

Using Volunteered Data in Land Cover Map Validation: Mapping West African Forests

Giles M. Foody, *Fellow, IEEE*, and Doreen S. Boyd

Abstract—Accuracy assessment should be a fundamental part of a programme that maps land cover from remotely sensed imagery but this activity is often constrained by the lack of high quality ground reference data. Here, two sources of volunteered data are used to illustrate the potential of amateur or neogeographical activity in map validation. Ground based photographs acquired from an internet-based collaborative project and interpreted by a set of four further volunteers provided the reference data to support evaluation of the Globcover map's representation of tropical forest in West Africa. Although the results highlight some concerns with volunteered data, notably the low levels of inter-volunteer agreement they also show that such imperfect data may be used to derive credible estimates on accuracy from both a site and non-site specific perspective. Specifically, the estimates of the producer's accuracy of forest and of forest extent derived from the free and volunteered data using a latent class model were of comparable magnitude to those derived in a formal validation by experts; the estimate of forest extent was within 1.38–9.08% of reference estimates while the difference in estimated producer's accuracy from that derived in an authoritative assessment of map accuracy was 2.82% and 0.34% for the forest and non-forest classes respectively.

Index Terms—Accuracy, classification, land cover, volunteered geographic information.

I. INTRODUCTION

LAND cover is one of the most critical environmental variables. It has, for example, numerous direct and indirect effects on environmental properties and processes that strongly impact on human health and well-being. Forests, for example, dominate the terrestrial carbon cycle and form the world's largest bank of species diversity making information on forest cover critical to the two greatest societal concerns of the day: climate change and the conservation of biodiversity. Changes in forest cover are, therefore, critical to scientific studies of carbon cycling and species conservation. There is, therefore, considerable need for accurate and timely information on land cover classes such as forest and the only practical means to derive this information over large areas and on a frequently repeatable basis is via satellite remote sensing.

Land cover mapping is one of the most common applications of satellite remote sensing. Considerable research has addressed a wide variety of issues connected with this application, from

sensor design through image-preprocessing to map production and evaluation. The latter issue is of particular importance, especially if remote sensing is to provide land cover information to support to major scientific and policy applications such as contributing to the United Nations collaborative initiative on Reducing Emissions from Deforestation and forest Degradation (REDD). Accuracy assessment is typically undertaken on a non-site or, more commonly, a site specific basis and has evolved greatly over the last four decades [1]. It is now widely accepted that every map should be accompanied by a rigorously derived accuracy statement, otherwise it is no more than one untested hypothesized representation [2], perhaps being little more than a pretty picture [3].

Although guidelines exist that specify good practices for accuracy assessment in remote sensing [2] the methods are often not followed [4], [5]. Indeed, because of the many challenges encountered in a mapping programme [6]–[10] the best practices are sometimes impractical to implement. The validation of a map, in which its accuracy is assessed, is a distinctly non-trivial task [6], [7]. A key concern is that it is often extremely difficult to acquire a suitable ground reference data set upon which to base an accuracy assessment. Concerns about ground reference data abound especially in relation to the quantity and quality of the reference data used to evaluate the map [11]–[14]. Thus while the main recommended approach to accuracy assessment [2] requires the collection of a potentially large sample of high quality ground reference data at sites selected in accordance to carefully designed probability sampling design the implementation of such approach is often extremely difficult. A range of research agendas have been defined to meet the challenges faced in accuracy assessment with activity directed at issues ranging from the provision of data sets that may aid an accuracy assessment [15] to the evaluation of methods for accuracy assessment without reference data [14], [16]. A further alternative explored here is to exploit the potential of reference data provided by volunteers, often amateurs or neogeographers collecting information as part of collaborative projects sometimes hosted on the internet.

The recent rise in citizen science and, in particular, citizen sensing [17], [18] offers an attractive, if complicated, alternative means to ground reference data collection to support map validation activities. The potential of volunteered data in accuracy assessment has been explored [19], [20] but many issues remain to be fully addressed. There are many concerns with volunteered data sets, notably their variable, and typically unknown, quality together with a set of ethical and legal concerns associated with its collection and use [21], [22]. Here, the aim is to further illustrate the potential of volunteered data in a way that explicitly

Manuscript received September 28, 2012; revised January 16, 2013; accepted February 12, 2013. Date of publication April 03, 2013; date of current version June 17, 2013.

The authors are with the School of Geography, University of Nottingham, Nottingham, NG7 2RD, U.K. (e-mail: giles.foody@nottingham.ac.uk; doreen.boyd@nottingham.ac.uk).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSTARS.2013.2250257

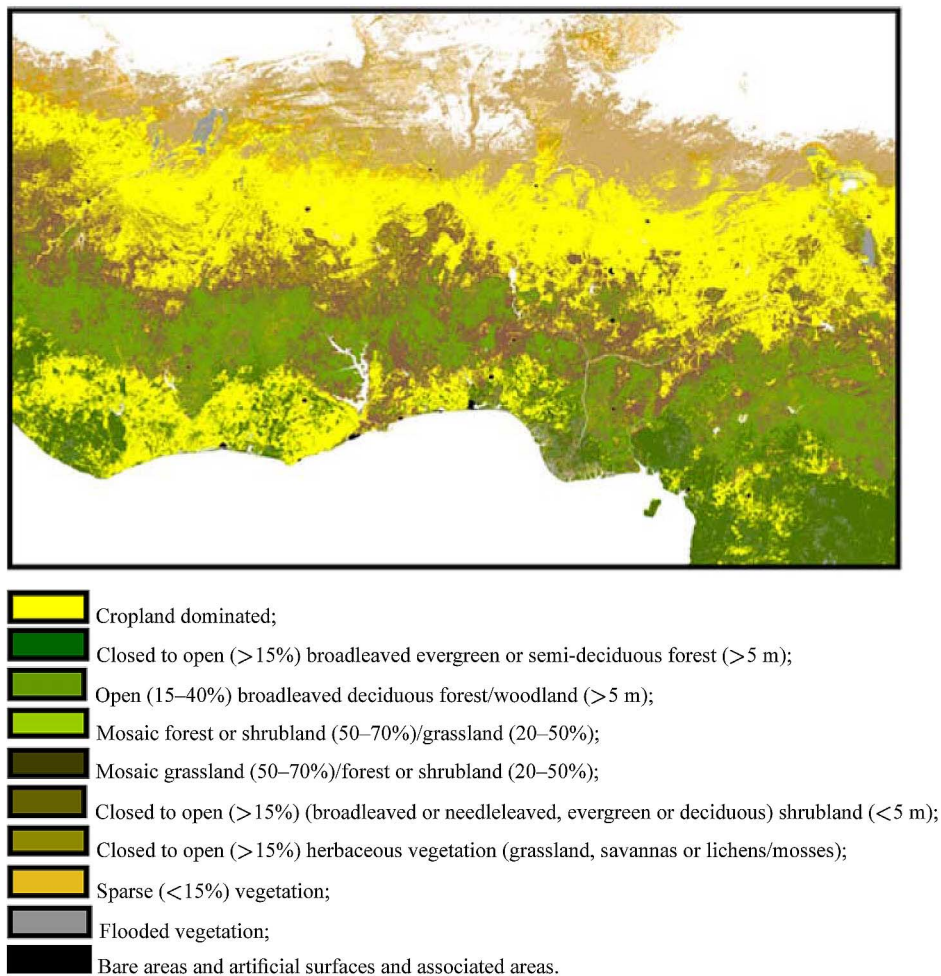


Fig. 1. Extract of the Globcover map for the study area. The classes have been adjusted for illustrative purposes.

recognizes its typically unknown and imperfect nature. Specifically, this article provides a brief summary of research into the validation of the forest cover representation provided by the European Space Agency's Globcover map using two sources of volunteered data. Section II outlines the data used in this research. Section III provides a summary of the methods used and especially latent class modeling. Section IV provides the results of the analyses and a discussion before the main conclusions are drawn in Section V.

II. DATA

Attention focused on forests in West Africa and their representation in the Globcover map (Fig. 1); the map is available from <http://dup.esrin.esa.int/globcover>. The Globcover map was selected as it provides a contemporary and relatively fine spatial resolution (300 m) representation of land cover. The map was produced using data from the MERIS sensor acquired over the period December 2004 to June 2006. The map has also been validated and so there is a guide to the accuracy of the map. The latter is critical in that it provides targets for the research to achieve.

Although the Globcover map has a relatively high thematic resolution the focus here was on a single class in order to reduce uncertainties, especially those associated with associated with

class definitions. Here, the focus was upon only forests in West Africa. To reduce uncertainties about different forest types, the forest classes that occurred in the region were aggregated into a single class. There are many definitions of what constitutes a forest [23], [24] and here a site was considered to be forested if it had at least 15% canopy cover. The forest class was formed by aggregating a suite of mosaic, closed and open forest classes contained in the original Globcover representation of the region.

A rigorous validation of the Globcover map informed by published documentation on good practices has been undertaken [25]. This validation involved 16 experts, 3 focusing on Africa, who labeled 3167 points selected from the map. The Globcover map was estimated to have an overall accuracy of $\sim 67.1\%$ rising to $\sim 79.2\%$ if expressed on a class area-weighted basis and using only the ground reference data with a high confidence in the labeling [25]. The classes aggregated to form the forest class of interest in this work may also be aggregated in the confusion matrix used in the estimation of Globcover map accuracy. Following this aggregation of classes, the producer's accuracy for forest and non-forest classes was estimated as 84.04 and 74.89% respectively. These values provided benchmarks for use in the evaluations of the estimates derived using the volunteered data.

Inspection of the confusion matrix used in the formal validation of the Globcover map [25] indicated that the omission

and commission errors for the forest class were asymmetric resulting in the extent of forest being overestimated by a factor of ~ 1.42 . A variety of methods to correct for misclassification bias are available [5], [26], [27]. Here, the ratio of the commission to omission errors was used to rescale downwards the estimated extent in order to adjust for the asymmetry in the observed classification errors. It was, however, apparent that some of the inter-class confusion evident in the matrix derived over the globe involved classes that did not occur in the region of study. Additionally, it is well known that map accuracy can vary spatially [28] with large regional variation observed in global maps [29], suggesting that the use of more local information is often desired. An evaluation of the Globcover map for the continent of Africa suggests that omission and commission are still imbalanced across the region [30] but less so than indicated by the formal validation of the entire map. The latter suggests that the extent of forest may be overestimated by a factor of ~ 1.10 . Critically, the extent of forest depicted in Globcover appears to be exaggerated by a factor of ~ 1.10 to ~ 1.42 . These latter values may be used to rescale or revise downward the estimated extent of forest.

The key focus of this article is on the potential of using volunteered data as a means to validate the map and derive useful information without conventional ground reference data. With the formal validation providing authoritative estimates of map accuracy and forest extent, the volunteered data were used to provide a guide to the accuracy of the forest representation provided by Globcover. This analysis used the data acquired by volunteers at 99 locations across West Africa.

Two sources of volunteered data were used to generate the ground reference data set. First, freely available photographic data on ground conditions at the point of intersection of lines of latitude and longitude provided by the Degrees of Confluence project (www.confluence.org), a web-based collaborative project, were used as a source of spatially extensive information on ground surface cover (Fig. 2). An aim within the project is for each confluence point to be visited by a volunteer and a set of photographs of the site obtained; typically the data set provided for each location contained four photographs viewing north, south, east and west from the confluence point. Additional data are often provided, including evidence of the photographs having been acquired at the correct location as well as text that could also reveal useful information about the site. Here, attention focused solely on the set of photographs acquired at or near the point of confluence (Fig. 3). The photographs for all 99 successfully visited and documented confluence points available at the time of the research (winter 2011) for the Ivory Coast, Ghana, Togo, Benin and Nigeria were used (Fig. 2). While the set of field visits from which the ground photographs were acquired were undertaken over a long period, the potential of forest seasonality effects and the time gap of up to seven years between map production and field visits are concerns. However, given the focus on low latitude forests and with the annual rate of forest area change being $< 0.5\%$ for Western and Central Africa in the period 2000–2010 [31] the potential for major error is reduced. A second set of data was derived from volunteers at Nottingham University who labelled each confluence point as forest or non-forest based on the set of photographs

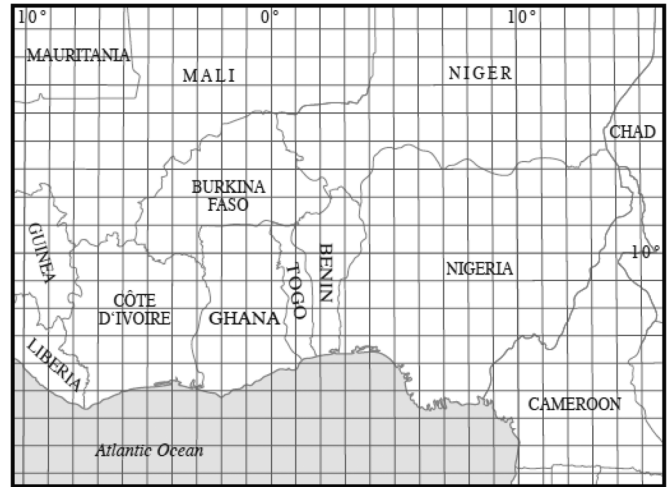


Fig. 2. Map of West Africa showing the 1° lines of latitude and longitude.

available for it. Each site was labelled as forest (1) or non-forest (0) independently by a total of four people (A-D). Thus for 99 systematically distributed points across West Africa, volunteered data on forest cover were available to compare against the labeling depicted in the Globcover map.

Aside from the use of land cover class labels generated by volunteers rather than experts, the fundamental nature of the sample used conforms to recommended practice in accuracy assessment. The systematic sample design, for example, is an equal probability design that is easy to implement and one of the recommended best-practice methods for use in relation to mapping large areas [2], [32]. The size of the sample was also anticipated to be suitable for credible accuracy assessment purposes. The required sample size can be calculated from sampling theory and is not a function of the size of the data set but of the desired precision in estimation and selected level of statistical confidence [33]. While the latter are a function of the specific objectives of a study, the sample size used, 99, is large enough to allow estimation with a margin of error $< 6.0\%$ assuming that the popular target accuracy of 85% producer's accuracy can be used as a prior estimate of the accuracy and working at the 90% level of confidence.

III. METHODS

For each of the 99 points of confluence, a set of four labels indicating forest presence (1) or absence (0) was available from the volunteers (A-D). These labels supplemented the set extracted from the Globcover map (E). The labels derived by each volunteer were compared against those from each other volunteer and in the map to provide a guide to the degree of agreement between the volunteer derived reference data and the depiction in the Globcover map. The level of agreement between pairs of labels was expressed as a percentage cases agreeing in label and also by the kappa coefficient of agreement [34]. The latter was calculated from,

$$\hat{\kappa} = \frac{p_o - p_c}{1 - p_c} \quad (1)$$

where p_o is the observed proportion of agreement and p_c the proportion of agreement expected by chance. Although the

kappa coefficient is unsuitable as a measure of accuracy in remote sensing [35], [36] it is used here primarily as an index of the level of inter-rater agreement and not of accuracy.

The key focus of this paper is on using the volunteered data to validate the Globcover map, including the estimation of forest extent (a non-site specific measure of map accuracy). Here, it is recognised explicitly that all five representations are imperfect; some of the volunteers, for example, had minimal relevant experience and hence error might be expected. It was shown in [16] that by casting the problem of estimation from imperfect data in terms of a latent class analysis a set of imperfect variables may be used to derive information on the accuracy of land cover maps. Here, the set of four labels derived from all volunteers were combined with those depicted in the map to yield information on forest cover and the accuracy of its representation. A key focus was on the estimation of forest extent, expressed as a percentage of the region covered by forest, and the accuracy of the map, expressed here by the producer's accuracy of the forest and non-forest classes.

The quality of imperfect data may be estimated by a latent class analysis [37], [38]. In a latent class analysis it is assumed that each of the input variables is an imperfect indicator of the unobserved, and so latent, variable of interest but that the observed associations among them can be explained by a latent variable [37]. Here, the latent variable is the forest class, F , while the observed or manifest data of the analysis are the sets of class labels derived from the four volunteers and the Globcover map. With the labels from sources A, B, C, D and E being represented by $a, b, c, d, e = 0, 1$ a latent class model for the probability of obtaining a pattern of class labels over the sources may be written as

$$\pi_{abcdef} = \pi_f^F \pi_{abcde|F}^{ABCDE} \quad (2)$$

where, based on the assumption of conditional independence,

$$\pi_{abcde|F}^{ABCDE} = \pi_{a|f}^{A|F} \pi_{b|f}^{B|f} \pi_{c|f}^{C|F} \pi_{d|f}^{D|F} \pi_{e|f}^{E|F} \quad (3)$$

is the conditional probability that the pattern of class labels is (a, b, c, d, e) given that the case has a forest class status f (1 or 0) and π_f^F is the probability that a case has the forest class status f [37], [38]. Assuming that the model fits the data and reflects the class information as planned, the two sets of parameters of the latent class model equate, therefore, to key measures of classification accuracy used in remote sensing [16]. Attention here focuses on π_1^F as an estimate of forest extent and $\pi_{11}^{E|F}$ and $\pi_{00}^{E|F}$ which represent respectively the producer's accuracy for the forest and non-forest classes in the Globcover map. The latter measures provide a detailed description of the accuracy of the binary classification and are the typical measures of interest in mapping studies as well as benefit from a theoretical independence on class abundances [12]. The fit of a latent class model to the observed data is commonly evaluated with the likelihood ratio chi-squared statistic, L^2 ; with a model typically viewed as fitting the data if the value of L^2 is sufficiently small to be attributable to the effect of chance [39].

The basic model defined by (2) and (3) may provide a poor fit to the data if the assumption of conditional independence that underlies it is not satisfied. However, the model in (1) can

be adapted to allow for conditional dependence between all or some of the manifest variables. This is useful since it might be expected that the volunteers would all tend to correctly label the cases that were clearly non-forest (e.g. location 5°N 3°W in Fig. 3) and struggle most with the labeling of points that had some tree cover, probably close to the 15% tree cover threshold value used in the definition of forest. A test for conditional dependence was undertaken using a modified version of the log-odds ratio check method using the CONDEP programme (acquired from <http://www.john-uebersax.com/stat/condep.html>). With this method, the log-odds ratio for the observed and expected data are compared and a z score calculated. The calculated value of z indicates the extent of conditional dependence. Large values of z suggest conditional dependence occurs and the values may be interpreted against standard tabulated critical values for specified levels of statistical significance if desired.

IV. RESULTS AND DISCUSSION

Of the 99 locations, 50 were labelled as belonging to forest in the Globcover map. This indicates 50.50% of the region is covered with the class defined here as forest. However, the validations of the map reported in the literature [25], [30] indicate that this is likely to be an overestimate; this may in part be related to the breadth of the class definition used. Re-scaling the estimate of forest extent for the observed overestimation factor derived with the global confusion matrix in [25] suggests that the extent of forest cover is actually 35.36% while using the rescaling factor derived from the accuracy assessment for Africa reported by [30] suggests that the extent of forest is 45.82%.

There was only a relatively low degree of agreement between the volunteers in terms of their labeling of the locations from the photographs (Table I). In total, the volunteers agreed unanimously on the labeling of 48 locations (19 forest and 29 non-forest) with disagreement evident for the remaining 51 locations (Fig. 4). Pairwise comparison of the set of labels generated for the 99 locations showed that the four volunteers varied greatly in terms of labelling. The degree of agreement between pairs of labellers varied between 62.66 and 79.79% or, in terms of the kappa coefficient of agreement from 0.282 to 0.553 (Table I). These relatively low levels of agreement lead to an initial assessment that highlights some concerns for the use of volunteered data in accuracy assessment. Critically, while interpretation is limited by the absence of a gold-standard reference, the low levels of agreement may suggest that the individual volunteers differ greatly in terms of their perception of forest and so might be taken to suggest that such data have little useful role in map validation, especially as there is no obvious means to select between the volunteers.

The limitations of the volunteer data are also evident when compared against the Globcover map. Again, only low levels of agreement between volunteered and mapped labels were observed, with agreement varying from 58.58% to 77.77% or 0.170 to 0.556 in terms of the kappa coefficient of agreement (Table I). These results suggest only relatively low degrees of agreement exist in the labelling provided by the different sources of data on forest cover. The results may suggest that none of the individual volunteers was able to provide data that could be confidently used in accuracy assessment. There

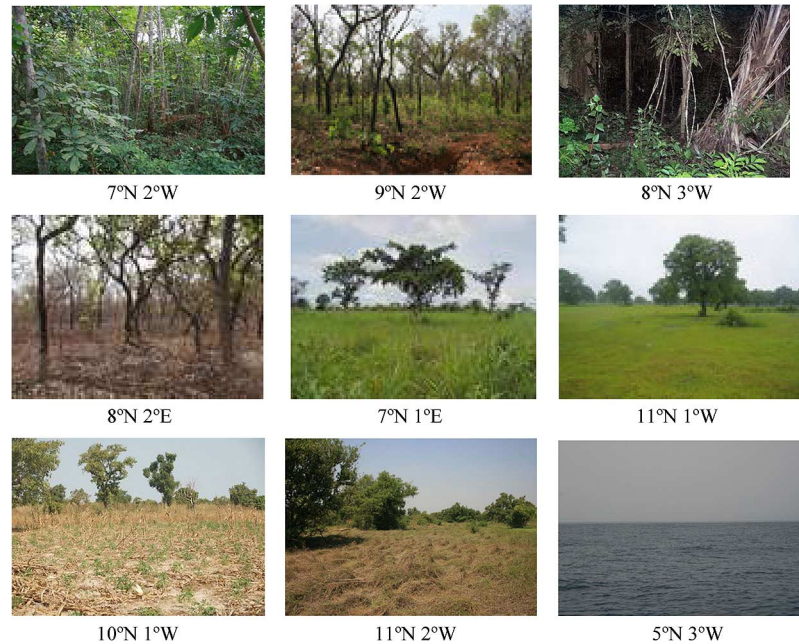


Fig. 3. Examples of photographs acquired at or near confluence points; the size/shape of some photographs have been adjusted for presentational reasons. The photographs were acquired from the Degrees of Confluence Project, contributed by N. Bieger, A. Kovacs, R. Mautz, H. Resch, D. Wood and reproduced with their permission).

TABLE I
KAPPA COEFFICIENTS INDICATING DEGREE OF INTER-RATER AGREEMENT

	B	C	D	E
A	0.367	0.495	0.553	0.235
B		0.424	0.282	0.170
C			0.458	0.556
D				0.154

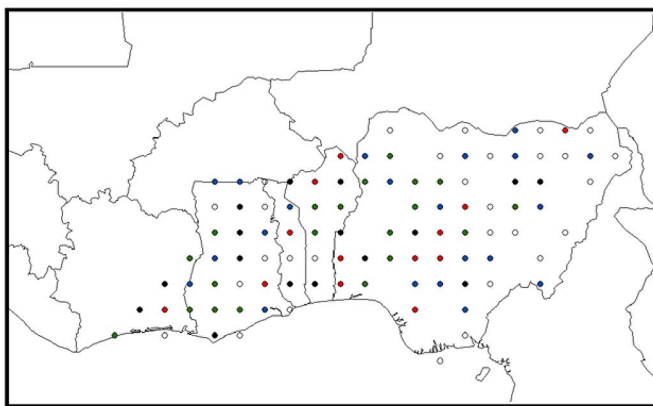


Fig. 4. Summary of the degree of agreement between the volunteers for each point of confluence; gaps occur for sites not visited. Open circle—all agree non-forest; green circle—all agree forest; blue, black and red locations that 1, 2 and 3 volunteers respectively labeled as forest.

are, of course, many potential sources of uncertainty and error in the analysis including problems with the photographic reference data, its labelling, the time period between the mapping and photography as well as error in the Globcover map. None-the-less, the low levels of agreement depicted in Table I do not initially suggest a positive role for any of the volunteers in a validation programme.

The estimated extent of forest over the test site derived from the volunteered data ranged from 32.33% to 57.57%. The wide range in the estimates is also apparent when viewed relative to the anticipated actual range of 35.36%–45.82%. These results further highlight some concerns with the use of volunteered data and suggest little confidence in using directly the volunteered data from any one source in map validation.

A variety of approaches might be used to derive an enhanced estimate of forest extent and map accuracy. One simple approach is to derive the average of the estimates from the 4 volunteer labellers. This approach yielded an estimate of 42.67% forest cover. Similarly, it would be possible to combine the individual classifications in a basic ensemble method. For example, by allocating each point to the class associated with the most frequently allocated label, and making a random class allocation for tied cases, yielded an estimate of 40.40% forest cover. Both of these estimates lie within the 35.36%–45.82% range anticipated and show that the set of imperfect labels derived by the multiple volunteers can be used to derive a credible estimate of forest extent. Here, however, attention is focused especially on the potential of the latent class modelling approach.

Using the set of four class labels for each point derived by the volunteers (A-D) and the map label (E), a latent class model based on (1) and (2) yielded a model that provided a good fit to the observed data; $L^2 = 22.73$ ($p = 0.20$). However, as volunteers might be expected to find some of the labelling tasks equally challenging with some locations obviously forest (or non-forest) while a set of others with some evidence of tree cover more problematic (see Fig. 3) it may be expected that the conditional independence assumption does not hold. The levels of inter-rater agreement observed (Table I) also suggest that some degree of dependence may exist between some of the

TABLE II
RESULTS OF TEST OF CONDITIONAL INDEPENDENCE ASSUMPTION

Variables compared	Expected log odds	Observed log odds	z
A - B	1.03	2.22	2.52
A - C	2.47	2.47	-0.00
A - D	1.04	2.62	3.55
A - E	1.29	1.12	-0.36
B - C	1.99	1.99	-0.00
B - D	0.89	1.41	1.17
B - E	1.02	0.70	-0.74
C - D	2.09	2.09	0.00
C - E	2.58	2.58	-0.00
D - E	1.12	0.68	-1.01

sources. A test of assumption of conditional independence indicated that for some of the variables a degree of conditional dependence existed (Table II). Based on Table II the latent class model was adjusted such that:

$$\pi_{abcdef}^{ABCDE|F} = \pi_{abdf}^{ABD|F} \pi_{cf}^{C|F} \pi_{ef}^{E|F} \quad (4)$$

In (4) the three variables indicated as being conditionally dependent (A, B and D) are those with the largest values of the z score in Table II. The model defined by (1) with (4) was applied to the data. This model appeared to fit the data slightly more closely than the original model that was based on the assumption of conditional independence; $L^2 = 14.20$ ($p = 0.28$). The parameters of this model were used to derive estimates of the accuracy of the Globcover map from the non-site and site specific perspectives that are widely used in the validation of maps derived by remote sensing.

From a non-site specific perspective attention is focused on the estimate of forest extent and so the magnitude of the π_f^F of the fitted model. The model based on (1) and (4) gave an estimate of 44.44% forest cover. Rigorous evaluation of this estimate is difficult in the absence of a true gold-standard reference data but it is very close to the value derived from Globcover map, especially after allowance for the known misclassification errors in the map. Indeed the model based estimate lay between the two values derived, after rescaling for misclassification errors, from conventional accuracy assessments: 35.36% and 45.82%. The latent class model, therefore, appears able to provide a credible estimate of the extent of forest cover in the region.

From a site specific perspective, the conditional probabilities of the latent class model provided estimates of the producer's accuracy for the forest and non-forest classes. The estimates for the forest and non-forest class were 81.22% and 74.55% respectively. Both are very close to the estimates derived from the confusion matrix produced in its formal validation [25] which were 84.04% and 74.89% for forest and non-forest respectively.

The results indicate that imperfect data arising from volunteered sources have the potential to provide useful information on map properties and accuracy. Specifically, the extent of forest (a non-site specific measure of accuracy) was estimated to within 1.38% and 9.08% of the rescaled estimates derived from the validation analysis provided by [30] and [25] respectively. The producers' accuracy of the forest and non-forest classes

were also estimated to within 2.82% and 0.34% of the values derived from the formal validation. It should also be noted that the volunteered data have other advantages. For example, the systematic sample design used ensures that data are derived over a widespread area and could be used to aid map production (e.g. by provision of prior information on class abundance) and the nature of the design is suitable for many map accuracy and comparison objectives [32]. The latent class model parameters could also be used to indicate the quality of the labels derived from the different volunteers.

There are, however, many issues to explore with the use of latent class analysis in remote sensing applications. Further work should seek to explore issues connected with the number of volunteers and their level of expertise as well as concerns linked to incomplete and uncertain labeling, especially for multi-class classifications. These are important issues in relation to the work reported as only a small number of volunteers were used and the uncertainties with the data (e.g. due to the time gap between map production and field visits and common problems such as those associated with the definition of the forest class [40]). The potential to also steer volunteer activity, forming active citizen sensors, should also be explored. It should also be stressed that while credible estimates have been derived the method should be used with caution and is, not yet, viewed as an alternative to standard good practice methods. Many of these issues are the topic of a recently launched European Union funded COST Action on 'Mapping and the Citizen Sensor' (TD1202—readers interested in participating or following its activities are encouraged to see www.cost.eu/domains_actions/ict/Actions/TD1202 for further details).

V. CONCLUSIONS

Volunteers have considerable potential to contribute constructively to land cover mapping programmes. This potential has grown rapidly over recent years, fostered by recent advances in geoinformation technologies and now offers great opportunities to enhance studies of the Earth's environment if it can be carefully used. Although the quality of volunteered data is often a concern it was shown that realistic estimates of forest cover and of map accuracy could be derived from easy to acquire and inexpensive, if not free, volunteered information.

This paper has confirmed the potential value of internet-based collaborative activities such as the Degrees of Confluence project for the provision of useful, spatially extensive, data to aid map evaluation. Moreover, it has shown that additional data from multiple interpreters can provide class labels for the data that can be used for validation purposes. Although these multiple classifications may be of unknown and uncertain quality they may be used together to aid the derivation of credible estimates of map accuracy. A latent class analysis may be applied to the sets of imperfect data to derive estimates of map accuracy. Although there were only relatively low levels of inter-rater agreement and each volunteer's labels showed poor agreement with the Globcover map, it was evident that the estimate of forest cover derived from the latent class analysis was close to that depicted in the map and accuracy estimates close to those derived from authoritative methods.

ACKNOWLEDGMENT

We gratefully acknowledge the help and support provided by the volunteers that contributed to this work which extends that reported in a paper presented at IGARSS 2012 in Munich and the referees for their comments on the article. In particular, we are grateful to the volunteers: contributors to the Degrees of Confluence project and students at Nottingham University. Permission to use the photographs reproduced in Fig. 3 was gratefully received from N. Bieger, A. Kovacs, R. Mautz, H. Resch, D. Wood. Thanks are also due to the creators and distributors of the software used: LEM for the latent class analyses and CONDEP for the test of conditional dependence. We are also grateful to ESA and the ESA GlobCover Project, led by MEDIAS-France, for GlobCover products. The sample locations are defined in the paper and the data sets used available from cited sources. The preparation of this article gained in part from support via an EPSRC grant (reference EP/J0020230/1) and the development and launch of COST Action TD1202 (Mapping and the citizen sensor) which focuses on aspects of volunteered geographic information for mapping applications.

REFERENCES

- [1] R. G. Congalton and K. Green, *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, 2nd ed. Boca Raton, FL, USA: Lewis Publishers, 2009.
- [2] A. H. Strahler, L. Boschetti, G. M. Foody, M. A. Friedl, M. C. Hansen, M. Herold, P. Mayaux, J. T. Morisette, S. V. Stehman, and C. E. Woodcock, "Global Land Cover Validation: Recommendations for Evaluation and Accuracy Assessment of Global Land Cover Maps," Joint Research Centre, Ispra, EUR 22156 EN Technical Report, 2006, 48 pp.
- [3] R. E. McRoberts, "Satellite image-based maps: Scientific inference or pretty pictures?," *Remote Sens. Environ.*, vol. 115, pp. 715–724, 2011.
- [4] N. M. Trodd, "Uncertainty in land cover mapping for modelling land cover change," in *Proc. Remote Sensing in Action, RSS'95*, Nottingham, U.K., 1995, pp. 1138–1145.
- [5] P. Olofsson, G. M. Foody, S. V. Stehman, and C. E. Woodcock, "Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation," *Remote Sens. Environ.*, vol. 129, pp. 122–131, 2013.
- [6] G. M. Foody, "Status of land cover classification accuracy assessment," *Remote Sens. Environ.*, vol. 80, pp. 185–201, 2002.
- [7] G. M. Foody, "Harshness in image classification accuracy assessment," *Int. J. Remote Sens.*, vol. 29, pp. 3137–3158, 2008.
- [8] J. Cihlar, "Land cover mapping of large areas from satellites: Status and research priorities," *Int. J. Remote Sens.*, vol. 21, pp. 1093–1114, 2000.
- [9] F. J. Gallego, "Remote sensing and land cover area estimation," *Int. J. Remote Sens.*, vol. 25, pp. 3019–3047, 2004.
- [10] R. R. Rindfuss, S. J. Walsh, B. L. Turner, J. Fox, and V. Mishra, "Developing a science of land change: Challenges and methodological issues," in *Proc. National Academy of Sciences USA*, 2004, vol. 101, pp. 13976–13981.
- [11] M. J. Carlotto, "Effect of errors in ground truth on classification accuracy," *Int. J. Remote Sens.*, vol. 30, pp. 4831–4849, 2009.
- [12] G. M. Foody, "Assessing the accuracy of land cover change with imperfect ground reference data," *Remote Sens. Environ.*, vol. 114, pp. 2271–2285, 2010.
- [13] R. G. Pontius and S. H. Petrova, "Assessing a predictive model of land change using uncertain data," *Environmental Modelling and Software*, vol. 25, pp. 299–309, 2010.
- [14] A. Baraldi, L. Bruzzone, and P. Blonda, "Quality assessment of classification and cluster maps without ground truth knowledge," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, pp. 857–873, 2005.
- [15] P. Olofsson, S. V. Stehman, C. E. Woodcock, D. Sulla-Menashe, A. M. Sibley, J. D. Newell, J. D. , M. A. Friedl, and M. Herold, "A global land-cover validation data set, part I: Fundamental design principles," *Int. J. Remote Sens.*, vol. 33, pp. 5768–5788, 2012.
- [16] G. M. Foody, "Latent class modelling for site and non-site specific classification accuracy assessment without ground data," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, pp. 2827–2838, 2012.
- [17] M. F. Goodchild, "Citizens as sensors: The world of volunteered geography," *GeoJournal*, vol. 69, pp. 211–221, 2007.
- [18] C. E. Jones, N. J. Mount, and P. Webber, "The rise of the GIS volunteer," *Trans. GIS*, vol. 16, pp. 431–434, 2012.
- [19] S. Fritz, I. McCallum, C. Schill, C. Perger, L. See, D. Schepaschenko, M. van der Velde, F. Kraxner, and M. Obersteiner, "Geo-Wiki: An online platform for improving global land cover," *Environmental Modelling and Software*, vol. 31, pp. 110–123, 2012.
- [20] K. Iwao, K. Nishida, T. Kinoshita, and Y. Yamagata, "Validating land cover maps with Degree Confluence Project information," *Geophys. Res. Lett.*, vol. 33, p. L23404, 2006, doi 10.1029/2006GL027768.
- [21] M. F. Goodchild and J. A. Glennon, "Crowdsourcing geographic information for disaster response: A research frontier," *Int. J. Digital Earth*, vol. 3, pp. 231–241, 2010.
- [22] "Committee on Strategic Directions for the Geographical Sciences in the Next Decade," in *Understanding the Changing Planet: Strategic Directions for Geog. Sciences*. Washington, DC, USA: National Academies Press, 2010.
- [23] O. Ahlqvist, "Extending post-classification change detection using semantic similarity metrics to overcome class heterogeneity: A study of 1992 and 2001 US National Land Cover Database changes," *Remote Sens. Environ.*, vol. 112, pp. 1226–1241, 2008.
- [24] A. Comber, P. Fisher, and R. Wadsworth, "What is land cover?," *Environment and Planning B*, vol. 32, pp. 199–209, 2005.
- [25] P. Bicheron, P. Defourny, C. Brockmann, L. Schouten, C. Vancustem, M. Huc, S. Bontemps, M. Leroy, F. Achard, M. Herold, F. Ranera, and O. Arino, *Globcover: Products Description and Validation Report* Medias France. Toulouse, 47pp, 2008.
- [26] D. H. Card, "Using map category marginal frequencies to improve estimates of thematic map accuracy," *Photogramm. Eng. Remote Sens.*, vol. 49, pp. 431–439, 1982.
- [27] D. L. B. Jupp, "The stability of global estimates from confusion matrices," *Int. J. Remote Sens.*, vol. 10, pp. 1563–1569, 1989.
- [28] G. M. Foody, "Local characterization of thematic classification accuracy through spatially constrained confusion matrices," *Int. J. Remote Sens.*, vol. 26, pp. 1217–1228, 2005.
- [29] J. Scepan, G. Menz, and M. C. Hansen, "The DISCover validation image interpretation process," *Photogramm. Eng. Remote Sens.*, vol. 65, pp. 1075–1081, 1999.
- [30] L. Bai, "Comparison and Validation of Five Global Land Cover Products Over the African Continent," degree thesis, Lund University, Lund, Sweden [Online]. Available: http://www.natgeo.lu.se/ex-jobb/exj_195.pdf
- [31] *Global Forest Resources Assessment 2010*, FAO, Rome, Italy, FAO Forestry Paper 163, 2010.
- [32] S. V. Stehman, "Sampling designs for accuracy assessment of land cover," *Int. J. Remote Sens.*, vol. 30, pp. 5243–5272, 2009.
- [33] G. M. Foody, "Sample size determination for image classification accuracy assessment and comparison," *Int. J. Remote Sens.*, vol. 30, pp. 5273–5291, 2009.
- [34] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, vol. 20, pp. 37–46, 1960.
- [35] G. M. Foody, "On the compensation for chance agreement in image classification accuracy assessment," *Photogramm. Eng. Remote Sens.*, vol. 58, pp. 1459–1460, 1992.
- [36] R. G. Pontius and M. Millones, "Death to kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment," *Int. J. Remote Sens.*, vol. 32, pp. 4407–4429, 2011.
- [37] I. Yang and M. P. Becker, "Latent variable modelling of diagnostic accuracy," *Biometrics*, vol. 53, pp. 948–958, 1997.
- [38] J. K. Vermunt, *Log-linear Models for Event Histories*. Thousand Oaks, CA, USA: Sage, 1997.
- [39] J. Magidson and J. K. Vermunt, "Latent class models, In Kaplan," in *The SAGE Handbook of Quantitative Methodology for the Social Sciences*. Thousand Oaks, CA, USA: Sage, 2004, pp. 175–198, D (editor).
- [40] P. Magdon and C. Klein, "Uncertainties of forest area estimates caused by the minimum crown cover criterion—A scale issue relevant to forest cover monitoring," *Environ. Monit. Assess.*, 2012, doi 10.1007/s10661-012-2950-0, open access.



Giles M. Foody (M'01–SM'10–F'13) received the B.Sc. and Ph.D. degrees from the University of Sheffield, U.K., in 1983 and 1986, respectively.

He is currently Professor of Geographical Information Science at the University of Nottingham, U.K. His main research interests focus on the interface between remote sensing, ecology, and informatics. Topics of particular interest relate to image classification for land cover mapping and monitoring applications, addressing issues at scales ranging from the sub-pixel to global. His publication

list includes 8 books and more than 160 refereed journal articles.

Prof. Foody serves as editor-in-chief of the *International Journal of Remote Sensing* and, as founding editor-in-chief, of *Remote Sensing Letters*. He holds additional editorial roles on eight other journals including *Ecological Informatics*, *Landscape Ecology*, *Remote Sensing of Environment*, *Geocarto International*, and the *International Journal of Applied Earth Observation and Geoinformation*. He has refereed for over 110 journals and served on numerous committees and panels, notably the U.K.'s two most recent research assessment exercises: RAE2008 and REF2014. He is chair of the European Union COST Action on 'Mapping and the Citizen Sensor' (TD1202) which involves researchers from more than 25 countries.



Doreen S. Boyd received the B.Sc. degree in geography from the University of Wales, U.K., in 1992 and received the Ph.D. degree from the University of Southampton, U.K., in 1996.

She is currently an Associate Professor in the School of Geography, University of Nottingham, U.K., having held Lectureships at Manchester, Kingston, and Bournemouth Universities. Between 2004 and 2006, she was Senior Research Leader in Research and Innovation at the Ordnance Survey, GB. Her main research interests are in the remote

sensing of terrestrial ecosystems and the monitoring of environmental change. She serves on the editorial board of the *Journal of Maps*. She is co-founder and co-chair of the laser scanning and LiDAR special interest group of the Remote Sensing and Photogrammetry Society.