Article:

See, L, Schepaschenko, D, Lesiv, M et al. (21 more authors) (2015) Building a hybrid land cover map with crowdsourcing and geographically weighted regression. ISPRS Journal of Photogrammetry and Remote Sensing, 103. pp. 48-56. ISSN 0924-2716

https://doi.org/10.1016/j.isprsjprs.2014.06.016

Building a Hybrid Land Cover Map with Crowdsourcing and Geographically Weighted Regression

[1]Linda See, [1]Dmitry Schepaschenko, [2,3]Myroslava Lesiv, [1]Ian McCallum, [1]Steffen Fritz, [4]Alexis Comber, [1]Christoph Perger, [5]Christian Schill, [6]Yuanyuan Zhao, [7]Victor Maus, [8]Muhammad Athar Siraj, [9]Franziska Albrecht, [10,11]Anna Cipriani, [1]Mar'yana Vakolyuk, [12]Alfredo Garcia, [13]Ahmed H. Rabia, [14]Kuleswar Singha, [15]Abel Alan Marcarini, [16]Teja Kattenborn, [14]Rubul Hazarika, [17]Maria Schepaschenko, [18]Marijn van der Velde, [1]Florian Kraxner and [1]Michael Obersteiner

[1]International Institute of Applied Systems Analysis (IIASA), Ecosystem Services and Management Program, Schlossplatz 1, Laxenburg, A-2361, Austria Email: see@iiasa.ac.at
Tel: +43 2236 807 423

[2]Lviv Polytechnic National University, Lviv, Ukraine

[3]Systems Research Institute of Polish Academy of Sciences, Newelska 6, 01-447 Warsaw, Poland

[4]Department of Geography, University of Leicester, Leicester, UK

[5]Felis, University of Freiburg, Germany

[6]Ministry of Education Key Laboratory for Earth System Modelling, Centre for Earth System Science, Tsinghua University, Beijing 100084, China

[7]National Institute for Space Research (INPE), São José dos Campos, Brazil

[8]Institute of Geographical Information System, National University of Sciences and Technology (IGIS-NUST), Islamabad, Pakistan

[9]University of Vienna, Department of Geography and Regional Research, Vienna, Austria

[10]Dipartimento di Scienze Chimiche e Geologiche, Università degli Studi di Modena e Reggio Emilia, Largo S. Eufemia 19, 41121 Modena, Italy

[11]Lamont Doherty Earth Observatory of Columbia University, Palisades, New York, 10964, USA

[12]National Institute of Agricultural Technology (INTA), Buenos Aires, Argentina

[13]Faculty of Agriculture, Alexandria University, Egypt

[14]Gauhati University, Assam, India

[15]AgroParis Tech, Nancy, France

[16]Fakulty of Forestry and Environmental Sciences, Albert-Ludwigs University, Tennenbacher 4, 79085, Freiburg, Germany

[17]Russian Institute of Continuous Education in Forestry, Pushkino, Russia

[18]Joint Research Centre of the European Commission, Via Fermi 2749, TP 266, Ispra, Italy

Abstract

Land cover is of fundamental importance to many environmental applications and serves as critical baseline information for many large scale models e.g. in developing future scenarios of land use and climate change. Although there is an ongoing movement towards the development of higher resolution global land cover maps, medium resolution land cover products (e.g. GLC2000 and MODIS) are still very useful for modelling and assessment purposes. However, the current land cover products are not accurate enough for many applications so we need to develop approaches that can take existing land covers maps and produce a better overall product in a hybrid approach. This paper uses geographically weighted regression (GWR) and crowdsourced validation data from Geo-Wiki to create two hybrid global land cover maps that use medium resolution land cover products as an input. Two different methods were used: a) the GWR was used to determine the best land cover product at each location; b) the GWR was only used to determine the best land cover at those locations where all three land cover maps disagree, using the agreement of the land cover maps to determine land cover at the other cells. The results show that the hybrid land cover map developed using the first method resulted in a lower overall disagreement than the individual global land cover maps. The hybrid map produced by the second method was also better when compared to the GLC2000 and GlobCover but worse or similar in performance to the MODIS land cover product depending upon the metrics considered. The reason for this may be due to the use of the GLC2000 in the development of GlobCover, which may have resulted in areas where both maps agree with one another but not with MODIS, and where MODIS may in fact better represent land cover in those situations. These results serve to demonstrate that spatial analysis methods can be used to improve medium resolution global land cover information with existing products.

Keywords (maximum of 6): land cover, validation, crowdsourcing, map integration, global land cover

1 Introduction

Spatially explicit information about land cover is of fundamental importance for many applications including nature protection and biodiversity, forest and water management, urban and transport planning, natural hazard prevention and mitigation, and the evaluation of agricultural policies. The importance of global land cover is recognized through its status as an essential climate variable (GCOS, 2013), where this information serves as a critical input to the monitoring of climate change. Global land cover forms a key input to large scale economic land use models (e.g. Havlík et al., 2011), which are used to determine important quantities such as the amount of land available for agricultural expansion, afforestation projects and biofuel production or whether reducing emissions from deforestation and forest degradation (REDD) are the most cost-effective solutions. A critical gap in accurate land cover and land use, which is needed to monitor ecosystem services and change over time, has also been highlighted recently by Tallis et al. (2012).

A number of different coarse to medium resolution global land cover products exist, e.g. the GLC2000 (Fritz et al., 2003), MODIS (Friedl et al., 2010) and GlobCover (Bicheron et al., 2008). These products, which vary from 1km to 300m resolution at the equator, have been developed using data from different satellite sensors and using different classification algorithms with varying degrees of automation. Although the published accuracies of these products vary between 68.5 to 74.8%, recent studies have shown that when these maps are compared, there are significant amounts of spatial disagreement across different land cover types, in particular in the cropland and forest domains even when taking semantic differences in the legend definitions into account (Fritz and See, 2008; Fritz et al., 2011a). Research has also shown that model outcomes can vary significantly when different land cover products are used in the same modelling exercise (Quaife et al., 2008; Seebach et al., 2012) while Fritz et al. (2012a) have demonstrated the value associated with reducing the uncertainty in land cover with regards to the cost of different climate mitigation options.

With the opening up of the Landsat archive (Wulder et al., 2012), one of the most recent trends in global land cover mapping has been to produce higher resolution products, i.e. at 30m (Gong et al., 2013; Yu et al., 2013b), with others currently in the pipeline by groups in China and the USA. The accuracies of these recently produced 30m products range from 63.7% to 66.0%. The technology and algorithms for classifying Landsat will undoubtedly improve in the future, and there will be new higher resolution sensors coming online soon where the data will be freely available (e.g. Sentinel II). Moreover, there are other multi-temporal and/or multi-sensor classification efforts ongoing (Lu et al., 2011; Roy et al., 2010). Despite this relatively positive outlook for land cover mapping in the future, there is still an urgent need for better land cover maps at the present time. Medium resolution products are also still extremely useful from a modelling and assessment point of view where the issue is not one of needing to improve the resolution for many applications but simply improving the accuracy.

One method which can be used to address this issue of accuracy is to merge existing land cover maps to create an integrated or hybrid product where the resulting accuracy should be higher than the accuracies of the individual products. Data fusion and soft computing are domains which are based on the integration of data from a variety of sources (e.g. from different sensors, models or approaches) so this idea is not new in itself. For example, Jung et al. (2006) developed a fuzzy agreement scoring method to determine the synergies between global land cover products for modelling the carbon cycle while Fritz et al. (2011b) employed this synergy concept in combination with expert knowledge to rank land cover products in order to combine them into a single cropland map of Africa. Iwao (2011) integrated the GLC2000, MODIS and the older University of Maryland (UMD) land cover product using a simple majority voting approach and validated the resulting map with data from the Degrees of Confluence project. However, the resulting improvements in accuracy were not statistically significant. More recently, Yu et al. (2013a) used a decision tree to combine two 30 m cropland products with a 250 m cropland probability layer to produce a global cropland mask. All of these approaches have demonstrated the increased accuracy that has resulted from the integration of existing products.

What these types of integration methods need are much larger amounts of data for training and validation. One potential source is Geo-Wiki, which is a visualisation, crowdsourcing and validation tool developed to help improve global land cover maps (Fritz et al., 2012b, 2009) where crowdsourcing is the use of the volunteers (which can also be experts) to help collect and analyse data (Howe, 2006; Heipke, 2010). Using Google Earth, volunteers are asked to indicate the land cover types that are visible from the images displayed in Geo-Wiki. Samples have been collected through a number of different Geo-Wiki campaigns that have run over the past few years (Perger et al., 2012; See et al., 2014a) and then used in subsequent research, e.g. Fritz et al. (2013a) used data from the first campaign to downgrade estimates of land availability for biofuels. This database, which represents a valuable source of data for both training and validation of land cover, continues to grow, with more than 4.5 million samples collected recently on the presence of cropland using a game version of Geo-Wiki (See et al., 2014b). Initial attempts were undertaken by Comber et al. (2013) to map the areas of highest correspondence between the Geo-Wiki crowdsourced data and the GLC2000, MODIS and GlobCover for one land cover type, i.e. tree cover, for a section of western Africa. The authors employed crowdsourced data from the first Geo-Wiki competition (Perger et al. 2012) and geographically weighted regression (GWR), which is a spatial extension to linear regression in which the coefficients of the regression equation are able to vary across space, which captures any effects due to location (Brunsdon et al., 1998).

The aim of this paper is to extend the research of Comber et al. (2013) in a number of ways. Firstly, we apply the method globally to all land cover types using larger amounts of crowdsourced data from four campaigns, including one campaign focussed on data collection in areas where all global land cover maps disagree with one another. Secondly, we implement a second approach in which the method is only applied to those areas where there is complete disagreement between land cover products, taking agreement between two or more land cover products as the land cover type at all other locations. Finally, we use an independent crowdsourced dataset to validate the products, using the sampling scheme of the global validation dataset created by Zhao et al. (2014) as the basis for collecting data via Geo-Wiki. A quality assured subset of this validation data is used.

An overview of the input data is provided in the next section, i.e. three global land cover maps and the crowdsourced data from Geo-Wiki, which are used in the hybrid methodology. The validation methods as outlined in Pontius and Millones (2011) are then presented, which are applied to the hybrid land cover products and the individual global land cover maps using a combination of a crowdsourced and expert derived validation data set as outlined below. Advantages and disadvantages of the methodology are then discussed along with ideas for further research.

2 Materials and Methods

2.1 Global Land Cover Maps

Three global land cover datasets have been used in the creation of the hybrid land cover map. The first is the GLC2000, which was developed by the Joint Research Centre (JRC) of the European Commission using 14 months of SPOT Vegetation data (Mayaux et al., 2006). This mapping effort was divided into regional windows with research teams around the world contributing to this effort. Maps created at the regional level were then harmonised to create a single global product. The exercise was intended to provide baseline information for the environmental year 2000 as a one-off land cover mapping exercise. The GLC2000 has the coarsest resolution of 1km at the equator. The MODIS global land cover product is developed by Boston University using data from the Moderate Resolution Imaging Spectroradiometer (MODIS) on board the NASA Terra satellite. In contrast to the GLC2000, this land cover product is created using an automated classification algorithm, which allows the MODIS team to create a new product on an annual basis and update previous years with the latest methodological developments (Friedl et al., 2010). The land cover map for 2005 at a resolution of 500 m is used here. At 300 m, GlobCover is the finest resolution product available for the year 2005-2006 (Bicheron et al., 2008). This product was intended to update and complement other existing comparable global products. As with MODIS, the intention was to repeat the exercise on a regular basis. Although later temporal versions of GlobCover and MODIS are available, the idea is to create a hybrid land cover map for the year 2005.

Since the resolutions of the individual global land cover products differ, a common resolution of 300 m was chosen in order to retain the maximum information content. Therefore, the GLC2000 and MODIS land cover maps were resampled to match the grid with the highest resolution, i.e. GlobCover. The legends associated with each product were mapped onto a simple legend with 10 land cover classes where the lookup table is provided in Table 1. This simple legend is based on Herold et al. (2008), in which the authors show how different global land cover maps, including the GLC2000 and MODIS products, can be harmonized and therefore compared at an aggregated level using the Land Cover Classification System (LCCS).

Table 1: Lookup table for a simple 10 class legend and each of the global land cover products. The legend descriptions corresponding to the GLC2000, MODIS and GlobCover can be found in Fritz et al. (2003), Friedl et al. (2010) and Bicheron et al. (2008)

Class	Simple Legend Description	GLC2000	MODIS	GlobCover
1	Tree cover	1-10	1-5, 8, 9	40, 50, 60, 70, 90, 100, 110, 160
2	Shrub cover	11, 12	6,7	130
3	Herbaceous vegetation / Grassland	13	10	120, 140
4	Cultivated and managed	16	12	11, 14
5	Mosaic of cultivated and managed / natural vegetation	17, 18	14	20, 30
6	Flooded / wetland	15	11	170,180
7	Urban	22	13	190
8	Snow and ice	21	15	220
9	Barren	14, 19	16	150, 200
10	Open water	20	17	210

2.2 Crowdsourced Data from Google Earth for Training and Validation

Geo-Wiki was originally developed as a way of displaying global land cover maps (GLC2000, MODIS and GlobCover) and layers of spatial land cover disagreement on top of Google Earth. Volunteers were then asked to indicate the quality of the different land cover maps based on comparing the land cover type at a given pixel with the land cover visible from Google Earth (Fritz et al., 2009). Since then Geo-Wiki has been used in a number of crowdsourcing campaigns where we have moved away from determining the quality of particular global land cover maps to collecting land cover information from selected samples on the Earth's surface. Figure 1 provides a screenshot from the Geo-Wiki system from the second campaign, which shows a sample pixel drawn on top of Google Earth along with the types of information that we ask volunteers to collect, e.g. the land cover type, what percentage of the pixel this covers and the human impact visible from this pixel (see See et al. (2013) for further information on human impact).

<Insert Figure 1 here>

The crowdsourced data used in this paper come from four different Geo-Wiki campaigns as outlined in Table 2. Data from two of the campaigns were used for training while data from the other two were used for independent validation. The first hybrid map was trained using data from the first competition on human impact, which is a globally distributed random sample. The second hybrid map was trained using data from the second competition, where the sample was stratified by disagreement. In particular, the areas sampled were in locations where the three global land cover maps disagree with one other. The spatial distribution of these two training datasets is shown in Figure 2.

<Insert Figure 2 here>

Table 2: Geo-Wiki crowdsourced data used in the development of the hybrid land cover maps

Competition	**Purpose of the Competition**	**Use in the Hybrid Methodology**
1 Human Impact	To validate a map of land availability for biofuel production (Fritz et al., 2013a)	Training dataset for Hybrid Map 1
2 Hotspots of Map Disagreement	To collect validation points in the areas were the GLC2000, MODIS and GlobCover disagree with one another	Training dataset for Hybrid Map 2
3 Wilderness	To collect land cover and human impact in order to determine the amount of global wilderness. The locations used were the same as that of the Chinese 30 m land cover map (See et al., 2014a)	Combined with the Global Validation Data set to validate both hybrid maps
4 Global Validation Dataset	To collect data at the same locations as the validation data assembled for the Chinese 30 m land cover map (Zhao et al., 2014)	Combined with the Wilderness dataset to validate both hybrid maps

More details of the validation data are provided in section 2.4 but the samples from the two competitions (Table 2) were purposely chosen to align with the location of the global validation dataset developed by Zhao et al. (2014) in their evaluation of the Chinese 30-m global land cover product.

2.3 Methodology for Creation of the Hybrid Land Cover Maps

To combine the three global land cover products, GWR is employed as previously implemented for western Africa and tree cover in Comber et al. (2013). GWR estimates model parameters at each geographical location by using a kernel. In addition, the observations are weighted by distance, so those closer to the studied location have more influence on the parameter estimates. The basic GWR equation is:

$$y(u,v) = b_0(u,v) + b_1(u,v)x_1 + \varepsilon(u,v) \tag{1}$$

where y is the dependent variable with a Gaussian distribution; x is the independent variable; u,v are the coordinates of the data; b_0 is the intercept term; b_1 is the coefficient being estimated and ε is the random error term.

GWR also has extensions of generalized linear models, including logistic and Poisson regressions (Fotheringham et al., 2002). Here we used logistic regression to calculate the probability of correspondence between the validation data and the global datasets at each pixel of a 300 m grid according to:

$$P(y_i = 1) = \log\mathrm{it}\,(b_{(u_i, v_i)}) \tag{2}$$

where $P(y_i = 1)$ is the probability that a given global land cover product defines the same land cover type as the validation data at each location i; logit is a logistic regression; (u_i, v_i)

is the two-dimensional vector of location i; and $b_{(u_i,v_i)}$ is the intercept. The optimal size of the window in terms of how many validation samples to use in each local instance of logistic regression is first calculated based on the overall number of points and their spatial distribution. The observed outcome is a probability associated with each global land cover product where the highest probability determines which land cover product to choose. This procedure was implemented in R in two steps: (1) estimating the intercept $b_{(u_i,v_i)}$ – for input we used a vector with the information whether or not a global land cover product defines the same land cover type as the crowdsourced data at each observed location; and (2) estimating the probabilities by implementing eqn (2).

Using GWR two different hybrid land cover maps were created:

- Hybrid Map 1: Using the method described above, the three land cover maps were trained using crowdsourced data at each pixel using a coarser 0.25 degree grid. The data used to train the map were from the first Geo-Wiki competition (see Table 2). The optimal number of training points suggested by the algorithm at each location was seven, which were then used to evaluate the correspondence to each individual land cover product. After running the procedure, three probability maps were created, one for each land cover product. These three maps were then combined into a single layer showing which of the three land cover maps had the highest probability at each location. This probability layer was then used to create the final hybrid map, selecting the land cover type from the product at a 300m resolution.
- Hybrid Map 2: The three maps were first compared and where two maps agreed on a land cover type, this value was used as the land cover at that cell. Where the three land cover maps disagreed, the method described for Hybrid Map 1 was applied, this time using data specifically from the second competition which was focused on collecting validation samples specifically in these disagreeing areas (see Table 2).

2.4 Map Validation

In order to estimate the quality of the two hybrid products and to determine if they have achieved a measureable improvement over the existing products (i.e. GLC2000, MODIS and Globcover), it is necessary to validate all of the datasets against a common sample dataset. The validation data set used here was developed from three different sources in an effort to ensure the highest quality. The first was an external dataset that was developed specifically for the validation of the Chinese 30 m global land cover map (Gong et al., 2013; Zhao et al., 2014), which contains 38,664 sample units globally based on Landsat and Google Earth image interpretation. The sample design utilised equal-area stratified random sampling, which partitioned the global land area into approximately 7,000 equal-area hexagons, inside of which 5 random samples were chosen from each one. This data set represents the land cover at an x,y location on the Earth's surface. The x,y locations of the validation data from this exercise were then used as the sample that was provided to volunteers in two different Geo-Wiki crowdsourcing campaigns, where volunteers were asked to indicate the percentage of land cover in a 1 km pixel (the pixel was centred on the x,y location). The samples from the two different Geo-Wiki crowdsourcing campaigns were then compared with one another where there was more than one answer provided at each sample pixel. Only those pixels where there was agreement between the two competitions were then kept. Finally, we selected only homogeneous pixels where the land cover type agreed with the x,y location class. This reduced the final validation dataset to a total of 5,096 1km^2 pixels.

Owing to the steps outlined above, a potential bias is present in the remaining sample, i.e. oversampling in the tree cover class. Although this can increase overall accuracies, we are actually more interested in relative performance so we will only judge the products in relation to one another rather than in terms of absolute performance.

We have applied the standard validation test of a cross-tabulation matrix with both hybrid products and the three existing products, using the above mentioned validation sample dataset. Initially we converted the observed sample matrix into an estimated unbiased population matrix which represents the entire study area (Pontius and Millones, 2011). From this we calculated two measures, i.e. the quantity disagreement, q, and the allocation disagreement, a, calculated as follows for each land cover type, g:

$$q_g = \left| \sum_{i=1}^{J} x_{ig} - \sum_{j=1}^{J} x_{gj} \right| \tag{2}$$

$$a_g = 2 * \min\left[\sum_{i=1}^{J} x_{ig} - p_{gg}, \sum_{j=1}^{J} x_{gj} - p_{gg} \right] \tag{3}$$

where J is the total number of categories, x_{ij} are the elements of a standard confusion matrix, i refers to the map being evaluated and j is the reference data, in this case the crowdsourced validation data from Geo-Wiki. The total quantity disagreement and total allocation disagreement are then calculated by summing across all land cover categories and dividing by 2 since these errors are double counted due to the way in which they are formulated. The total disagreement is then the sum of the quantity disagreement and the allocation disagreement. We have also reported the percentage correct or overall accuracy since this is one of the most commonly reported accuracy measures in the literature.

3 Results

This study has resulted in the first global hybrid land cover maps obtained with the aid of crowdsourced training data. Figure 4 shows the land cover map chosen at each location based on the highest probability calculated by the GWR for Hybrid Map 1. Probability is assigned to the three datasets as: GLC2000 (33%), MODIS (41%) and GlobCover (25%). While the resulting map is quite heterogeneous, global patterns nonetheless emerge. GLC2000, together with GlobCover, have the highest probability in the far north. This could be explained by the fact that there are missing MODIS data in the northern high latitudes, which can lead to degradation in the performance of the classification algorithm (Friedl et al., 2010). The continental US is ranked highest by MODIS.

Based on the results demonstrated in Figure 4, two hybrid maps were derived (Figures 5 and 6). Significant differences are visible between the two resulting hybrid products. Patterns visible in the probability map (Figure 4) are sometimes apparent in Hybrid Map 1 (Figure 5). An example of this is Australia where barren areas appear in stark contrast to herbaceous areas. Hybrid Map 2, with its focus on areas of complete disagreement among the input datasets, produces a map which differs less from the original input datasets (Figure 6).

A quantitative validation of the hybrid products was performed using cross-tabulation. Results are presented here from sample matrices for Hybrid Map 1 and Hybrid Map 2 (Tables 3 and 4). Note that because of edge effects, the different resolutions of the input maps, and the methodologies used to produce Hybrid Maps 1 and 2, around 200 validation samples fell outside of a land boundary and were therefore not used in the final validation. For this reason there are also slight differences in the total number of validation samples used in validating Hybrid Maps 1 and 2.

For both maps, the users and producers accuracies are highest in the forest classes. Surprisingly the class cultivated and managed also performed well. In both maps the greatest confusion occurs between these two classes and the shrub cover and herbaceous vegetation/grassland classes.

Table 3: The cross-tabulation matrix for Hybrid Map 1

	Reference Class										
Map Class	1	2	3	4	6	7	8	9	10	Row Total	User's Accuracy
1	2653	99	27	38	5		3	5	5	2835	0.94
2	37	240	20	38		4	4	75	2	420	0.57
3	28	96	68	84	2		3	55	2	338	0.20
4	18	23	28	669	1	5		6	2	752	0.89
6	27	1	4	1	3	1			5	42	0.07
7				2		5			1	8	0.63
8							57	1	3	61	0.93
9	12	18	25	13			10	170	4	252	0.67
10	4	2	1		1		2	3	186	199	0.93
Col. Total	2779	479	173	845	12	15	79	315	210	4907	
Producer's Accuracy	0.95	0.50	0.39	0.79	0.25	0.33	0.72	0.54	0.89		

Table 4: The cross-tabulation matrix for Hybrid Map 2

	Reference Class										
Map Class	1	2	3	4	6	7	8	9	10	Row Total	User's Accuracy
1	2676	119	35	45	5		3	4	4	2891	0.93
2	40	174	18	30		1	2	50	2	317	0.55
3	15	112	50	92	2		5	68	1	345	0.14
4	12	17	28	661	1	4		7	2	732	0.90
6	18	2	1	3	2	1			2	29	0.07
7				1		7			1	9	0.78
8							59	5	2	66	0.89
9	7	54	37	17	1	2	8	182	1	309	0.59
10	2	2	1	1	1		2	3	200	212	0.94
Col. Total	2770	480	170	850	12	15	79	319	215	4910	
Producer's Accuracy	0.97	0.36	0.29	0.78	0.17	0.47	0.75	0.57	0.93		

Table 4 lists the performance measures as outlined in section 2.4 for each of the individual land cover products (resampled to the 300m resolution of GlobCover and reclassified to the harmonized simple legend as outlined in section 2.1) and Hybrid Maps 1 and 2. A map can be judged as better in relation to another the lower the three disagreement measures are and the higher the percentage correct. Overall, Hybrid map 1 performs better in relation to Hybrid Map 2 across all measures and in relation to the individual global land cover products. Both hybrid products perform similarly in terms of the quantity disagreement, with lower values than any of the individual global land cover products where both GLC2000 and GlobCover show the highest quantity disagreement. The allocation disagreement, however, shows a different picture. For this measure, Hybrid Map 1 has the lowest value while MODIS has the second lowest allocation disagreement. Thus when viewed from the perspective of total disagreement and percentage correct, the MODIS land cover product is slightly better than Hybrid Map 2. There are a number of reasons why this may be the case. Firstly, GLC2000 was used in the development of GlobCover so they are not completely independent of one another. Thus there may be situations where both products agree but disagree with MODIS yet MODIS may actually better represent the land cover at those locations. Secondly, since the training algorithms for MODIS continue to be improved and applied retrospectively to previous land cover products, the MODIS land cover product may have improved over time. Thus an alternative would have been to take only those situations where all three maps agree or where MODIS agrees with at least one of the other two followed by GWR on disagreeing areas.

Table 4: Performance measures for the individual land cover maps and the two hybrid products

Land Cover Map	Quantity Disagreement	Allocation Disagreement	Total Disagreement	Percentage Correct
GLC2000	13.4	12.7	26.1	73.9
MODIS	7.6	9.0	16.6	83.4
GlobCover	13.4	14.5	27.9	72.1
Hybrid Map 1	5.9	6.2	12.1	87.9
Hybrid Map 2	5.6	11.6	17.2	82.8

4 Conclusions

This paper demonstrated how existing global land cover maps can be integrated into a hybrid product using GWR and a training dataset obtained through crowdsourcing. Two different methods were used, one involving a global data set from the first Geo-Wiki competition to determine the best land cover map to use at a given location, and the second focussed only on correcting areas where all three land cover maps disagree. These hybrid products would generally represent the year 2005 as MODIS and Globcover apply to that year. Although the GLC2000 is for the year 2000, the majority of spatial disagreements between the products are not about land cover change but about incorrect classifications. Thus the merging of the products is really about trying to find the best land cover representation more generally for that time period. The two hybrid products were compared with the individual global land cover products using performance metrics suggested by Pontius and Millones (2011) as well as overall accuracy as an additional relative measure. The first hybrid map outperformed the individual land cover maps based on the validation data set used while the second hybrid map was not as good as the individual MODIS land cover product on three out of four performance measures. We offered potential explanations for this including the use of the GLC2000 in the development of GlobCover and the continued improvement of MODIS over time. Other variations that might be tried are to: take the land cover at a given point only when MODIS agrees with one of the other land cover products (or all three agree); and use the two crowdsourced training datasets and GWR at all locations to create a single hybrid product.

The use of crowdsourced data for both calibration and validation of the hybrid products as well as for the development of future land cover products represents an area of great potential. Although quality is clearly a very important issue, some initial research in this area has shown that the data are of sufficient quality for use in further scientific research or that methods can be put in place to correct for errors and biases automatically (Foody et al., 2013; See et al., 2013). There is much to be learned from ongoing quality control measures implemented in ecological citizen science projects that are using the data for rigorous scientific research as well as from businesses who need to control for quality (Delaney et al., 2007; Bonter and Cooper, 2012; Allahbakhsh et al., 2013). We will continue to collect further crowdsourced data via Geo-Wiki and expand this data source. Moreover, the opening up of calibration and validation data sets via the GOFC/GOLD portal and through collaborations with Chinese colleagues also means that the amount of data becoming openly available is rapidly expanding.

Only three global land cover maps were used in the creation of the hybrid product but we could easily extend this exercise to include other hybrid maps, e.g. the hybrid cropland product that integrates cropland information from more than 25 countries (Fritz et al., 2013b), as well as the new 30m land cover products that are starting to appear. Incorporation of regional and national maps, e.g. CORINE land cover and countries mapped by the Africover initiative, could also be attempted to improve the hybrid product further. While there will continue to be new mapping initiatives, new sensors appearing and advances in classification algorithms, the hybrid approach represents a simple, low cost and promising way of improving information on global land cover in the short term while we look to improved accuracies from remote sensing in the future. It is therefore important to develop methods and tools that optimally use and integrate existing products rather than focussing all of our efforts and funds on producing only new products. In fact, new mapping efforts should focus on improving representation of areas where uncertainty is the highest or where land cover is changing quickly.

Acknowledgements

This research was supported by the Austrian Research Funding Agency (FFG) via the LandSpotting (No. 828332) and FarmSupport (No. 833421) projects and via the EU-funded AGRICISTRADE (No. 612755) project.

References

Allahbakhsh, M., Benatallah, B., Ignjatovic, A., Motahari-Nezhad, H.R., Bertino, E., Dustdar, S., 2013. Quality control in crowdsourcing systems: Issues and directions. IEEE Internet Comput. 17, 76–81. doi:10.1109/MIC.2013.20

Bicheron, P., Defourny, P., Brockmann, C., Schouten, L., Vancutsem, C., Huc, M., Bontemps, S., Leroy, M., Achard, F., Herold, M., Ranera, F., Arino, O., 2008. Globcover: Products description and validation report.

Bonter, D.N., Cooper, C.B., 2012. Data validation in citizen science: a case study from Project FeederWatch. Front. Ecol. Environ. 10, 305–307. doi:10.1890/110273

Brunsdon, C., Fotheringham, S., Charlton, M., 1998. Geographically Weighted Regression-Modelling Spatial Non-Stationarity. J. R. Stat. Soc. Ser. Stat. 47, 431–443. doi:10.2307/2988625

Comber, A., See, L., Fritz, S., Van der Velde, M., Perger, C., Foody, G., 2013. Using control data to determine the reliability of volunteered geographic information about land cover. Int. J. Appl. Earth Obs. Geoinformation 23, 37–48. doi:10.1016/j.jag.2012.11.002

Delaney, D.G., Sperling, C.D., Adams, C.S., Leung, B., 2007. Marine invasive species: validation of citizen science and implications for national monitoring networks. Biol. Invasions 10, 117–128. doi:10.1007/s10530-007-9114-0

Foody, G.M., See, L., Fritz, S., Van der Velde, M., Perger, C., Schill, C., Boyd, D.S., 2013. Assessing the Accuracy of Volunteered Geographic Information arising from Multiple Contributors to an Internet Based Collaborative Project. Trans. GIS n/a–n/a. doi:10.1111/tgis.12033

Fotheringham, A.S., Brunsdon, C., Charlton, M., 2002. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. John Wiley & Sons, Chichester.

Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sens. Environ. 114, 168–182. doi:10.1016/j.rse.2009.08.016

Fritz, S., Bartholomé, E., Belward, A., Hartley, A., Stibig, H.-J., Eva, H., Mayaux, P., 2003. Harmonisation, mosaicing and production of the Global Land Cover 2000 database (Beta Version) (No. EUR 20849EN). Office for Official Publications of the European Communities, Luxembourg.

Fritz, S., Fuss, S., Havlik, P., Szolgayova, J., McCallum, I., Obersteiner, M., See, L., 2012a. The value of determining global land cover for assessing climate change mitigation options, in: Laxminarayan, R., Macauley, M.K. (Eds.), The Value of Information: Methodological Frontiers and New Applications in Environment and Health. Springer, Dordrecht, pp. 193–230.

Fritz, S., McCallum, I., Schill, C., Perger, C., Grillmayer, R., Achard, F., Kraxner, F., Obersteiner, M., 2009. Geo-Wiki.Org: The use of crowdsourcing to improve global land cover. Remote Sens. 1, 345–354. doi:10.3390/rs1030345

Fritz, S., McCallum, I., Schill, C., Perger, C., See, L., Schepaschenko, D., van der Velde, M., Kraxner, F., Obersteiner, M., 2012b. Geo-Wiki: An online platform for improving global land cover. Environ. Model. Softw. 31, 110–123. doi:10.1016/j.envsoft.2011.11.015

Fritz, S., See, L., 2008. Identifying and quantifying uncertainty and spatial disagreement in the comparison of Global Land Cover for different applications. Glob. Change Biol. 14, 1057–1075. doi:10.1111/j.1365-2486.2007.01519.x

Fritz, S., See, L., McCallum, I., Schill, C., Obersteiner, M., van der Velde, M., Boettcher, H., Havlík, P., Achard, F., 2011a. Highlighting continued uncertainty in global land cover maps for the user community. Environ. Res. Lett. 6, 044005. doi:10.1088/1748-9326/6/4/044005

Fritz, S., See, L., van der Velde, M., Nalepa, R.A., Perger, C., Schill, C., McCallum, I., Schepaschenko, D., Kraxner, F., Cai, X., Zhang, X., Ortner, S., Hazarika, R., Cipriani, A., Di Bella, C., Rabia, A.H., Garcia, A., Vakolyuk, M., Singha, K., Beget, M.E., Erasmi, S., Albrecht, F., Shaw, B., Obersteiner, M., 2013a. Downgrading recent estimates of land available for biofuel production. Environ. Sci. Technol. 47, 1688–1694. doi:10.1021/es303141h

Fritz, S., See, L., You, L., Justice, C., Becker-Reshef, I., Bydekerke, L., Cumani, R., Defourny, P., Erb, K., Foley, J., Gilliams, S., Gong, P., Hansen, M., Hertel, T., Herold, M., Herrero, M., Kayitakire, F., Latham, J., Leo, O., McCallum, I., Obersteiner, M., Ramankutty, N., Rocha, J., Tang, H., Thornton, P., Vancutsem, C., van der Velde, M., Wood, S., Woodcock, C., 2013b. The Need for Improved Maps of Global Cropland. Eos Trans. Am. Geophys. Union 94, 31–32. doi:10.1002/2013EO030006

Fritz, S., You, L., Bun, A., See, L., McCallum, I., Schill, C., Perger, C., Liu, J., Hansen, M., Obersteiner, M., 2011b. Cropland for sub-Saharan Africa: A synergistic approach using five land cover data sets. Geophys. Res. Lett. 38, L04404. doi:10.1029/2010GL046213

GCOS, 2013. GCOS Essential Climate Variables [WWW Document]. URL http://www.wmo.int/pages/prog/gcos/index.php?name=EssentialClimateVariables (accessed 9.3.13).

Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., Li, C., Li, X., Fu, W., Liu, C., Xu, Y., Wang, X., Cheng, Q., Hu, L., Yao, W., Zhang, H., Zhu, P., Zhao, Z., Zhang, H., Zheng, Y., Ji, L., Zhang, Y., Chen, H., Yan, A., Guo, J., Yu, L., Wang, L., Liu, X., Shi, T., Zhu, M., Chen, Y., Yang, G., Tang, P., Xu, B., Giri, C., Clinton, N., Zhu, Z., Chen, J., Chen, J., 2013. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. Int. J. Remote Sens. 34, 2607–2654. doi:10.1080/01431161.2012.748992

Havlík, P., Schneider, U.A., Schmid, E., Böttcher, H., Fritz, S., Skalský, R., Aoki, K., Cara, S.D., Kindermann, G., Kraxner, F., Leduc, S., McCallum, I., Mosnier, A., Sauer, T., Obersteiner, M., 2011. Global land-use implications of first and second generation biofuel targets. Energy Policy 39, 5690–5702. doi:10.1016/j.enpol.2010.03.030

Heipke, C., 2010. Crowdsourcing geospatial data. ISPRS J. Photogramm. Remote Sens. 65, 550–557. doi:10.1016/j.isprsjprs.2010.06.005

Herold, M., Mayaux, P., Woodcock, C.E., Baccini, A., Schmullius, C., 2008. Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. Remote Sens. Environ., Earth Observations for Terrestrial Biodiversity and Ecosystems Special Issue 112, 2538–2556. doi:10.1016/j.rse.2007.11.013

Howe, J., 2006. The rise of crowdsouring. Wired Mag. 14.

Iwao, K., Nasahara,, K.N., Kinoshita, T., Yamagata, Y., Patton, D., Tsuchida, S., 2011. Creation of new global land cover map with map integration. J. Geogr. Inf. Syst. 03, 160–165. doi:10.4236/jgis.2011.32013

Jung, M., Henkel, K., Herold, M., Churkina, G., 2006. Exploiting synergies of global land cover products for carbon cycle modeling. Remote Sens. Environ. 101, 534–553. doi:10.1016/j.rse.2006.01.020

Lu, D., Li, G., Moran, E., Dutra, L., Batistella, M., 2011. A Comparison of Multisensor Integration Methods for Land Cover Classification in the Brazilian Amazon. GIScience Remote Sens. 48, 345–370. doi:10.2747/1548-1603.48.3.345

Mayaux, P., Eva, H., Gallego, J., Strahler, A.H., Herold, M., Agrawal, S., Naumov, S., De Miranda, E.E., Di Bella, C.M., Ordoyne, C., Kopin, Y., Roy, P.S., 2006. Validation of the global land cover 2000 map. IEEE Trans. Geosci. Remote Sens. 44, 1728–1737.

Perger, C., Fritz, S., See, L., Schill, C., van der Velde, M., McCallum, I., Obersteiner, M., 2012. A campaign to collect volunteered geographic information on land cover and human impact, in: Jekel, T., Car, A., Strobl, J., Griesebner, G. (Eds.), GI_Forum 2012: Geovisualisation, Society and Learning. Herbert Wichmann Verlag, Berlin / Offenbach, pp. 83–91.

Pontius, R.G., Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. Int. J. Remote Sens. 32, 4407–4429. doi:10.1080/01431161.2011.552923

Quaife, T., Quegan, S., Disney, M., Lewis, P., Lomas, M., Woodward, F.I., 2008. Impact of land cover uncertainties on estimates of biospheric carbon fluxes. Glob. Biogeochem. Cycles 22.

Roy, D.P., Ju, J., Kline, K., Scaramuzza, P.L., Kovalskyy, V., Hansen, M., Loveland, T.R., Vermote, E., Zhang, C., 2010. Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. Remote Sens. Environ. 114, 35–49. doi:10.1016/j.rse.2009.08.011

See, L., Comber, A., Salk, C., Fritz, S., van der Velde, M., Perger, C., Schill, C., McCallum, I., Kraxner, F., Obersteiner, M., 2013. Comparing the quality of crowdsourced data contributed by expert and non-experts. PLoS ONE 8, e69958. doi:10.1371/journal.pone.0069958

See, L., Fritz, S., Perger, C., Schill, C., Albrecht, F., McCallum, I., Schepaschenko, D., Van der Velde, M., Kraxner, F., Baruah, U.D., Saikia, A., Singh, K., de Miguel, S., Hazarika, R., Sarkar, A., Marcarini, A.A., Baruah, M., Sahariah, D., Changkakati, T., Obersteiner, M., 2014a. Mapping human impact using crowdsourcing, in: Carver, S. (Ed.), Mapping Wilderness: Concepts, Techniques and Applications of GIS. Springer, Dordrecht; Heidelberg.

See, L., Sturn, T., Perger, C., Fritz, S., McCallum, I., Salk, C., 2014b. Cropland Capture: A gaming approach to improve global land cover, in: 17th AGILE International Conference on Geographic Information Science. Castellon, Spain.

Seebach, L., McCallum, I., Fritz, S., Kindermann, G., Leduc, S., Böttcher, H., Fuss, S., 2012. Choice of forest map has implications for policy analysis: A case study on the EU biofuel target. Environ. Sci. Policy 22, 13–24. doi:10.1016/j.envsci.2012.04.010

Tallis, H., Mooney, H., Andelman, S., Balvanera, P., Cramer, W., Karp, D., 2012. A Global System for Monitoring Ecosystem Service Change. BioScience 62, 977–986. doi:10.1525/bio.2012.62.11.7

Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R., Woodcock, C.E., 2012. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. Remote Sens. Environ. 122, 2–10. doi:10.1016/j.rse.2012.01.010

Yu, L., Wang, J., Clinton, N., Xin, Q., Zhong, L., Chen, Y., Gong, P., 2013a. FROM-GC: 30 m global cropland extent derived through multisource data integration. Int. J. Digit. Earth 6, 521–533. doi:10.1080/17538947.2013.822574

Yu, L., Wang, J., Gong, P., 2013b. Improving 30 m global land-cover map FROM-GLC with time series MODIS and auxiliary data sets: a segmentation-based approach. Int. J. Remote Sens. 34, 5851–5867. doi:10.1080/01431161.2013.798055

Zhao, Y., Gong, P., Yu, L., 2014. Towards a common validation sample set for global land cover mapping. Int. J. Remote Sens. In press.

Figure Captions

Figure 1: A screenshot from the Geo-Wiki system showing how validation data are collected

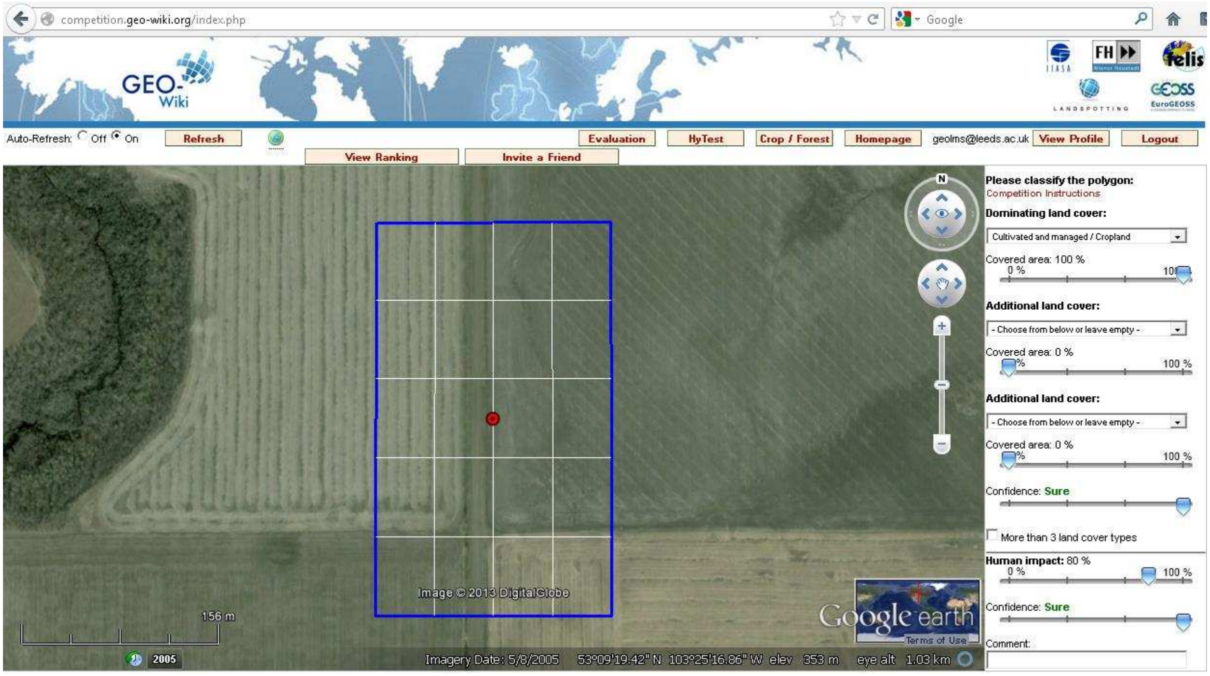

Figure 2: Distribution of the training data for creation of the hybrid global land cover maps

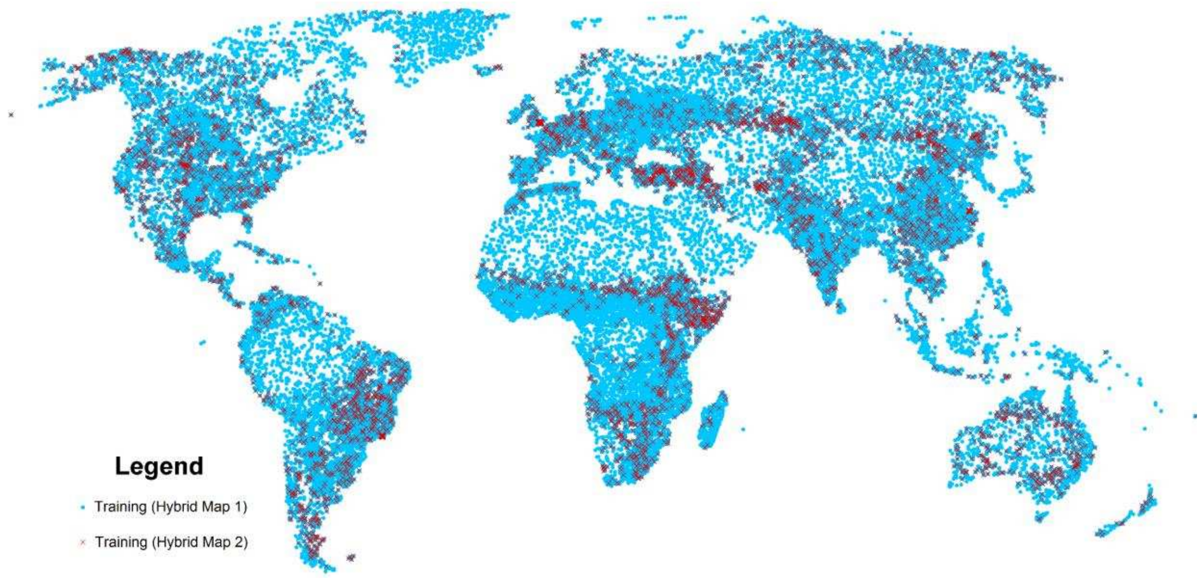

Figure 3: Distribution of the validation data (with the training data provided for reference)

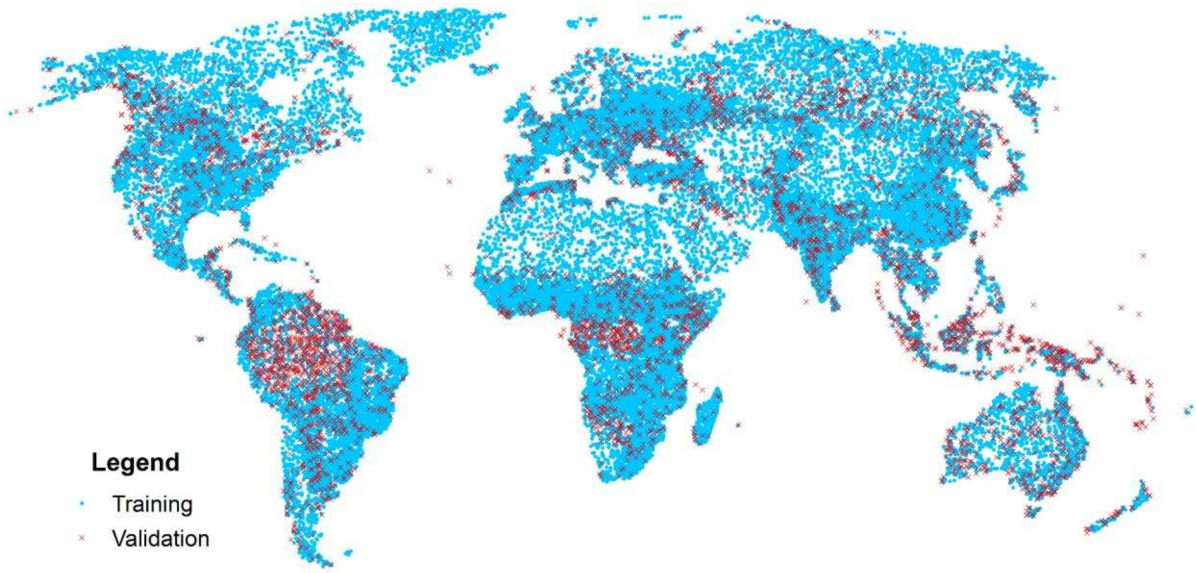

Figure 4: Area when one or another land cover product shows the best agreement with the Geo-Wiki global training dataset

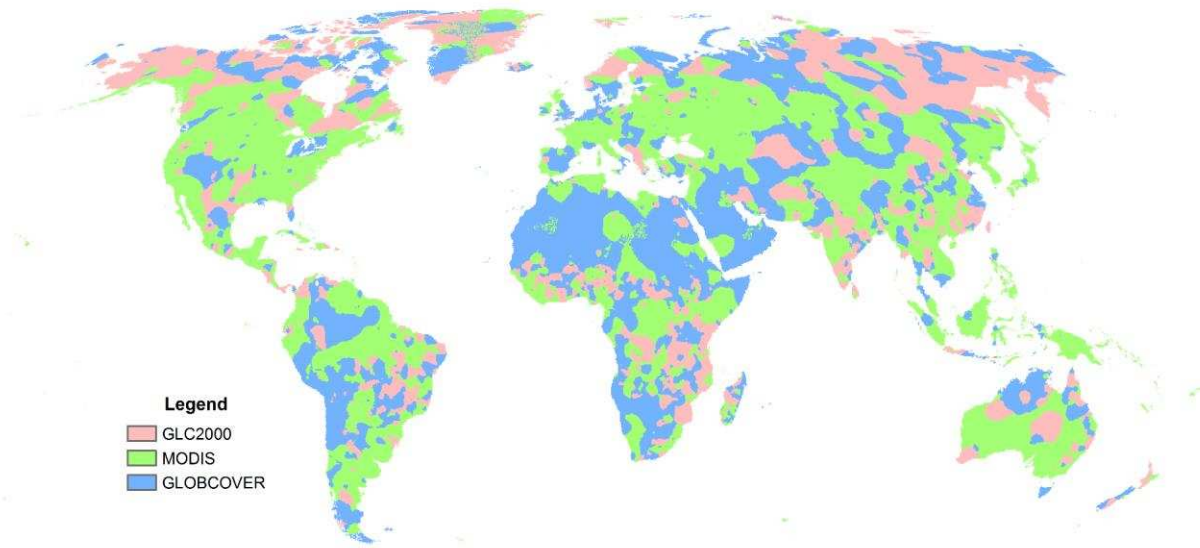

Figure 5: The hybrid land cover map based on applying GWR trained with Geo-Wiki training data from Human Impact competition 1

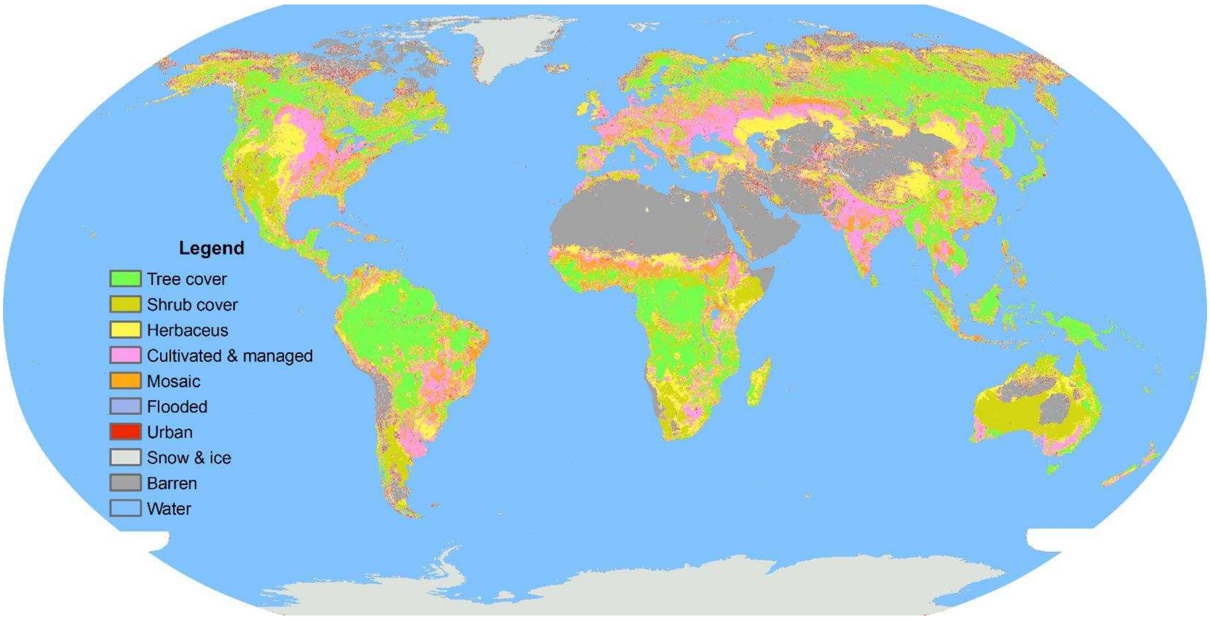

Figure 6: The hybrid land cover map based on Geo-Wiki training points from the Hotspot Map Disagreement competition 2