a weighted function of both feature matching (common to both items) and mismatching (belonging to one item but not the other) (Tversky, 1977).

In the case of two objects (C and R), the more criteria that match between C and R, the greater the similarity is between the two objects. These measures require reference objects for comparison against classified objects. A major limitation with this type of reference data is the need for objects to be of a similar spatial scale to the classification. If the reference data are of a coarser scale than the classification they will lack the spatial variability of the classification. Alternatively, if the reference data are of a finer scale than the classification there will be greater spatial variability than the classification. Both cases may affect the perceived accuracy of the classification. There are limitations associated with temporal differences that also need consideration.

Ideally, to implement spatial accuracy measures there should be one-to-one correspondence between C and R objects (Clinton et al., 2010). A C object and corresponding R are established if there exists overlap between the two objects. In a comparison of a land cover map to a reference layer of objects there will always be spatial correspondence between objects from the two layers, although there may be thematic differences. In a single class (or feature detection), where a C object exists with no corresponding R, it is a false positive (Whiteside et al., 2011b) and that instance contributes to a class's commission error. Where an R object exists with no corresponding C object, then it is a non-positive and the instance contributes to a class's omission error. There may also be instances where more than one R object corresponds to a C object, and vice versa. As the measures used here are area-based, the sum of the overlap is used.

2. Methods

2.1. Location accuracy

Location-based accuracy measures assess the similarity in location between a classified or extracted object and its corresponding reference object. Measures that define the distance between a classified object and the corresponding reference object can be considered measures of object accuracy. Within certain parameters, the distance from the centre of the classified object to the centre of the reference object is inversely proportional to the location accuracy. Conversely, the smaller the distance between the central points, the greater the location accuracy of the classified object relative to the reference object.

The *Loc* measure (Eq. (2)) utilised by Zhan et al. (2005) is based on the Euclidean distance between centroids to provide location accuracies for extracted objects within a scene to relation to their reference counterparts:

$$Loc_{C_i,R_i} = \sqrt{(x_{c_i} - x_{R_i})^2 + (y_{c_i} - y_{R_i})^2} \quad (2)$$

where C_i and R_i are the ith extracted and corresponding reference objects respectively, x_{C_i} and y_{C_i} are the x and y coordinates of the centroid of C, and x_{R_i} and y_{R_i} are the x and y coordinates of centroid of R. *MeanLoc* is the mean distance (representing overall quality) while *StDevLoc* is the standard deviation of the measure.

Another horizontal accuracy measure that could be used is the root mean square error (RMSE) (Congalton and Green, 2009),