Fuzzy difference and data primitives: a transparent approach for supporting different definitions of `forest` in the context of REDD+

Alexis Comber[1], Werner Kuhn[2]

[1] School of Geography, University of Leeds, Leeds, LS2 9JT, UK

email: a.comber@leeds.ac.uk

[2] Department of Geography, University of California Santa Barbara, USA

email: werner@ucsb.edu

Abstract

This paper explores the use of fuzzy difference methods in order to understand the differences between forest classes. The context for this work is provided by REDD+, which seeks to reduce the net emissions of greenhouse gases by rewarding the conservation of forests in developing countries. REDD+ requires that local inventories of forest are undertaken and payments are made on the basis of the amount of forest (and associated carbon storage). At the most basic level this involves classifying land into forest and non-forest. However, the critical issues affecting the uptake, buy-in and ultimately the success of REDD+ are the lack of universally agreed definition of forest to support REDD+ mapping activities, and where such a definition is imposed, the marginalization of local community voices and local landscape conceptualizations. This tension is at the heart of REDD+. The methods described in this paper to address these issues link methods to quantify fuzzy land cover change to notions of data primitives, which have been previously proposed as an approach to move between land cover classes with different semantics. Together these methods provide for transparent translations between alternative conceptualizations of forest, allowing for plural notions of forest to be mapped and quantified. In particular, they allow for moving from an object-based notion of forest (and land cover in general) to a field-based one, entirely avoiding the need for forest boundaries.

1. Introduction

This paper suggests and applies an alternative approach to classifying and mapping land cover, using the example of forest mapping in the context of Reducing Emissions from Deforestation and forest Degradation (REDD and REDD+). REDD+ initiatives seek to reduce the net emissions of greenhouse gases by rewarding the conservation of forests in developing countries (Angelsen, 2009). REDD+ requires that local inventories of forest are undertaken and payments are made on the basis of the amount of forest (and associated carbon) that are mapped. At the most basic level this involves classifying land into forest and non-forest.

The problem encountered by many forest monitoring strategies under REDD+ is how to accommodate divergent conceptualizations of *forest* such that local interpretations are reflected in forest mapping while providing a transparent tool for carbon accounting. The International Panel on Climate Change (IPCC) recommends the use of an internationally recognized forest definition (e.g., that of the Food and Agricultural Organisation, FAO) but even within the United Nations Framework Convention on Climate Change (UNFCCC) there are no agreed definitions (Romijn et al, 2013). The flexibility of REDD+ allows forest to be defined locally, but ultimately requires mappings to be transparent (Gupta et al, 2014; Ryan et al., 2014), in order for forest estimates to be *accurate* and payments to be justified. There are well established reasons for not having a standard definition of forest (see Comber et al 2007, for an extensive discussion on the problems caused by the imposition of standards in land cover mapping). Not least among these reasons is that local considerations and landscape conceptualizations would be ignored (Hoeschele, 2000; Robbins 2001) to the detriment of local stakeholders. However, there can be considerable disagreement over the amount of forest, when mapped under different classifications. For example, Romijn et al (2013) compared the forest stock levels in Indonesia when mapped using local forest definitions and FAO definitions and found the extent of forest to be 27% higher when using the national definition.

The critical issue addressed by this paper is how to allow for the divergence of landscape conceptualizations, avoid divergent and possibly inaccurate estimates of forest extent. The paper proposes a method based on a mix of data primitives and soft classifications that allow the critical dimensions associated with land cover and land use to be captured **and** the inherent relative uncertainty when any two mappings are compared. This approach provides a shared, transparent, robust and well understood method for mapping and reporting forested areas. It is transparent and supports global policy initiatives such as REDD+, whilst at the same time accommodating local divergence in forest semantics - and the way that 'forest' (and therefore the boundary with non-forest) is conceptualized spatially.

Our analysis uses existing classifications of forest from the Global Land Cover (GLC) and from FAO and shows how these can be re-interpreted using soft classifications to explore the overlaps and boundaries between different versions of forest. In this way the paper develops approaches that could be used in the context of REDD+ initiatives to support multiple views of the same landscape to support divergent objectives. Such transparency can overcome some of the problems of exclusion and back door land reforms that are commonly experienced during REDD+ initiatives and mapping (Ezzine-de-Blas, 2011), as well as the issues highlighted by Romijn et al (2013) of how to define 'forest'. In our approach, one can define forest however one wants and link it to another definition. # 2. Background

To situate our approach, we first provide the necessary background on land cover classification, the problems associated with forest definitions, and the method of fuzzy boundaries and fuzzy differences.

2.1 Land cover classification

Classification, the process of sorting real world phenomena into categories or classes, is a core activity within geography and many other disciplines. Classes describe groups of social and environmental phenomena with broadly similar characteristics. Land cover classes are used to describe the physical characteristics of the earth's surface and land use classes describe the socio-economic activities thereon. Typical land cover classes include water, grass, and forest; typical land use classes include agriculture, forestry, and urban. There is an extensive literature on the different philosophical underpinnings of the way that land cover and land use classes are conceptualized (e.g.~Comber et al., 2005; Fisher et al, 2005), also for specific classes such as forest (Bennet, 2001, Comber et al, 2008). This application based literature is complemented by more general considerations of the nature of classification (e.g.~Lakoff, 1987) and by more formal research considering the nature of how to represent objects, processes, and relationships of geographic phenomena in computer models and analyses (e.g.~Kuhn 2005, 2012).

Classification typically allocates each item uniquely to one class and this is true for land cover and land use, although much theoretical work exists suggesting alternative, soft classifications such as those supported by Fuzzy Sets (e.g.~Fisher 2010). All global land cover data produced for monitoring purposes and used as inputs to global climate change models adopt a crisp classification with hard and crisp boundaries between the resulting land cover objects of different classes. Thus, land cover classification is a process of homogenization that ignores any within class variation and, more critically, excludes any consideration of the potential for uncertainty in boundaries between classes. This contrasts with a well developed literature on boundary uncertainty within geography (Smith, 1995; Burrough and Frank 1996), reflecting a vernacular definition of geography as the art of drawing lines on a map that do not exist in reality.

Classifications with crisp boundaries provide a convenient and familiar framework for representing environmental and social processes that are readily understood by policy makers and initiatives. For example, in the sphere of climate change, REDD+ and climate models seek to standardize or

```
harmonize' definitions of forest and develop crisp mappings of forested areas that may
not exhibit crisp edges in reality and whose classification may be deeply contested.
```

This culture of class definition standardization for crisp mapping creates baseline inventories that ignore local variations in how land cover classes such as forest are conceptualized, their semantics and the functions associated with them. This can have serious political implications for the way that land is managed locally and for local political discourses around land. This has become evident with the roll out of REDD+ in many developing countries, where fixed IPCC / UN definitions offorest', for the purpose of carbon accounting, are being used to drive land reform and privatization agendas. Local land cover interpretations are excluded from local and global decision making.

2.2 Forest definitions and REDD+

The REDD and REDD+ initiatives (Reducing Emissions from Deforestation and forest Degradation), as set up by the United Nations Framework Convention on Climate Change (UNFCCC) in 2005 and enhanced in 2007 to stimulate actions related to conservation and sustainable forest management (hence the +'), aim at mitigating climate change by reducing net emissions of greenhouse gases by financially rewarding enhanced forest management in developing countries. In simple terms, poor countries with forests are to be paid to \emph{not} cut down their trees, with payments based on the amount of trees and carbon storage capacity they have, although the precise details are still to be decided. In order to support payments, the UN have suggested that countries should develop robust and transparent forest monitoring system to record forest and carbon stocks and changes using a combination of remote sensing and field surveys (UNFCCC, 2009), with each country defining what constitutes \emph{forest}. This was to overcome the deeply contested nature of what is aforest' (and land cover / land use classification in general), well established in the literature (see for example Hoeschele 2000; Bennet 2001; Robbins 2001; Smith and Mark 2001; Comber et al, 2005) and reflects commendable courage by the remote sensing community that is driving the REDD+ initiative. However, REDD+ activities require forest to be demarked, presenting a number of challenges, two of them being of critical importance.

First, any particular definition of forest and the associated mapping will determine the reported amount of forest changes, which will in turn affect the estimates of carbon storage (Magdon and Kleinn, 2012; Romijn et al, 2013). There are hundreds of forest definitions as documented by Lund (2015) and explored conceptually by Bennett (2001) and Comber (2005) with profoundly different threshold parameters for height, stripwidth, minimum area, species (some include grasses such as bamboo), canopy, management / plantation cover etc. These are international, national and sub-national and reflect locally important variations in the concept of forest' and its semantics - what the concept is in the local (not just REDD+) context. In the context of REDD+, Morales-Barquero et al (2014) highlight the difficulties in operationalizing local and national level REDD+ projects and programmes in Mexico. They present a framework to support local definitions and measurements of forest degradation that support both local and national objectives and suggest the use of comparative biophysical benchmarks for assessing degradation. More recently, Chazdon et al (2016) examined historical forest concepts and definitions and documented how these relate to variations in socio-economic activity. They note a number of problems with top-down forest definitions for assessing global changes in forest stocks: their failure to distinguish between natural / plantation forests and the lack of consideration of the qualities and trajectories of forested areas in standard approaches. In order to try to accommodate this divergence in the concept of forest, and to support consistency in carbon accounting, the UN requires robust and transparent national reporting of the measurement systems, the remote sensing and ground based forest carbon inventory data, the methods and the forest definition used in national measurements (Herold et al., 2012a, 2012b). The impacts of such semantic variations are well documented. Romijn et al (2013) compared measures of deforestation using 3 definitions of forest in Indonesia - the FAO definition, anatural forest' definition and a national forest definition - and found large differences between the amounts of deforestation.

Second, national definitions and associated mappings are deeply political exercises that are frequently used to drive secondary local political agendas that exclude and are in conflict with local communities. Mapped forests, however they are defined, are presented locally as 'facts', which are then manipulated towards the interests of the state, excluding community views, perceptions and opinions from the classification and measurement activity. For example, there are cases of REDD+ projects being used locally to support the violent eviction of people from the land in many countries (Himmelfarb, 2012; Cavanagh and Benjaminsen, 2014; Grainger and Geary, 2011; Nel and Hill, 2013; Lyons and Westoby, 2014; Forest Peoples Programme, 2014; Beymer-Farris and Bassett, 2012). In other cases, national definitions are imposed and the subsequent mappings are used to support land privatization agendas and forest commodification with the objective of removing commonly held land from collective ownership (Ece, 2015). Critical to the success of REDD+, there are also situations identifying the financial beneficiaries of carbon payments because the forested land lacks natural ownership divisions. A further negative dimension of REDD+ is that it promotes the obfuscation and overlooking of the opinions and views of forest-dependent communities and indigenous forest-dwelling populations in developing countries. Local understanding and knowledge are important for the forest conservation objectives of REDD+. For example, Bong et al. (2016) compared community knowledge of deforestation and degradation drivers amongst different Indonesian villages and found local knowledge to be key to understanding the local impacts of deforestation drivers. They highlighted the importance of incorporating local knowledge and conceptualizations within definitions of forest to develop more locally appropriate REDD+ monitoring systems. However, some communities in likely REDD+ countries do not conceptualize forest at all (Niclas Burenhult, pers comm) and where they do this has very nuanced and spiritual meaning. Specifically, although the people of Jahai in the in the Malay Peninsula live in forests, they have no concept of forest (Burenhult, 2009). Instead it is their home and the nearest terms they have for forest like things describe leaves and trees as well as canopy and floor, covered area and exposed area.

The conclusion is that the requirement of REDD+ to demarcate forest is problematic in many places. Crisp mappings of forest do not reflect the many cultural, linguistic and ownership aspects of forest, however the category is defined. Initiatives like REDD+, however well meaning they are, result in a clash of categories with the consequence that important forest related parameters are overlooked. These wide ranging negative implications for forest-dependent communities and indigenous forest-dwelling populations threaten forest conservation in developing countries and as a consequence have been referred to as the dark side of REDD+ (Cavanagh et al., 2015).

2.3 Fuzzy Boundaries and difference

Using fuzzy land classifications each piece of land (for example, the piece corresponding to a pixel in a satellite image) is allocated a degree of membership to each class in the range {[}0,1{]} (Fisher and Pathirana, 1990). This allows each piece of land to have partial membership to more than one land cover class. For these reasons fuzzy approaches are frequently referred to as *soft* classifications as the fuzzy outputs include and represent some of the inherent uncertainty associated with allocating image elements (such as pixels) to classes (Fisher, 1997). This is in contrast to *hard* or Boolean approaches, which do not accommodate any uncertainty. For these reasons fuzzy sets and fuzzy classification algorithms have been suggested as appropriate approaches in remote sensing analyses for representing land cover objects which may or may not fit neatly into the classification scheme and / or the sampling frame (pixel).

In a standard land cover classification, class summaries or centres are created that describe the (proto-)typical class properties usually in an n-dimensional feature space, for example representing different remote sensing image bands, or as in this case data primitives, described below. In a Boolean classification each piece of land is allocated to the class to which it is nearest. Different algorithms create the class centres in different ways and measure distance in n-dimensional space in different ways. The issue that Fuzzy classifications seek to address is that a piece of land may be *near* to more than one class, suggesting that it contains some of their properties and resulting in a degree of ambiguity (Fisher et al., 2006a) in the allocation.

A second commonly arising problem that fuzzy classifications seek to address relates to the sampling frame. The assumption in mapping land cover using remote sensing is that the objects of interest on the ground are adequately described and captured by the imagery. The inference is that the spatial scale of the processes on the ground (such as land cover) are larger that then the pixel and that contiguous pixels will describe the extent of these processes. In reality there are many sub-pixel objects on the ground (Fisher, 1997) arising from a mismatch between the scale of observation and the granularity of the processes being observed (Fisher, 2007). Fuzzy sets seeks to accommodate a number of problems related to the pixel: sub-pixel objects (less than the size of a pixel), trans-pixel objects (linear objects crossing pixels), intergrades (where land cover types merge into each other over space) and boundaries (where multiple covers meet within one pixel).

There is a long standing body of research describing the application of fuzzy sets to land cover and this has been extended by some researchers who have developed methods for related analyses: fuzzy accuracy, fuzzy change detection, Type-2 and Type-N fuzzy approaches (Fisher, 2010; Fisher et al., 2007). Conceptually, fuzzy approaches have potential value to operational land cover monitoring exercise such as REDD+ because they can explicitly accommodate the uncertainty inherent to classification. Here these the concepts of fuzzy Minimum Interval, fuzzy Bounded Difference and fuzzy Loss and Gain are extended to examine the differences between two fuzzy classifications as described in the next section.

2.4 Land cover objects vs land cover fields

The land cover classification process described above results in crisp and hard boundaries between classes in attribute space as well as between the resulting regions on the ground. It is therefore tempting to interpret these regions as land cover objects, with boundaries and identity. To quantify a country's amount of, for example, forest, one simply adds the surface areas of all forest objects in that country at a given time.

Yet, for the reasons stated in the previous subsection, such an outcome of land cover classification is not satisfactory. In particular, for the case of forests, it is unsuitable to deal with partial deforestation and forest degradation. Forests do not simply disappear "from their boundaries inward" by converting pieces of forests to non-forests. Rather, deforestation and forest degradation proceed as gradual reductions of *forestness* of pieces of land. They may occur, for example, through a systematic thinning of tree density.

It turns out that fuzzy boundaries and differences, or rather fuzzy membership values for pieces of land in a land use category, provide exactly the mathematical model needed to deal with such an idea of forestness. As the following analysis section will show, they generate feature vectors for each piece of land corresponding to, for example, a pixel. In Geographic Information Systems (GIS), such arrays of vectors of observed values are known as vector fields. Thus, using the classic dichotomy of fields vs objects (Couclelis 1992), we conclude that forests are much more adequately conceptualized as fields rather than objects.

3. Analysis

o illustrate the proposed method, this paper uses the case study of translating land cover data describing forest in North America to another classification. In this case we convert the 6 GLC forest classes to the FAO class of forest. The approach is to characterize the classified data using data layers referred to as data primitives (Comber et al., 2008; Wadsworth et al, 2008) which enable each class to be described at the most fundamental level. In the terminology of GIS field models, data primitives are the dimensions of the forest field vector. Here we use data describing vegetation height, photo-synthetic activity, soil wetness, human disturbance and seasonality. The positions of each class in these dimensions are extracted. Then a new class of forest is defined using the data primitives, and subsequently mapped. The differences between the two soft, fuzzy mappings of 'forest' within the primitives are explored and shown to provide a transparent and translatable comparison of different definitions of forest.

3.1 Data

The Global Land Cover (GLC2000) data were downloaded and extracted for the USA. Five datasets to provide measures that approximated to the dimensions of vegetation height, photo-synthetic activity, soil wetness, human disturbance and seasonality were obtained from the following open data portals:

Table 1: The sample number and total distribution of the GLC2000 classes.

	sample	total no
Tree Cover, broadleaved, deciduous, closed	24219	104142
Tree Cover, needle-leaved, evergreen	32841	141220
Tree Cover, mixed leaf type	13407	57652
Herbaceous Cover, closed-open	3592	15449
Sparse herbaceous or sparse shrub cover	149	641
Cultivated and managed areas	23030	99030
Water Bodies	1190	5118
Artificial surfaces and associated areas	1569	6748

Table 2: The median values for each GLC2000 in the 5 primitive dimensions: VH = Vegetation Height; PAR = Photosynthetic Activity; SW = Soil Wetness; LHD = Lack of Human Disturbance; GS = Seasonality.

	vh	pr	sw	hd	gs
Tree Cover, broadleaved, deciduous, closed	24.6	8018.6	339.3	5.4	204.1
Tree Cover, needle-leaved, evergreen	20.1	7673.4	281.2	5.8	214.3
Tree Cover, mixed leaf type	27.5	9701.6	304.9	10.3	200.7
Herbaceous Cover, closed-open	13.9	3415.1	235.2	2.6	205.2
Sparse herbaceous or sparse shrub cover	6.0	130.0	298.9	2.1	192.7
Cultivated and managed areas	11.5	2407.0	337.1	0.0	204.3
Water Bodies	13.1	8160.3	303.7	1.8	212.1
Artificial surfaces and associated areas	7.3	367.3	283.4	0.0	214.5

- **Vegetation height**: from *Existing Vegetation Height* data from http://www.landfire.gov/version_comparison.php?mosaic=Y
- **Biomass**:from http://landcarbon.org/categories/biomass-c/download/ g Carbon/m2.
- **Soil wetness / soil moisture**: from http://giovanni.gsfc.nasa.gov/giovanni for the month of June for 0-100cm depth
- **Human disturbance**: proximity to mapped Urban and Managed land from GLC (classes 16, 17, 18, 22)
- **Seasonality**: phenology growing season data from the MODIS archive: Duration DUR Length of photosynthetic activity (number of days): https://lta.cr.usgs.gov/emodis_phen

These were all resampled to 0.0217 of a degree with the mean value used to aggregate the data, and clipped to a smaller case study area. The spatial distributions of these variables and the GLC2000 data are shown in Figures 1 and 2.

3.2 Class centres

The typical characteristics of each of the GLC2000 classes in each of the five dimensions were extracted and are shown in Table 1. This was done by selecting a sample of GLC2000 pixels from each class, with the sample number weighted by the distribution of the class in the study area.

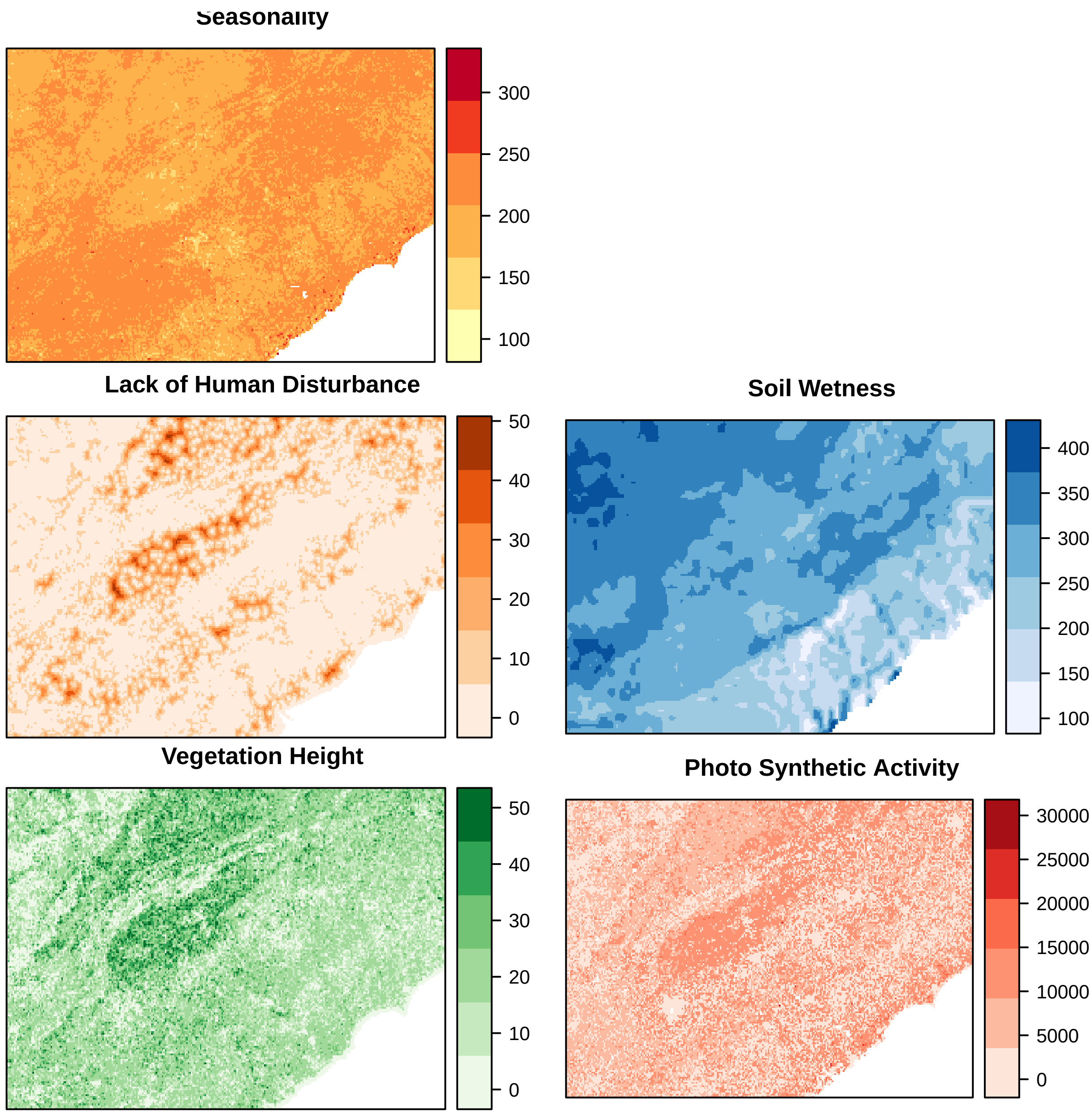

Figure 1: The data primitives.

Table 3: The median values for each FAO Forest classes in the 5 primitive dimensions.

	vh	pr	sw	hd	gs
FAO Forest	25	8665.5	309.2	7.3	205.7
FAO not-Forest	1	130.0	348.6	0.0	206.3

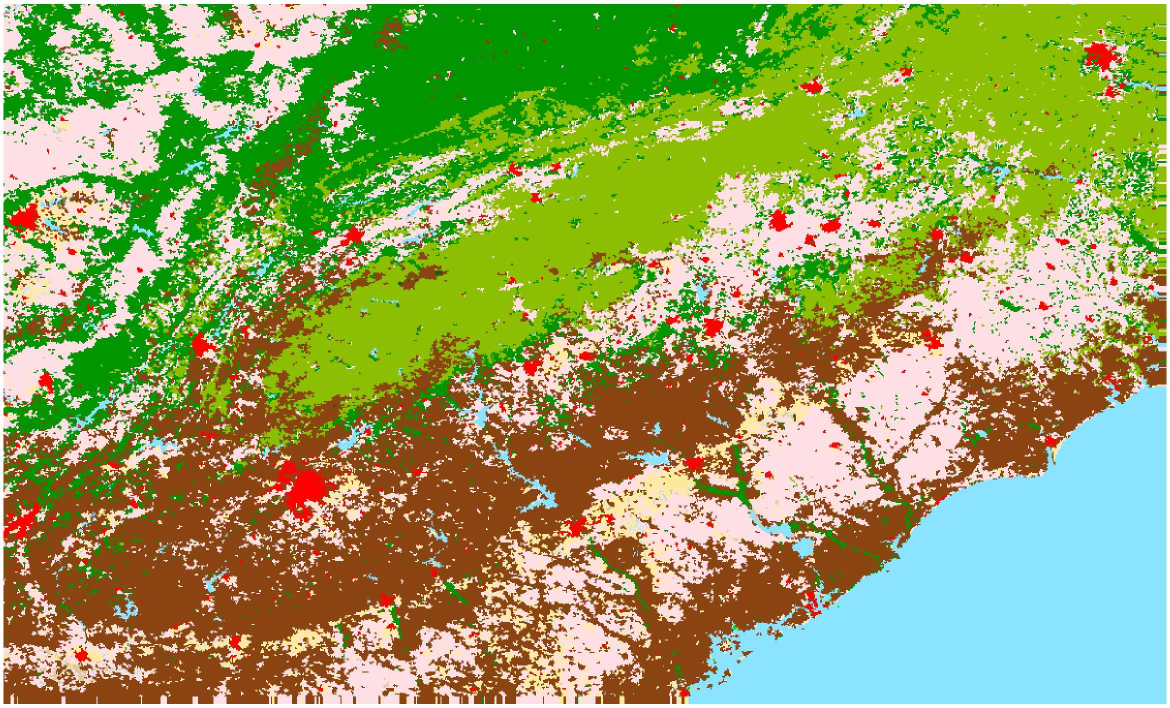

Figure 2: GLC2000 data: green and brown are forest classes, pink is agriculture and red is urban.

3.3 Fuzzy Classification: GLC2000 and FAO

A fuzzy c-Means classifier was used to generate soft classifications for each pixel in the study area. This algorithm used was the implementation of the c-Means fuzzy clustering method (e.g.~Bezdek, 1981) in the `e1071` R package (Meyer et al., 2012). The classifier generates fuzzy memberships in the range of {[}0, 1{]} to each class for each pixel. In this case the class centres were those in Table 2 and the classifier generated fuzzy memberships to the et of each class in the five primitive dimensions. The fuzzy c-means algorithm seeks to to minimise the objective function:

$$\sum_i \sum_j w_i u_{ij}^m d_{ij} \tag{1}$$

where w_i is the weight of observation (pixel) i, u_{ij} is the membership of observation i in cluster (class) j, and d_{ij} is the Euclidean distance (dissimilarity) between observation i and centre j. The fuzzy cmeans algorithm generates a membership to each class for each pixel, based on their closeness to the class centres in Table 2.

The fuzzy memberships can be used to extract crisp values of the amount of forest and non-forest, depending on the classes of forest that are preferred, through the application of different alpha cuts. This in turn allows the amount of carbon to be measured relative to the way that *forest* is defined.

Then by combining all of the memberships for the three forest classes, a fuzzy class of forest from the GLC2000 classes was generated.

Finally, a second forest class was introduced. This was based on the he FAO forest classes forest class definition (http://www.fao.org/docrep/006/ad665e/ad665e06.htm) and included the following key characteristics: tree crown cover, more than 10 percent canopy cover, minimum height of 5 metres although young and 'temporarily unstocked' (!) land can be included. By selecting a sample of forest pixels with the characteristics of being greater than 5m and have greater than 10% canopy cover, a second set of class centres in the five primitives was generated. These are shown in Table 3. A second fuzzy classification of forest was then generated using the c-Means algorithm above, from the class centres.

3.5 Fuzzy boundaries and difference

The method for quantifying fuzzy differences between 2 classification is based on that suggested by Fisher et al (2006b). In this approach, the land cover at any location is considered to have potential memberships to each of the different classes being considered. That is, class membership is considered as a fuzzy set. Fuzzy change is described using the fuzzy change matrix which supports the calculation of fuzzy losses and gains. This process is exemplified below with a walk-through example.

The fuzzy confusion matrix, describing the intersection between the two fuzzy forest classes can be derived using the minimum interval or using the bounded difference of fuzzy loss and fuzzy gain. As Fisher (2006b) noted the minimum interval is the standard approach in fuzzy sets but is counter-intuitive when it is used to compare different classes – it only makes sense in the context of fuzzy land cover when comparing fuzzy sets of the same class but captured at different times (i.e.~the diagonal in a correspondence matrix). For these reasons a number of alternative operators have been suggested (Klir and Yuan, 1995; Leung, 1988; Zadeh, 1965) and as Fisher et al (2006b, p166) note *Of these, the Bounded Difference between two fuzzy sets is the simplest operator for which the results make the most sense.*

The notion of bounded difference underpins the fuzzy change matrix as follows: the minimum interval is used to calculate the fuzzy changes between fuzzy sets of the same class and the bounded difference is used to determine all other fuzzy changes (i.e.~off-diagonal changes in change matrix). Fuzzy loss and gain are derived from the marginal totals of a change matrix describing transitions between land covers at Time 1 and Time 2. From the estimates of the amount (area) of fuzzy change between each class can be obtained as well as the total amounts of loss and gain.

Minimum Interval

In a standard fuzzy sets approach the minimum operator is used to determine the intersection between 2 fuzzy classifications, C_2 and C_2:

$$\mu(C_2, C_2) = min(\mu(C_1, C_2)) \tag{2}$$

In the analyses below, this is used to determine the diagonal elements in the fuzzy difference matrices.

Bounded Difference

The bounded difference describes the fuzzy intersection of 2 land cover types. It provides a conservative estimate of change because of its logic and mathematics: essentially if the sum of fuzzy memberships at any location are less than 1 then the membership to the intersection – i.e.~fuzzy change – will be zero. The bounded difference between 2 fuzzy sets A and B is defined as:

$$\mu(A \cap B) = max(0, \mu(A) + \mu(B) - 1) \tag{3}$$

Loss and Gain

In the methods suggested by Fisher et al (2006) the bounded difference (Equation 3) is used to quantify the off-diagonal loss and gain at each location (pixel) from the marginal totals of the fuzzy change matrix.

Worked example

Consider 2 fuzzy classifications at 2 different classifications (C_1 and C_2) of 3 classes each (a, b and c) on a 4 by 4 pixel grid, as described in Table 3. The diagonals on the fuzzy change matrix are populated by the sum of the minimum intersections and the off diagonals by the sums of the bounded difference.

Table 4: Example fuzzy memberships for a hypothetical classification over 4 by 4 pixel dataste with 3 classes (a1, b1 and c1)

a1				b1				c1			
0.5	0.50	0.43	0.33	0.33	0.33	0.14	0.17	0.17	0.17	0.43	0.5
0.5	0.38	0.29	0.33	0.33	0.25	0.29	0.17	0.17	0.38	0.43	0.5
0.2	0.20	0.40	0.40	0.60	0.60	0.20	0.20	0.20	0.20	0.40	0.4
0.2	0.17	0.40	0.40	0.60	0.50	0.20	0.20	0.20	0.33	0.40	0.4

Table 5: Example fuzzy memberships for a second hypothetical classification, again with 3 classes (a2, b2 and c2)

a1				b1				c1			
0.5	0.50	0.43	0.33	0.33	0.33	0.14	0.17	0.17	0.17	0.43	0.5
0.5	0.38	0.29	0.33	0.33	0.25	0.29	0.17	0.17	0.38	0.43	0.5
0.2	0.20	0.40	0.40	0.60	0.60	0.20	0.20	0.20	0.20	0.40	0.4
0.2	0.17	0.40	0.40	0.60	0.50	0.20	0.20	0.20	0.33	0.40	0.4

To illustrate this, the table of minimum interval for class a in classes $a1$ and $a2$ is shown in Table 5, along with the bounded difference between classes $b1$ and $c1$. The derived fuzzy correspondence matrix is shown in Table 6.

4. Results

4.1 Fuzzy Forest

The fuzzy memberships to the three forest classes, whose class centres are listed in Table 2, are shown in Figure 3, along with a combined membership for all GLC2000 forest classes. The class centres for the forest class arising from FAO class definition are shown in shown Table 8. The result of the fuzzy C-Means classification is shown in Figure 4.

4.2 Boundaries

Having generated the fuzzy classifications, the resulting boundaries between the two forest classes (GLC2000 and FAO) can now be compared. It is instructive to consider a sub-region in greater detail, in this case an area to the North and West of Atlanta, Georgia. Figure 5 shows the GLC-derived and FAO-derived fuzzy forest distributions, with greater detail in Figure 6. It is evident that the FAO membership values are greater in many places than the GLC memberships and this is reflected in the summary statistics in Table 8. Figure

Table 6: The minimum interval for class a and the bounded difference between classes $b1$ and $c1$.

a1-a2				b1-c2			
0.5	0.5	0.25	0.25	0	0	0	0
0.38	0.38	0.25	0.25	0	0	0	0
0.14	0.14	0.4	0.4	0.03	0.03	0	0
0.14	0.14	0.25	0.4	0.03	0	0	0

Table 7: The fuzzy difference matrix with loss and gain for the worked example.

	a	b	c	loss	gain
a	4.77	0.00	0.00	0.00	0.00
b	0.00	4.45	0.09	0.09	0.00
c	0.00	0.00	4.40	0.00	0.09

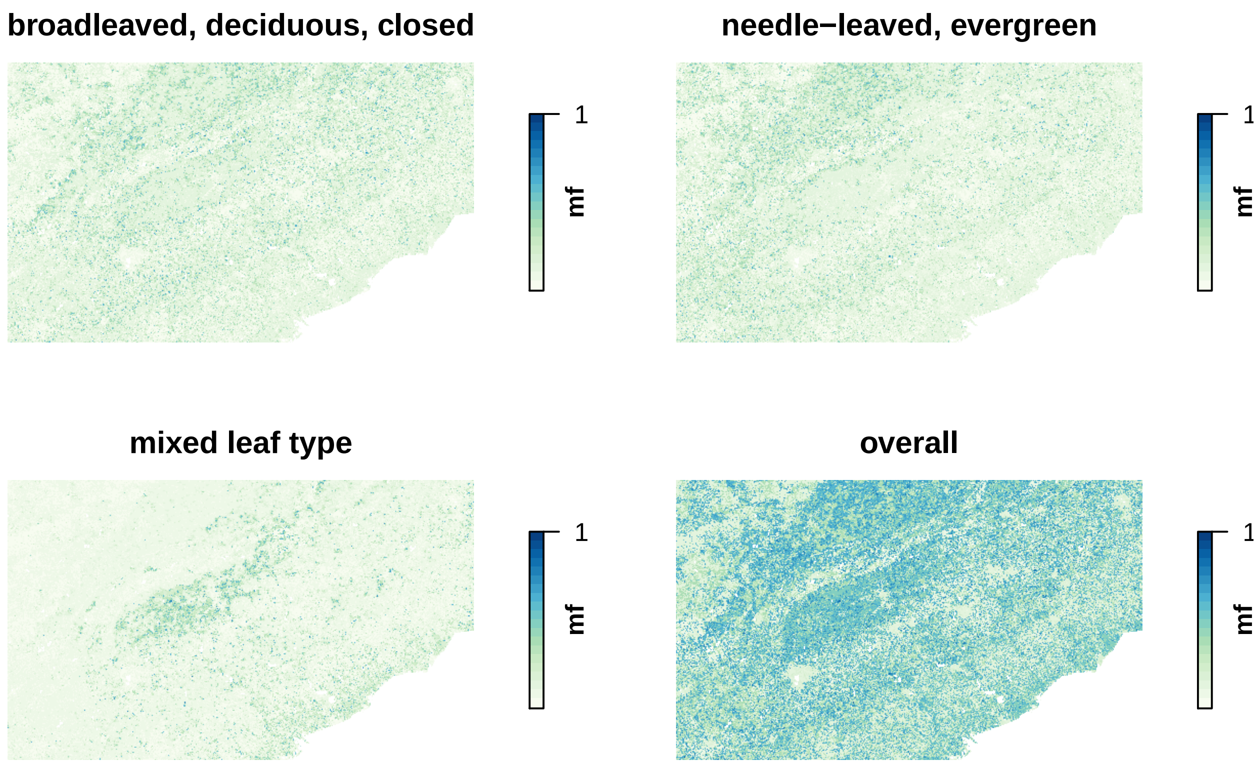

Figure 3: The fuzzy membership to the 3 GLC2000 forest classes in study area, and their combined fuzzy memberships to an overall forest class.

Table 8: The median values for the FAO Forest and not-Forest classes in the 5 primitive dimensions.

	vh	pr	sw	hd	gs
FAO Forest	25	8665.5	309.2	7.3	205.7
FAO not-Forest	1	130.0	348.6	0.0	206.3

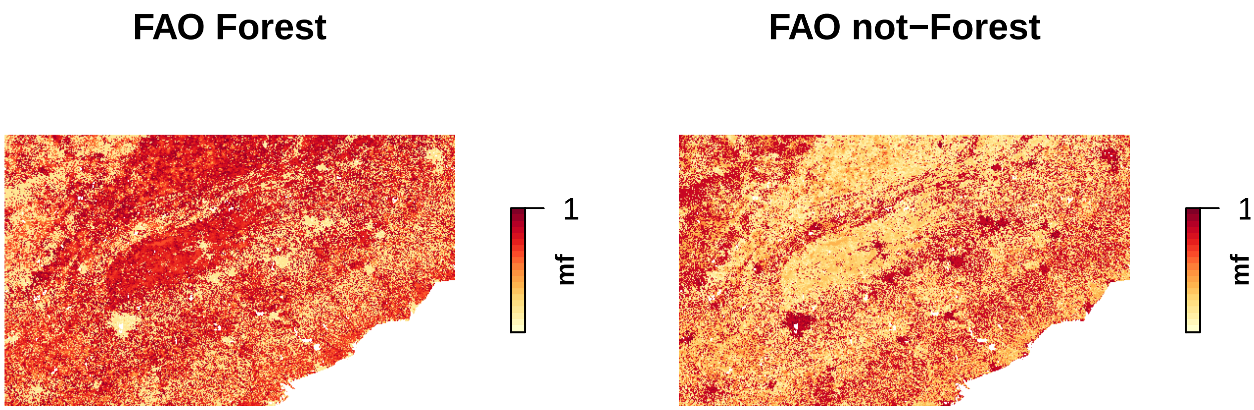

Figure 4: The fuzzy membership to the the FAO Forest and not-Forest classes in study area.

Table 10: The fuzzy difference matrix for the GLC (rows) and FAO (columns) classifications of forest.

	Forest	Not-Forest
Forest	42016.14	218.71
Not-Forest	21362.08	50417.08

5 shows that, broadly, the presence of forest is indicated in the same places (erroneously or not!) by both definitions of forest, with varying degrees of membership, suggesting different intensities of forest. The detail in Figure 6 illustrates the local variation in those membership intensities.

Table 9: The summaries of the Fuzzy memberships GLC and FAO derived Forest classes.

	GLC derived Forest	FAO derived Forest
Min.	0.04	0.02
1st Qu.	0.14	0.17
Median	0.36	0.68
3rd Qu.	0.57	0.79
Max.	0.91	0.98
Total	42234.84	63378.22

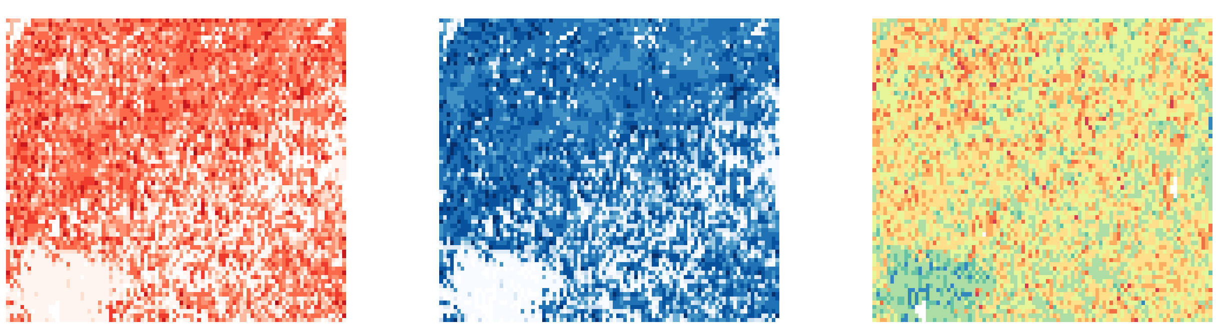

Figure 5: Detail around Atlanta, Georgia of the differences in the fuzzy memberships to the the GLC (red) and FAO (blue) forest classes, with a map of difference, with yellow indicates similar fuzzy membership values, red where GLC fuzzy forest memberships are greater and Blue where FAO are greater.

4.3 Fuzzy difference

The classic approach to examining the intersection between two fuzzy sets is to use fuzzy bounded difference, with explanatory fuzzy losses and gains (as in Table 6). The losses and gains are evident in the off diagonals

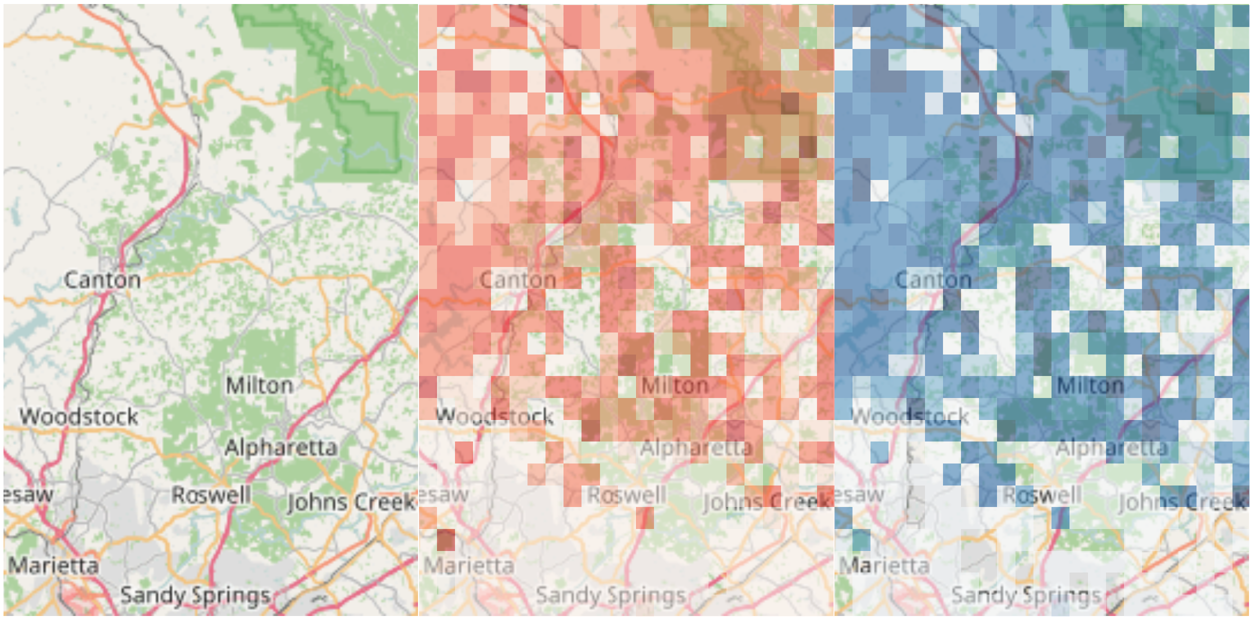

Figure 6: The boundary conditions between Forest and not-Forest arising from the differently defined forest classes, with context from an Open Street Map backdrop. The fuzzy GLC forest is mapped in red and the FAO fuzzy forest class is mapped in blue, both with a transparency term.

5. Discussuon and Conclusion

The REDD+ programme aims to reduce carbon emissions from deforestation and forest degradation. A key tenet of the programme is to financially reward less developed countries for their preservation (or increase) of forest stocks. However, any mapping, including land cover mapping, is only a representation of some view of %I added "some view of" reality (Comber et al., 2005) and different views of reality exist between and within different stakeholders and communities, whether they be citizens, scientists or involved in policy making (see Harvey and Chrisman (1998) for the classic example of the negotiation of the differences in land cover definitions).

The problems for REDD+ are two-fold. First, differences in definitions of forest influence the amount of forest, and therefore carbon storage, that is mapped. This undermines the ambitions of REDD+. For example, Romijn et al (2013) found that different definitions of forest resulted in different levels of Indonesian deforestation being identified. They critically noted that any definition of forests that is not adapted to national circumstances could lead to large areas of deforestation being excluded and not accounted for under the REDD+ scheme, confusing the carbon accounting. Other research has identified similar problems (Van Noordwijk and Minan 2009; Magdon 2014; Gou, 2016). To better understand the implications and origins of this confusion, some research has sought to unpick forest classifications and how they relate to different management objectives (Chazdon et al 2016). Although these authors do not refer to semantics, their study clearly shows how different forest semantics are driven by initiatives (such as the IPCC, biodiversity conventions, or agroforestry fora), noting that the presence of different management objectives 'provides a perspective from which specific definitions are created' (p.539). Second, REDD+ forest mapping activities are being used to impose a specific view of the landscape, often at the expense of communities who live and work in forested areas (eg Bong et al., 2016). These communities, although key stakeholders, are frequently excluded and marginalized from the REDD+ mapping process, receiving little or none of the financial rewards, and often suffering from opportunistic land reforms arising on the back of REDD+. There is plenty of evidence of this situation in the literature from all over the world (eg Gizachew et al 2017, Astuti and McGregor 2016; Larson et al, 2013; Sunderlin et al, 2014; Lyster 2011; Ezzine-de-Blas et al 2011; Duchelle et al, 2014)

In this paper we have shown how data primitives and and fuzzy intersections can be used to quantify the differences between stocks of forest when different forest classifications are applied. The data primitive approach defines each forest class using data of soil wetness, vegetation height, biomass, human disturbance and seasonality. These provide a common set of orthogonal data layers (GIS field layers) and are fully described in Comber (2008) and Wadsworth et al (2008). Fuzzy intersection methods (Fisher, 2006b) were used to quantify the overlaps between the classes, specifically fuzzy bounded difference.

Thus, the problem of divergence in classification, resulting from diverse conceptualizations of forest (as identified by many authors in the context of data integration (eg Bennet 2001, Comber et al 2005), can be overcome, not through standardization (see Comber et al, 2007) as argued by many in the policy arena and the remote sensing community, but by allowing *divergent* forest classifications and mappings. Their relationships can be handled through commonly used and accepted uncertainty handling methods. Crucially, such an approach to forest mapping does not rely on any boundaries, be they crisp or vague. It only requires to be able to characterize for each piece of land considered the values of the dimensions (data primitives) in the forestness vector field.

Acknowledgements

The authors would like to thank Ross Purves and his team for their invitation to take part in a workshop in 2016, which led to the further development of these ideas. Some of the ideas in this work were developed under a Joint Nature Conservancy Committee project *Fuzzy approaches for Developing and Evaluating Earth Observation-enabled ecological land cover time series system* (JNCC Reference: C12-0171-0589) and Lex Comber would like to thank the late and great Prof Pete Fisher for some of the ideas in this paper: these were lifted from work that he and Pete never managed to finish. Werner Kuhn acknowledges support from the University of California, Santa Barbara, for the Center of Spatial Studies. All of the analyses and maps in this paper were undertaken in R, the open source statistics software. The code and data used to construct this analysis as well as the `RMarkdown` file used to create this paper are freely available at https://github.com/lexcomber/ForestPaper - have a play!

References

Angelsen, A. (Ed.). (2009). Realising REDD+: National strategy and policy options. CIFOR.\

Astuti, R., & McGregor, A. (2016). Indigenous land claims or green grabs? Inclusions and exclusions within forest carbon politics in Indonesia. The Journal of Peasant Studies, 1-22.\

Bennett B, 2001, What is a forest? On the vagueness of certain geographic concepts, Topoi 20 , 189-201\

Beymer-Farris, B.A., Bassett, T.J., 2012. The REDD menace: resurgent protectionism in Tanzania's mangrove forests. Glob. Environ. Change 22 (2), 332–341.\

Bezdek, J.C., (1981). Pattern Recognition with Fuzzy Objective Function Algorithms. New York: Plenum Press.\

Bong, I. W., Felker, M. E., & Maryudi, A. (2016). How Are Local People Driving and Affected by Forest Cover Change? Opportunities for Local Participation in REDD+ Measurement, Reporting and Verification. PloS one, 11(11), e0145330.\

Burenhult, N. (2009). Landscape terms and toponyms in Jahai: a field report. Working Papers in Linguistics, 51, 17-29.\

Burrough P A, Frank A U (Eds), 1996 Geographic Objects with Indeterminate Boundaries (Taylor and Francis, London)

Cavanagh, C., Benjaminsen, T.A., 2014. Virtual nature, violent accumulation: the 'spectacular failure' of carbon offsetting at a Ugandan National Park. Geoforum 56, 55-65.\

Cavanagh, C. J., Vedeld, P. O., & Trædal, L. T. (2015). Securitizing REDD+? Problematizing the emerging illegal timber trade and forest carbon interface in East Africa. Geoforum, 60, 72-82.\

Chazdon, Robin L., Pedro HS Brancalion, Lars Laestadius, Aoife Bennett-Curry, Kathleen Buckingham, Chetan Kumar, Julian Moll-Rocek, Ima Célia Guimaraes Vieira, and Sarah Jane Wilson. "When is a forest a forest? Forest concepts and definitions in the era of forest and landscape restoration." Ambio 45, no. 5 (2016): 538-550.\

Comber, A.J., Fisher, P.F., Wadsworth, R.A., (2005). What is land cover? Environment and Planning B: Planning and Design, 32:199-209.\

Comber, A., (2008). The separation of land cover from land use with data primitives. Journal of Land use Science, 3(4): 215–229\

Comber, A.J., Fisher, P.F., Wadsworth, R.A., (2007). Land cover: to standardise or not to standardise? Comment on 'Evolving standards in land cover characterization' by Herold et al. Journal of Land Use Science, 2(4): 287–291\

Couclelis, H. (1992). People manipulate objects (but cultivate fields): Beyond the raster-vector debate in GIS. In A. U. Frank, I. Campari, U. Formentini (Eds.), Theories and methods of spatio-temporal reasoning in geographic space. Berlin: Springer-Verlag: 65–77\

Duchelle, Amy E., Marina Cromberg, Maria Fernanda Gebara, Raissa Guerra, Tadeu Melo, Anne Larson, Peter Cronkleton et al. "Linking forest tenure reform, environmental compliance, and incentives: lessons from REDD+ initiatives in the Brazilian Amazon." World Development 55 (2014): 53-67.\

Ece, M. (2015). Representation through Privatisation. Responsive Forest Governance Initiative, working paper 23 – available from https://sdep.earth.illinois.edu/files/RFGI_Working_Papers/RFGI%2023_Melis%20Ece.pdf\

Ezzine-de-Blas, D., Börner, J., Violato-Espada, A. L., Nascimento, N., & Piketty, M. G. (2011). Forest loss and management in land reform settlements: Implications for REDD governance in the Brazilian Amazon. Environmental Science & Policy, 14(2), 188-200. \

Fisher, P.F. (1997). The pixel: a snare and a delusion. International Journal of Remote Sensing 18 (3), 679-685. \

Fisher, P. F., and Pathirana, S. (1990). The evaluation of fuzzy membership of land cover classes in the suburban zone. Remote Sensing of the Environment, 34, 121-132. \

Fisher, P., Comber, A., and Wadsworth, R., (2006a).~ Approaches to uncertainty in spatial data.~ In Fundamentals of Spatial Data Quality, edited by R.Devillers and R.Jeansoulin, ISTE, London. pp43-59 \

Fisher, P., Arnot, C., Wadsworth, R., and Wellens, J. (2006b). Detecting change in vague interpretations of landscapes. Ecological Informatics, 1, 163-178. \

Fisher, P., Cheng, T., & Wood, J. (2007). Higher order vagueness in geographical information: empirical geographical population of type n fuzzy sets. Geoinformatica, 11(3), 311-330. \

Fisher, P. F. (2010). Remote sensing of land cover classes as type 2 fuzzy sets. Remote Sensing of Environment, 114(2), 309-321. \

Fisher, P.F., Comber, A.J., Wadsworth, R.A., (2005). Land use and Land cover: Contradiction or Complement. Pp. 85-98 in Re-Presenting GIS, (eds. Peter Fisher, David Unwin), Wiley, Chichester. \

Forest Peoples Programme, 2014. Annual Report. Available form http://www.forestpeoples.org/tags/annual-reports/publication/2015/annual-report-2014 \

Gizachew, B., Astrup, R., Vedeld, P., Zahabu, E. M., & Duguma, L. A. (2017, March). REDD+ in Africa: contexts and challenges. In Natural Resources Forum. Blackwell Publishing Ltd. \

Gou, Y. (2016). The Role of Maps in Capturing Distal Drivers of Deforestation and Degradation: A Case Study in Central Mozambique. In Land Use Competition (pp.~91-109). Springer International Publishing. \

Grainger, M., & Geary, K. (2011). The New Forests Company and its Uganda plantations. Washington DC, USA: OXFAM International.

Gupta, A., Vijge, M. J., Turnhout, E., & Pistorius, T. (2014). Making REDD+ transparent: The politics of measuring, reporting, and verification systems. Transparency in Global Environmental Governance: Critical Perspectives, 181. \

Harvey F, Chrisman N, 1998, "Boundary objects and the social construction of GIS technology" Environment and Planning A 30 1683 -{} 1694 \

Herold, M., Verchot, L., Angelsen, A., Maniatis, D., Bauch, S., 2012a. A Step-wise Framework for Setting REDD+ Forest Reference Emission Levels and Forest Reference Levels. CIFOR Infobrief, no. 52, April 2012. \

Herold, M., Angelsen, A., Verchot, L.V., Wijaya, A., Ainembabazi, J.H., 2012b. A stepwise framework for developing REDD+ reference levels. In: Angelsen, A., Brockhaus, M., Sunderlin, W.D., Sunderlin, L.V. (Eds.),

Analysing REDD+; Challenges and Choices. CIFOR, Bogor, Indonesia, pp.~279–299. \

Himmelfarb, D., 2012. In the Aftermath of Displacement: A Political Ecology of Dispossession, Transformation, and Conflict on Mt. Elgon, Uganda. PhD Dissertation. University of Georgia, Athens. \

Hoeschele W, 2000, "Geographic information engineering and social ground truth in Attappadi, Kerala State, India" Annals of the Association of American Geographers 90 293 -{} 321 \

Klir, G.J., Yuan, B., (1995). Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice Hall, Englewood Cliff. \

Kuhn, W. (2005). Geospatial semantics: why, of what, and how?. In Journal on data semantics III (pp.~1-24). Springer Berlin Heidelberg. \

Lakoff, G. (2008). Women, fire, and dangerous things. University of Chicago press. \

Larson, A.M., Brockhaus, M., Sunderlin, W.D., Duchelle, A., Babon, A., Dokken, T., Pham, T.T., Resosudarmo, I.A.P., Selaya, G., Awono, A. and Huynh, T.B., 2013. Land tenure and REDD+: The good, the bad and the ugly. Global Environmental Change, 23(3), pp.678-689. \

Leung, Y.C., (1988). Spatial Analysis and Planning under Imprecision. Elsevier, New York. \

Lund H. Gyde (2015). Definitions of Forest, Deforestation, Afforestation, and Reforestation. http://home.comcast.net/~gyde/DEFpaper.htm \

Lyons, K., Westoby, P., 2014. Carbon colonialism and the new land grab: plantation forestry in Uganda and its livelihood impacts. J. Rural Stud. 36, 13–21. \

Lyster, R. (2011). REDD+, transparency, participation and resource rights: the role of law. Environmental science & policy, 14(2), 118-126. \

Magdon, P., Kleinn, C., 2012. Uncertainties of forest area estimates caused by the minimum crown cover criterion - a scale issue relevant to forest cover monitoring. Environmental Monitoring and Assessment, 1–16. \

Magdon, P., Fischer, C., Fuchs, H., & Kleinn, C. (2014). Translating criteria of international forest definitions into remote sensing image analysis. Remote Sensing of Environment, 149, 252-262. \

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., eisch, F., 2012. e1071: Misc Functions of the Department of Statistics (e1071), TU Wien, 2012. R package version, pp.~1–6. \

Morales-Barquero, L., Skutsch, M., Jardel-Peláez, E. J., Ghilardi, A., Kleinn, C., & Healey, J. R. (2014). Operationalizing the definition of forest degradation for REDD+, with application to Mexico. Forests, 5(7), 1653-1681. \

Nel, A., Hill, D., 2013. Constructing walls of carbon–the complexities of community, carbon sequestration and protected areas in Uganda. J. Contemp. Afr. Stud. 31 (3), 421–440. \

Robbins P, 2001, “Fixed categories in a portable landscape: the causes and consequences of land-cover categorization” Environment and Planning A 33 161 -{} 179 \

Romijn, E., Ainembabazi, J. H., Wijaya, A., Herold, M., Angelsen, A., Verchot, L., & Murdiyarso, D. (2013). Exploring different forest definitions and their impact on developing REDD+ reference emission levels: A case study for Indonesia. Environmental Science & Policy, 33, 246-259. \

Ryan, C. M., Berry, N. J., & Joshi, N. (2014). Quantifying the causes of deforestation and degradation and creating transparent REDD+ baselines: a method and case study from central Mozambique. Applied Geography, 53, 45-54. \

Smith B, Mark D, 2001, “Geographical categories: an ontological investigation” International Journal of Geographical Information Science 15 591 -{} 612 \

Smith B, 1995, “On drawing lines on a map”, in Lecture Notes in Computer Science 988 Spatial Information Theory Eds A U Frank, W Kuhn (Springer, Berlin) pp 475 -{} 484 \

Sunderlin, W. D., Larson, A. M., Duchelle, A. E., Resosudarmo, I. A. P., Huynh, T. B., Awono, A., & Dokken, T. (2014). How are REDD+ proponents addressing tenure problems? Evidence from Brazil, Cameroon, Tanzania, Indonesia, and Vietnam. World Development, 55, 37-52. \

UNFCCC, 2009. Methodological guidance for activities relating to reducing emissions from deforestation and forest degradation and the role of conservation, sustainable management of forests and enhancement of forest carbon stocks in developing countries. Decision COP 15/4. 2009. \

Van Noordwijk, Meine, and Peter Minang. If we cannot define it, we cannot save it: forest definitions and REDD. ASB PolicyBrief 15 (2009).

Wadsworth, R., Balzter, H., Gerard, F., George, C., Comber, A. and Fisher, P. (2008). An environmental assessment of land cover and land use change in Central Siberia using Quantified Conceptual Overlaps to reconcile inconsistent data sets. Journal of Land use Science, 3(4): 251–264. \

Wang, F. (1990). Fuzzy supervised classification of remote sensing images. Geoscience and Remote Sensing, IEEE Transactions on, 28(2), 194-201.\

Zadeh, L.A., (1965). Fuzzy sets. Information and Control 8, 338–353.\