Uncertainty in ecosystem mapping by remote sensing

Duccio Rocchini, Giles M. Foody, Harini Nagendra, Carlo Ricotta, Madhur Anand, Kate S. He, Valerio Amici, Birgit Kleinschmit, Michael Förster, Sebastian Schmiedlein, Hannes Feihlauer, Anne Ghisla, Markus Metz, Markus Neteler

Abstract

The classification of remotely sensed images such as aerial photographs or satellite sensor images for deriving ecosystem-related maps (e.g., land cover, land use, vegetation, soil) is generally based on clustering of spatial entities within a spectral space. In most cases, Boolean logic is applied in order to map landscape patterns. One major concern is that this implies an ability to divide the gradual variability of the Earth’s surface into a finite number of discrete non-overlapping classes, which are considered to be exhaustively defined and mutually exclusive. This type of approach is often inappropriate given the continuous nature of many ecosystem properties. Moreover, the standard data processing and image classification methods used will involve the loss of information as the continuous quantitative spectral information is degraded into a set of discrete classes. This leads to uncertainty in the products resulting from the use of remote sensing tools.

It follows that any estimated ecosystem property has an associated error and/or uncertainty of unknown magnitude, and that the statistical quantification of uncertainty should be a core part of scientific research using remote sensing. In this paper we will review recent attempts to take explicitly into account uncertainty when mapping ecosystems.

1. Introduction

Mapping and modeling the complexity of ecosystems and their changes over time is a key issue in spatial ecology and biogeography. Remote sensing has been acknowledged as one of the most powerful methods to map abiotic and biotic components of ecosystems (including land cover, land use, vegetation, soils) and estimate their changes over time. The mechanism used to create maps based on remote sensing data is to derive classification algorithms to label pixels. Formally speaking, let \( S = \{1, 2, 3, \ldots, n\} \) be a set of pixels; clustering algorithms seek to find out the most probable/possible partition of \( S \).

Quintana (2006) provides a detailed mathematical review of clustering algorithms. In most cases classes are derived relying on Boolean rules where classes are sharply defined and pixels are generally associated to a class on the basis of relative spectral similarity.

Regardless the method being used (raster-based or object-oriented classification), the assumptions for carrying out classification are associated with one major drawback: classes are mutually exclusive with discrete boundaries separating each other. Hence, processing and classifying images can result in a substantial loss of information, due to the degradation of continuous quantitative information into discrete classes (Foody, 2000; Palmer et al., 2002).

Many authors have attempted to produce better representations of the true complexity using improved methods for discrete boundaries, e.g., using multi-scale segmentation based...
on hierarchical patch dynamics (e.g., Blaschke, 2010). Nonetheless, it has been widely demonstrated in ecological studies that habitats vary in space in a continuum manner (Rocchini, 2010, for a review) and, in some cases, a standard discrete boundary provides an unrealistic representation (Foody, 1999). As an example, Fig. 1 represents an aerial photo (1 m spatial resolution) of a mountainous landscape (Monte di Mezzocorona and Valle dell'Adige, Trentino, Northern Italy, centre scene coordinates, \( \lambda 11^\circ 07'02''\text{E}, \phi 46^\circ 13'36''\text{N}, \) WGS84 datum, acquisition date May 2006, Fig. 1A). For each unit (e.g., a pixel or a polygon) assigning a membership (\( \mu \)) approaching 1 to each single class is unrealistic. As an example distinguishing grassland from shrubland in ecotones would be practically impossible (Fig. 1B). On the other hand, anthropic-dominated landscapes may show discrete borders among objects (e.g., fields versus roads). In this cases, depending on the scale, discrete mutually exclusive classification can be reliable (Fig. 1C). Similar examples considering different habitat types can be found Wood and Foody (1989, based on lowland heaths) and Rocchini (2010, based on Mediterranean forests). Ahlqvist et al. (2003) and Comber et al. (2005) provide robust critical reviews on the matter.

Regarding the aforementioned loss of information, classification can implicitly degrade information which in turn results in uncertainty in the data and related outcomes. The uncertainty related to the classification process often remains hidden in the output maps, thus it cannot be readily accounted for during further analysis. For instance, maximum-likelihood classification of remotely sensed data simply leads to pixel membership to each class, while additional information generated during the classification process, such as posterior probabilities, could be output (Foody et al., 1992). Although some attempts exist that aim to map and preserve this uncertainty for further analyses (e.g., Ohmann and Gregory, 2002; Ohmann et al., 2011), this issue requires further attention.

According to Rocchini and Ricotta (2007) we will generally refer to uncertainty as both (i) vagueness, namely the lack of sharpness of relevant distinctions, and (ii) ambiguity, arising from conflicting distinctions (discordance, Klir and Wierman, 1999). In this paper, we will review the progress made in geosciences and ecology for taking explicitly into account uncertainty when mapping ecosystems and related environmental phenomena.

2. Uncertainty related to input data for ecosystem mapping

An accurate supervised classification of remotely sensed images requires appropriate ground reference data which are often derived from field training sites. There are many sources of uncertainty in the training stage of a supervised classification, such as class definitions, subjectivity of field data collection and the mixed pixel problem.

Since plant species represent the bulk of habitat structure (Chiarucci, 2007), training sites are often derived from plant sampling-based field surveys, for which one of the main problems lies in the definition of plant communities, an issue raised as early as 1926 by Gleason (1926). A formal definition has not been and will not be accepted globally (Chiarucci, 2007). Moreover, there is an intrinsic difficulty in judging survey completeness (Palmer et al., 2002). This is generally true for all observational sciences; geosciences are not free from such uncertainty as a result of a partial input (Henley, 2006).

There are a number of provoking papers dealing with problems in the discrimination of species in the field, including operator bias (Bacaro et al., 2009), taxonomic inflation (Isaac et al., 2004; Knapp et al., 2005) and more generally taxonomic uncertainty (Guilhaumon et al., 2008; Cayuela et al., 2011), i.e., the subjectivity of field biologists in acquiring species lists which is expected to increase error variance instead of obtaining accurate information on field data.

The effect of imperfect species reference data have been discussed mainly in relation to species distribution modeling (Foody, 2011; Rocchini et al., 2011), in which labeling accuracy together with sample size and pseudo-absence data may lead to biased models of species distribution over space. The same reasoning applies to input field data for generating ecosystem/habitats maps.

Evidence exists about the possibility that abrupt classification of vegetation types, especially at the species hierarchical level, can present misleading or even erroneous results (Schmidtlein and Sassin, 2004). This is due to the often continuous transition of the vegetation assemblages due to changes in environmental gradients (e.g., moisture) and self-organization in vegetation. Alternative approaches like ordination methods aim to extract major floristic gradients describing the variation of the assemblages as metric variables, thus still retaining the continuous

---

Fig. 1. An aerial photo (spatial resolution—1 m, acquisition date: May 2006) of a mountainous landscape of Trentino (Monte Mezzocorona), Northern Italy (A). In case of high heterogeneity (B) the membership to a defined class (e.g., classA—woodland) is not exhaustive (high uncertainty). In case of a homogeneous landscape (C) the membership to a defined class (e.g., classA—crops) may be complete equaling 1. Similar examples can be found in Wood and Foody (1989) and Rocchini (2010).
character of the data (Trotted, 1996; Schmidtlein and Sassn, 2004). These gradients can be related to any sort of remote sensing data set using regression approaches such as generalized linear models or partial least square regression (Wold et al., 2001; Feilhauer et al., 2011).

A second major problem in input data sources for ecosystem mapping and its related uncertainty is the gap perceived between the scale of the field sampling, namely its grain (see Duncan et al., 2002), and the spatial resolution of the image being used, which appears to be a case of incompatible spatial data (Gotway and Atzberger, 2006; Kuusk et al., 2008; Förster et al., 2010). While the concept of a spectral library has been proven for spectrally homogeneous and stable features (e.g., geological formations at coarse spatial scale), the spectral response of species varies with plant phenology, stress, and environmental conditions (Kumar et al., 2001). This variation impairs the transferability of relations between vegetation and spectra and hence affects the use of spectral libraries (Feilhauer and Schmidtlein, 2011). However, if the complete vegetation period can be covered with measurements of field spectra, a relation between remote sensing imagery and a spectral library is possible for a given date of acquisition (Förster et al., 2011), assuming that environmental conditions and the occurrence of plant stress are homogeneous for the mapped area.

### 3. Models accounting for uncertainty to map ecosystems

The attempt to divide a continuous surface of spectral reflectance to create a map of ecosystem properties such as canopy cover, biomass, density and biodiversity is a generalization process with many pitfalls that may result in consequences like the misinterpretation of ecological pattern and process. In fact, such properties are supposed to be internally homogeneous, with sharp differences existing only across boundaries, which is rarely the case. Moreover, in general, apparently simple phenomena (in this case “ecosystem properties”) may show a surprisingly complicated pattern. Relying on Chaos theory (Lorenz, 1964), Li and Yorke (1975) demonstrated it for time series; this definitively applies also to spatial patterns.

It follows that any ecosystem property has an associated error of unknown magnitude, and that the statistical quantification of uncertainty should be a core part of scientific research including using remote sensing in ecosystem mapping.

A number of methods have been proposed to account for uncertainty in an explicit manner when mapping ecosystems. In this review paper we will focus on the most important methods, such as fuzzy set theory and spectral unmixing, providing for a brief overview of additional proposed methods (Table 1).

### Table 1

Recently developed (or rejoined) methods for mapping ecosystems taking uncertainty explicitly into account.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Empirical example</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayesian theory</td>
<td>Application of Bayesian posterior probabilities to classify remotely sensed images</td>
<td>Gonçalves et al. (2009)</td>
</tr>
<tr>
<td>Maximum entropy</td>
<td>Maximum entropy ecological niche modeling to classify different forest types in complex Mediterranean ecosystems</td>
<td>Ferguson (1973), Barry and Hartigan (1992)</td>
</tr>
<tr>
<td>Regression models</td>
<td>Kriging and Generalized Linear Models (GLMs) to account for spatial correlation in residential land use change modeling</td>
<td>Amini (2011)</td>
</tr>
<tr>
<td>Multivariate analysis</td>
<td>Mapping floristic gradients based on relationship between spectral values and floristic data in a Non-metric Multidimensional Scaling space (NMS)</td>
<td>Schmidtlein and Sassn (2004); Schmidtlein et al. (2007), Feilhauer et al. (2011)</td>
</tr>
<tr>
<td>Bootstrap-based procedures</td>
<td>Non-parametric RandomForest (RF) algorithm to map forest succession from LIDAR data</td>
<td>Falkowski et al. (2009)</td>
</tr>
</tbody>
</table>

Please cite this article as: Rocchini, D., et al., Uncertainty in ecosystem mapping by remote sensing. Computers & Geosciences (2012), http://dx.doi.org/10.1016/j.cageo.2012.05.022
3.1. Fuzzy set theory

The concept of fuzzy sets was first introduced by Zadeh (1965) for dealing with vagueness in complex systems. While in traditional discrete classifications information is represented in a one-pixel-one class method (Fig. 2A), dividing the gradual variability of the Earth’s surface into a finite number of mutually exclusive non-overlapping classes, the principle behind fuzzy set theory is that a ‘discrete’ pattern in which classes are mutually exclusive is hard to be found in nature (Rocchini and Ricotta, 2007), and hence the condition of one class being exactly right and all other classes being equally and exactly wrong usually does not exist. Rather, there is a gradual change from membership to non-membership (Gopal and Woodcock, 1994; Foody, 1999) such that fuzzy sets allow for varying levels of class membership for multiple thematic map categories.

Fuzzy sets have been used in a number of fields where abrupt thresholds (classes) cannot represent a suitable model of reality, including: massive data analysis and computation (Jasiwiec, 2011), expert knowledge (Janssen et al., 2010), air pollution (Guo et al., 2007), theoretical topological spaces (Chareeb, 2011), species habitat suitability modeling (Amici et al., 2010), geomorphology (Arrell et al., 2007), soil science (Burrough et al., 2011), expert knowledge (Janssen et al., 2010), air pollution (Guo et al., 2007), theoretical topological spaces (Ghareeb, 1997), and vegetation science (Foody, 1996).

Let $X$ be a universe of objects, whose elements are denoted by $x$. A fuzzy set $F \subseteq X$ is defined by a membership function $\mu_F(x)$ which associates with each $x \in X$ a membership level in the range $[0, 1]$ (the real range between 0 and 1). In the case of remotely sensed images, rather than assigning individual pixels to just a single class, each pixel may be associated with every class with variable degrees of class membership such that $0 \leq \mu_k \leq 1$ and $\sum \mu_k = N$, where $\mu_k$ is the class membership degree of a given pixel for the $k$th thematic map class. The problem of uncertainty assumes a central role. Explicit measures of fuzziness can be applied to each pixel (i) to discriminate between the allocation of mixed pixels to different classes, and (ii) to summarize in map format the additional information concerning residual class memberships. These measures are usually based on either information theory (De Luca and Termini, 1972; Ricotta, 2005), or on more tailored methods depending on the restrictions imposed on the membership degrees. A conceptual drawback related to fuzzy set theory is the deterministic relationship between each object and each class (Zadeh, 1965) which is a paradox since in this case the description of uncertainty is made with class membership values that are not necessarily possible (Schmidtlein and Sassini, 2004). A statement describing a vague phenomenon must be necessarily vague (Sorensen, 1985). This has led to the possibility of assuming for each level of the fuzzy membership function a fuzzy set membership (Fig. 3). This is also known as type 2 fuzzy set, or second order vagueness which has been extended to higher order vagueness concept by Williamson (1999) and Varzi (2003). Refer to Fisher et al. (2007a) and Fisher et al. (2007b) for applied example of higher order vagueness and type 2 fuzzy sets to mountain peak and coastal dune detection, respectively.

For a thorough review on fuzziness measures, see Klir and Wierman (1999).

3.2. Spectral unmixing

As previously stated, pixels can be mixed and contain different land cover classes therein (Fisher, 1997; Foody, 2000), overall in case of medium to low spatial resolution data. This prevents hard classification algorithms from producing accurate results. Quoting Small (2005) “the spectral heterogeneity at pixel scale violates the cardinal assumption of most statistical classification algorithms (e.g., maximum likelihood) wherein each thematic class is...”
Spectral unmixing derives fractions of endmembers (classes) from the spectral differences among classes. This has led to the possibility of assuming for each level of the fuzzy membership function ($s$ axis) a fuzzy set membership ($y$ axis). As an example both P1 (representing a pixel) and P2 show a high membership to a certain class (say class A, $\mu_{\text{first order}} = 0.7$). However, P2 has also a higher value of second order membership ($\mu_{\text{second order}} = 0.8$) to $\mu_{\text{first order}} = 0.7$. In other words, P2 is more certain to belong to the class A than P1.

Fig. 3. A statement describing a vague phenomenon must be necessarily vague. This has led to the possibility of assuming for each level of the fuzzy membership function (x axis) a fuzzy set membership (y axis). As an example both P1 (representing a pixel) and P2 show a high membership to a certain class (say class A, $\mu_{\text{first order}} = 0.7$). However, P2 has also a higher value of second order membership ($\mu_{\text{second order}} = 0.8$) to $\mu_{\text{first order}} = 0.7$. In other words, P2 is more certain to belong to the class A than P1.

3.3. Additional rejoined methods

Besides the techniques described above, additional approaches have been tested by means of elegant statistics to map ecosystems relying on, e.g., posterior probabilities of occurrence of habitats instead of discrete sets relying on e.g., Bayesian theory (Gonçalves et al., 2009; Amici et al., 2010), maximum entropy ecological niche modeling (Amici, 2011), posterior probabilities of

land use change by geostatistical modeling (Braimoh and Onishi, 2007), Dirichlet process partitioning (Ferguson, 1973), product partition models (Barry and Hargitan, 1992; see Quintana, 2006 for a review), covariance matrices analysis for finding homogeneous vectors (Meidow et al., 2009), indicator geostatistics (Boucher and Kyriakidis, 2006), generalized linear models for mapping tree cover continuous fields (Schwarz and Zimmermann, 2005), continuous mapping of vegetation properties relying on multivariate analysis (Schmidtlein et al., 2007), Monte-Carlo permutations to depict conservation priorities by uncertainty maps (Gallo and Goodchild, 2012), Random Forest algorithm to map forest succession detecting areas with high structural variability (Falkowski et al., 2009). Table 1 summarizes such methods with empirical examples.

4. Concluding remarks

Uncertainty as a core issue of ecosystem mapping has been repeatedly tested and recognized over the past decades. It is impossible to measure with precision all the facets of a complex phenomenon (Hartley, 1928; Schrödinger, 1935; Ricotta and Anand, 2006; Anand et al., 2010), because the act of measuring itself has an effect on the perception of such a phenomenon (Brown, 2004). Moreover, there is a lack of a comprehensive theory to explain the geometry of environmental variation (environmental gradients) over space (Palmer, 2007), with some environmental gradients varying smoothly and some others being more chaotic.

A way for taking into account the unpredictability of spatial variation is to rely on the probability of belonging to a certain class instead of abruptly applying a classification algorithm based on image thresholding, like in maximum likelihood classification. From a theoretical point of view this is similar to approximation theory in mathematics, in which, once searching for a function which best approximates a more complex one, the characterization of the introduced errors (uncertainty) is of primary importance (e.g., Fourier, 1822). In this paper, we reviewed the main models used in mapping ecosystems to properly account for uncertainty, overall in cases of high ecosystem complexity.

Models, such as fuzzy-sets, have been implemented in Free and Open Source Software (FOSS) like R (R Development Core Team, 2012, packages fuzzyFDR, Lewin, 2007, FKBL, Alvarez, 2007, reviewed in Meyer and Hornik, 2009) or GRASS GIS (Neteler et al., 2012, rs.fuzzy.system module, Jasiewicz, 2011) and are readily available for further improvement. The importance of FOSS has been widely acknowledged (Stallman, 1985) since it allows to reproduce complex algorithms with the possibility to access the code, which enables worldwide contributions from the scientific community. Refer to Rocchini and Neteler (2012) for a critique about the importance of open source code in ecological studies.

Efforts have also been made to extract reliable measures of the error and obtain effective ways of error visualization (Gonçalves et al., 2009), generally measured by posterior probability estimators (Stehman and Czaplewski, 1998). Box 1 presents good practices in reliable accuracy assessment of classification models used for ecosystem mapping.

This paper provided a review of some of the recent issues relating to uncertainty in ecosystem mapping from remote sensing, particularly focusing on vagueness and ambiguity (see Section 1). The reader is also referred to Blaschke (2010) for a review on geometric (boundary) uncertainty of recognized shapes and to Hay et al. (2005), Nol et al. (2008) and Schulp and Alkemade (2011) for empirical examples.

We hope to stimulate discussion about the potential of uncertainty-related models for mapping ecosystems which
should allow researchers in geosciences and spatial ecology to attain proper representations of ecosystems over space.

Box 1—Good practices in accuracy assessment of ecosystems maps

The accuracy of ecosystem maps may simply rely on raw accuracy, i.e. the congruence between map classes and classes identified in the field. Generally it is based on a representative (random) set of sampling units \( n = \{n_1, n_2, n_3, \ldots, N\} \) acquired by an appropriate sampling design whose class \( y_i \) is to be checked in the field (also known as post-classification sample, Steele, 2005).

Error matrices of observed versus predicted classes are then produced and accuracy measures are derived (see Congalton, 1991, Foody, 2002, on the user versus producer accuracy). The set \( n \) can also be achieved by resampling training sites and calculating final accuracy, i.e. reusing the training data for accuracy assessment. Among the methods being used for calculating accuracy, maximum posterior probabilities proposed by Steele (2005) seems to be the most straightforward since it can be adapted to any sampling design being used.

It is worth remembering that experts going in the field are generally obliged to specify one single class per plot by next making a binary comparison (right or wrong) with the map labels. Hence, this leads to difficulties in properly treating uncertainty. In this view, soft accuracy assessments of map classification have thus been proposed based on e.g. a linguistic value like ‘absolutely right’, ‘good answer’, ‘reasonable’, etc. (Woodcock and Gopal, 2000) or on different fuzzy accuracy levels considering the nature, magnitude, and frequency of errors at different levels of taxonomic resolution (LaBA et al., 2002). Additional examples include the use of fuzzy error matrices for accuracy assessment when field and/or classified data are in form of ‘multimembership’ values (Binaghi et al., 1999). Further, also in case of hard classification procedures, the adoption of variability-based quality assessment has been proposed relying on e.g. confidence intervals to accuracy estimates (Foody, 2009).

Refer also to Strahler et al. (2006) who set out best practices in terms of global mapping.

Acknowledgments

We are grateful to the Editor Prof. Jef Caers and two anonymous referees for useful insights which improved a previous version of this manuscript. Prof. Roberto Imperiale provided useful mathematical insights. DR is partially funded by the Autonomous Province of Trento (Italy), ACE-SAP project (registration number 23, June 12th 2008, of the University and Scientific Research Service). This study has been partially funded by the Earth and Space Foundation Award (http://www.earthandspace.org/).

References


