

Assessing the suitability of GlobeLand30 for mapping land cover in Germany

Jamal Jokar Arsanjani^a, Linda See^b and Amin Tayyebi^c

^aGIScience Research Group, Institute of Geography, Heidelberg University, Heidelberg, Germany;^bEcosystems Services and Management Program (ESM), International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria

^cUniversity of California-Riverside, College of Natural & Agricultural Sciences, Center for Conservation Biology, Riverside, CA, USA

Abstract

Global land cover maps have been widely employed as the base layer for a number of applications including climate change, food security, water quality, biodiversity, change detection and environmental planning. Due to the importance of land cover, there is a pressing need to increase the temporal and spatial resolution of global land cover maps. A recent advance in this direction has been the GlobeLand30 dataset derived from Landsat imagery, which has been developed by the National Geomatics Center of China (NGCC). Although overall accuracy is greater than 80%, the NGCC would like help in assessing the accuracy of the product in different regions of the world. To assist in this process, this study compares the GlobeLand30 product with existing public and online datasets, i.e. CORINE, Urban Atlas, OpenStreetMap and ATKIS for Germany in order to assess overall and per class agreement. The results of the analysis reveal high agreement of up to 92% between these datasets and GlobeLand30 but that large disagreements for certain classes are evident, in particular wetlands. However, overall, GlobeLand30 is shown to be a useful product for characterizing land cover in Germany, and paves the way for further regional and national validation efforts.

Keywords: Land cover/use mapping, Germany, GlobeLand30, CORINE, Global land cover maps, Accuracy assessment

1 Introduction

Over the last two centuries, humans have made considerable changes to their landscape. This includes unprecedented land surface changes such as increasing urbanization (Tian et al., 2011; Tayyebi et al., 2013; Pijanowski et al., 2014) and deforestation (Mas, 2004; Pineda Jaimes et al., 2010; Tayyebi et al., 2015). Areas of current scientific concern include understanding the global impacts of land use / land cover (LULC) changes on multiple ecosystem services (Meehan et al. 2013; Vaz 2016), analyzing landscape changes at the global scale (Tayyebi et al., 2014) and improving estimates of greenhouse gas emissions from the LULUCF (land use, land-use change and forestry) sectors (Schlamadinger et al. 2007). Information to support these analyses is extracted from LULC maps, which highlights their importance (Sexton et al. 2013). A number of land cover products covering different time periods and different spatial resolutions have been created from remote sensing, e.g. MODIS (Friedl et al. 2002), Africover and GLC-SHARE (Latham et al. 2014), GLC2000 (Fritz et al. 2003), IGBP (Loveland et al. 2000) and GlobCover (Arino et al. 2012). Many of these products are based on coarse resolution sensors, e.g. GLC2000 is at 1km, MODIS at 500 m and GlobCover is at a 300m resolution (Grekousis et al. 2015; Congalton et al. 2014). The spatial resolution of these products is not sufficient for detailed studies, their accuracy is not high enough for many applications (see e.g. Fritz et al. 2011; Batisani & Yarnal 2009), and change detection is not possible. For LULC change simulation and land policy resolutions (Haase & Nuisl 2007; Jokar Arsanjani, Helbich & Vaz 2013), more accurate, higher resolution products are needed. As noted

from recent land change modelling experiences, namely Tayyebi et al. (2011) and Arsanjani et al. (2014), while exploring the magnitude and type of land change such as expansion, shrinkage or intensification, the positional accuracy of temporal land cover maps cause major sensitivities across modelling approaches. Urbanization will see the population of many cities in the developing world grow (United Nations, Department of Economic and Social Affairs 2014), while shrinking cities in parts of the developed world are of particular economic concern, especially when changes are persistent and accompanied by an aging population (The Economist 2015b). For example, more than a third of German cities are currently shrinking and the number continues to grow (The Economist 2015a). Land cover can play an important role in modelling future changes as well as development of urban and regional policy.

At the European scale, extensive use of the European wide Corine Land Cover (CLC) map is made by landscape planners, ecologists, researchers, and policy makers (Pérez-hoyos et al. 2012; Pekkarinen et al. 2009; Neumann et al. 2007; Grekousis et al. 2015). The CLC was produced for four years: 1990, 2000, 2006 and more recently for 2012 but CLC2012 is still in the validation phase and therefore not yet available. Moreover, the product is provided at a resolution of 100 and 250 m, so there is a lack of detail, particularly in urban environments (Diaz-Pacheco & Gutiérrez 2013; Feranec et al. 2010). More recently, the National Geomatics Center of China (NGCC) released GlobeLand30 as a global land cover dataset at a 30m resolution for 2010 (see Section 1.2.1); this was possible because of the opening up of the Landsat archive (Giri et al. 2013). This product may address some of the limitations of CLC products including improved spatial resolution and additional temporal coverage.

The purpose of this study is to compare GlobeLand30 with existing authoritative and crowdsourced LC products for Germany in order to evaluate their agreement. In doing so, we will contribute to the validation of this new high resolution LC product and gain a better understanding of the potential contribution of this dataset to regional and global LC mapping. In the next section, the different LC products used in this study are described in more detail, after which the methodology for comparing the products is outlined. The results of the comparison are presented using measures of agreement and maps of agreement/disagreement.

2 Current sources of land cover information for Europe

Currently, there are several land cover inventories that cover the European continent. These products are generated by various institutions for specific purposes. Each of these datasets has its own strengths and limitations. A review of available global and regional land cover products along with their specifications has been provided recently by Grekousis et al. (2015).

2.1 CORINE Land Cover

Work on the CLC inventory began in 1985 to develop a product for 1990 as the reference year. Since then, several updates have been produced, for example, in 2000, 2006, and the latest 2012 update, which is still in the validation phase and has not yet been released. The CLC nomenclature includes 44 land classes with a minimum mapping unit (MMU) of 0.25 km² for areal objects and a minimum width of 100 m for linear objects (Büttner et al. 2002). The data provider, i.e., the European Environment Agency (EEA), has also generated a time series of land change maps with an MMU of 0.05 km² at a finer resolution than the base layers. In the production process, the EEA coordinates, integrates, and consolidates the national products from the National Reference Centres within EIONET (European Environment Information and Observation Network). At a national level, visual interpretation of high resolution satellite imagery was used to generate maps in the majority of countries while semi-automatic

approaches were used in the rest, with national in-situ data, satellite image processing, and GIS integration/generalization (Diaz-Pacheco & Gutiérrez 2013). This product has been extensively used for a wide variety of applications such as population mapping (Gallego 2010), environmental protection (Feranec et al. 2007), air pollution studies (Janssen et al. 2008), and spatial planning and decision support systems (Koschke et al. 2012). The products are generated in both vector and raster format at resolutions of 100m and 250m.

2.1.2 Global Monitoring for Environment and Security Urban Atlas: GMESUA

While CLC is widely used as a land cover dataset, GMESUA offers pan-European comparable land-use datasets for large urban areas above 100,000 inhabitants. To date, 305 urban regions within Europe have been mapped within this initiative and the classes reflect the monitoring needs of European cities. Earth Observation (EO) data are supported by other reference data, which are used in the classification process, e.g. commercial-off-the-shelf (COTS) navigation data and topographic maps. Ancillary data including (a) COTS navigation data such as points of interest (POIs), land use, land cover, and water bodies; (b) Google Earth for visual checking and interpretation; (c) local city maps for certain LULC classes; (d) local zoning and cadastral data; (e) on-site visits; and (f) high-resolution aerial imagery finer than 1 m resolution were used to verify the locational and thematic accuracy of the classification (European Union 2011; Jokar Arsanjani, Helbich, Bakillah, et al. 2013). The dataset has an MMU of 0.0025–0.01 km², and a minimum width of linear elements of 100 m with ± 5 m positional accuracy (European Union 2011; Seifert 2009) For more details, see the Urban Atlas mapping guide (European Union 2011).

2.1.3 Collaborative Mapping via Citizens

Along with the advancement of semi/fully automatic signal processing approaches using remote sensing data for mapping LULC (Mayaux et al. 2006; Neumann et al. 2007), citizen observatories have recently shown their competence in collecting LULC information (Fowler et al. 2013; Jokar Arsanjani & Vaz 2015). Due to the advancement of Web 2.0 technologies, citizen observations have been well organized, stored, and maintained through collaborative mapping projects (CMPs: (Rouse et al. 2007; Vaz & Jokar Arsanjani 2015)), whereby these observations have been made available to the public for downloading, revising and for commenting upon. This approach has helped to collect citizen observations not only in a raw form, but also updated and corrected by other citizens, who wish to contribute to these collections (Jokar Arsanjani, Mooney, Helbich, et al. 2015). These platforms provide citizens with user-friendly functionalities so that they can a) either collect information offline by GPS-enabled devices and then upload them to the platform, or b) allow them to map online based on visual interpretation of very high resolution satellite/aerial images (i.e. 40 cm resolution) embedded in the platforms. This type of information is named Volunteered Geographic Information (VGI: (Goodchild 2007)) and this process is called crowdsourcing (Heipke 2010). Studies show that these platforms have been quite successful in terms of attracting a large number of active users as well as collecting trustworthy contributions. Some of these platforms are even GIS-friendly, which facilitates further processing and exploitation (Sester et al. 2014). From a remote sensing viewpoint, visual interpretation of high resolution satellite images is a common way to collect data. There are numerous examples including OpenStreetMap (Jokar Arsanjani, Zipf, Mooney, et al. 2015a), Geo-Wiki (Fritz et al. 2012), and Wikiloc (Castelein et al. 2010). Geo-Wiki, in particular, was designed to help improve global land cover maps through volunteer mappers (See et al. 2015).

OpenStreetMap (OSM), with over 2.2 million registered users and billions of contributed node features (by October 2015), has strengthened its role in creating a free Wikipedia of maps (Jokar Arsanjani, Zipf, Mooney, et al. 2015b). Various studies (Estima & Painho 2015; Jokar Arsanjani & Vaz 2015; Jokar Arsanjani, Mooney, Zipf, et al. 2015; Jokar Arsanjani, Helbich, Bakillah, et al. 2013; Kalantari & La 2015) have demonstrated the usefulness of OSM for mapping information about land cover and land use. The OSM contributions are organized into several layers, such as POIs, places, roads, railways, waterways, buildings, and natural and land-use features, which are rendered as points, polylines, and polygons. For the purpose of this study, data from OSM were used because the other sources of LULC data from the aforementioned applications are not as complete.

2.1.4 Authoritative repositories: case of ATKIS data

At a local level, national public authorities are responsible for providing official geodata. Additionally, private sector and commercial firms provide geodata at some costs for a certain number of licenses. In Germany, the federal surveying and cartographic authorities of the German Federal States are responsible for the production of land cover maps. The landscape map of 2010 called Digitales Basis-Landschaftsmodell (Basis-DLM) provided by the Bundesamt für Kartographie und Geodäsie (BKG; Arnold, 2009) from the Amtliches Topographisch-Kartographisches Informationssystem (ATKIS) is available at scales of 1:10,000 and 1:25,000. This product has an MMU of 0.001-0.01 km² and geometric accuracy of 3m. The traditional update cycle is 5 years, but objects with high relevance are updated on a 3, 6 or 12 monthly cycle depending on the object type (Becker et al. 2010).

2.1.5 Globeland30

The NGCC under the “Global Land Cover Mapping at Finer Resolution” project has recently generated a global land cover map named GlobeLand30. The dataset covers two timestamps of 2000 and 2010 through the collection and classification of more than 10,000 scenes, primarily acquired from Landsat Thematic Mapper (TM) and Enhanced TM plus (ETM+) sensors, which were then coupled/checked with some local products. A pixel-object-knowledge-based (POK-based) classification approach was used to produce GlobeLand30 (Han et al., 2014; Chen *et al.* 2014). The overall accuracy of this product is reported to be above 80% as outlined by Chen et al. (2015). To the best of our knowledge, there have been few studies that have analyzed the usefulness of this dataset. Exceptions include the study by Brovelli et al. (2015), who concluded that this product compared well with national authoritative datasets in Italy, and the research by Ban et al. (2015), who demonstrated the high accuracy of the water class for Scandinavian countries, in particular Sweden. Thus there is a great deal of scope for comparing GlobeLand30 with existing LULC products in other regions of the world.

2.1.6 Technical comparisons between the existing datasets

In this section, a summary of the five datasets used in this study is provided in Table 1. To be more specific, a comparison between them in terms of MMU, linear width, nomenclature richness (i.e., number of classes), temporal coverage and update period, spatial coverage (globally and in Germany), and positional accuracy is detailed. From an MMU viewpoint, ATKIS and CORINE account for the best and worst examples for capturing MMU and positional accuracy; however, OSM also provides very fine geometric information as outlined in Touya & Reimer (2015). GlobeLand30 is potentially a better data source of land cover information than CORINE in terms of MMU and positional accuracy. However, from a temporal coverage perspective, a wider range of time is covered by CORINE, while GlobeLand30 covers only two timestamps of 2000 and 2010 so far. In terms of nomenclature, ATKIS and OSM offer a wider

range of land classes (>62 classes), while GlobeLand30 lacks diversity in land classes. This factor limits the use of GlobeLand30 for detailed land cover characterization but makes it suitable for applications that require only ten broad land cover classes. However, GlobeLand30 spatially covers the whole Earth, which is a great data source for land cover assessment at a global scale. Similarly, OSM covers the whole globe although the contributions across space vary significantly. Therefore, GlobeLand30 and OSM seem to be good datasets for monitoring global landscapes at a better spatial resolution than other current global land cover maps although further assessments of GlobeLand30 regionally and nationally are still required (Chen et al. 2015; Jokar Arsanjani & Vaz 2015).

Table 1: A technical comparison between several land cover datasets

	MMU (km ²)	Linear width	Nomenclature richness	Temporal coverage and update period	Spatial coverage	Positional accuracy	Spatial coverage in Germany
CORINE	0.25	100m	44 classes	1990, 2000, 2006, 2012	Europe	100 m	Complete
GlobeLand30	0.008-0.09	90m	10 classes	2000, 2010	Global	±75 m	Complete
Urban Atlas	0.0025–0.01	10m	44 classes	2006, 2012*	Europe (partially)	5m	Major urban areas with a population > 100K
ATKIS	0.001-0.01	undefined	>62 classes	Regular updates (latest 2012)	National (Germany)	±3 m	Complete (0.5% in our study)
OpenStreetMap	undefined	undefined	63 classes (+ user defined)	Heterogeneous **	Global		25.6%

* Version 2012 of this product is scheduled for release in 2015.

** OSM gets updated continuously, and by looking at the history of each feature, its temporal coverage can be identified.

3 Materials and methods

3.1 Study site and data

The area selected for study is the country of Germany due to availability of different data sources for comparison as well as the diversity of its urban and rural landscapes. Germany covers an area of over 357,000 km² with a population of greater than 80 million (The Federal Statistical Office 2011). The five data sources used in this study are shown in Figure 1, which are GMESUA, GlobeLand30, CORINE, OSM and ATKIS as described in the previous section. While GlobeLand30 and CORINE cover the whole country, OSM, ATKIS, and UA have only partial coverage as document in Table 1. The OSM dataset was downloaded on February 5, 2014. Features tagged with “*Land-use*” and “*Natural*” represent LULC information.

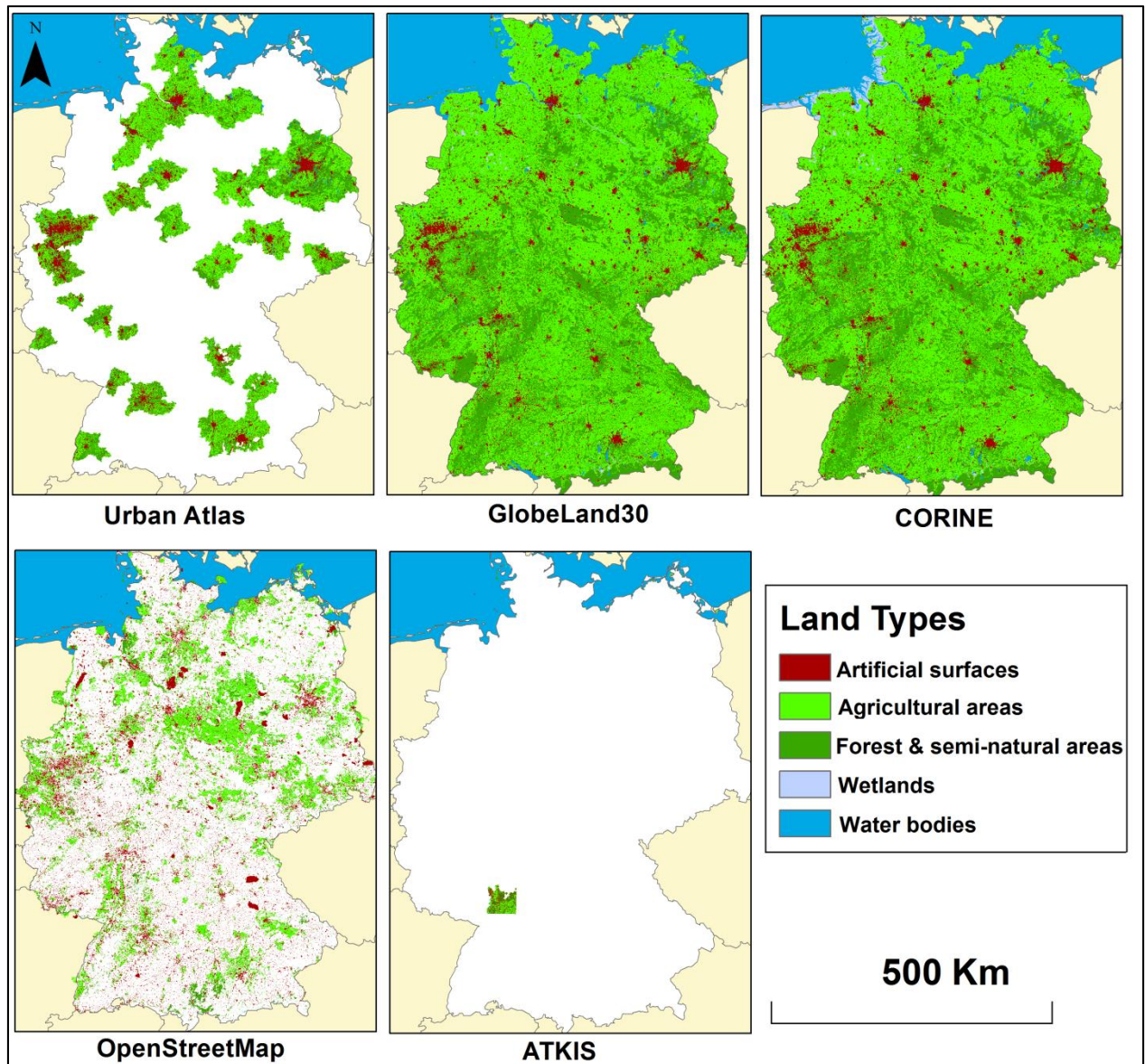


Figure 1: Different input datasets for Germany, including the Urban Atlas, GlobeLand30, CORINE, OpenStreetMap and ATKIS

3.2 Methods

3.2.1 Pre-processing of the input data

Figure 2 highlights the main tasks applied within this study. Prior to conducting the comparison between map products, they were first pre-processed, i.e. merged, mosaicked and harmonized due to heterogeneities in data formats, nomenclatures and projection systems. The GlobeLand30 image tiles were downloaded from www.globallandcover.com and mosaicked together. The vector data from the UA, OSM, and ATKIS were collected from their respective providers and merged separately. These five datasets were then resampled to match the 30m resolution of GlobeLand30, and the classes from each product were subsequently mapped onto CORINE level 1 nomenclature in order to facilitate the

comparison between products. These include five main land cover types: artificial surfaces, agricultural areas, forests and semi natural areas, wetlands, and water bodies. The reason for choosing this nomenclature is that it is widely used at a regional level, i.e., European, and the use of these generic high level classes meant that harmonization was more straightforward. Finally, the harmonized land cover maps were compared against one another as outlined in section 3.2.2 below.

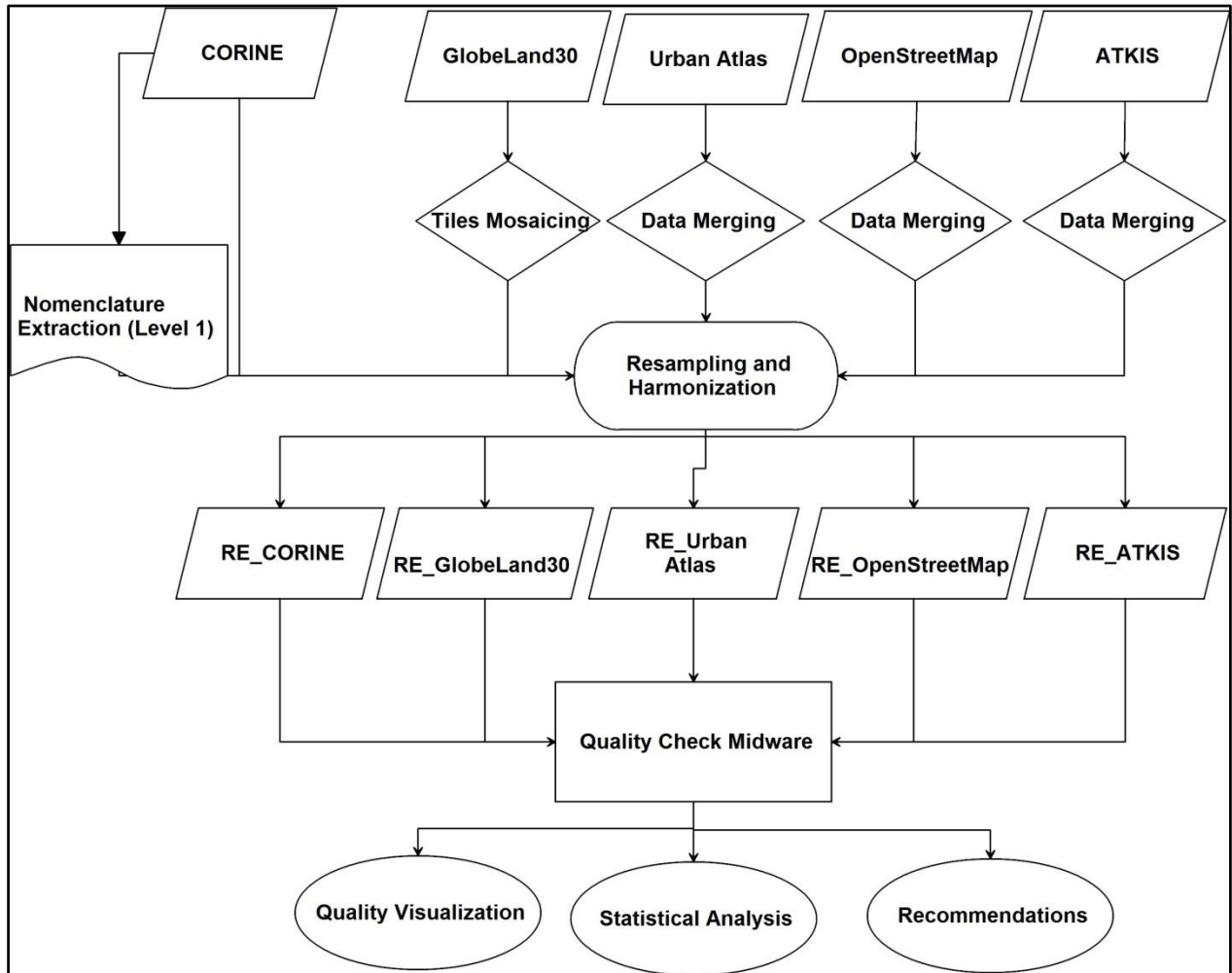


Figure 2: Flowchart of the workflow for map comparisons

3.2.2 Calculation of agreement

Once the pre-processing of the target datasets was carried out, they were imported into quality checking middleware, which calculates a confusion matrix, overall agreement, user's and producer's accuracy, and Kappa coefficient measures. Maps that represent agreement and disagreement between each pair of datasets were then generated. Finally, based on the computed indicators as well as the agreement/disagreement maps, recommendations were drawn regarding the suitability of GlobeLand30 for Germany. Although the UA and ATKIS datasets only cover part of the country, comparison was made on the overlapping areas.

4 Results and discussion

This section presents the results of the comparison between the GlobeLand30 dataset and the other available European land cover datasets for Germany.

4.1 GlobeLand30 compared with CORINE

One of the key differences between the two datasets is in their nomenclatures. While CORINE has 44, 15, and 5 different land cover classes at the third, second, and first levels of classification, respectively, GlobeLand30 has only 10 major land cover classes with no further hierarchy. Therefore, CORINE is much more thematically detailed than GlobeLand30. In contrast, the MMU of GlobeLand30 is 0.0009 km², which is finer than the 0.01 km² MMU of CORINE and may therefore pick up finer resolution features than CORINE. Table 2 provides the calculated confusion matrix as well as the measures of agreement between the two products where CORINE is used as the reference. The overall agreement is high at 92.5% with a Kappa index of 87.7%. Looking at the user's accuracies for each land cover type, all are above 80% except for wetlands, which shows a poor correspondence between the wetland class in GlobeLand30 and CORINE at 27.6%. The producer's accuracies are similarly high although wetland is once again lower at 70%.

Table 2: Confusion matrix and agreement values for CORINE (as reference) and GlobeLand30 (as source)

		GlobeLand30 (1000 m ²)						Producer's accuracy
		Artificial surfaces	Agricultural areas	Forests and semi-natural areas	Wetlands	Water bodies	Total	
CORINE	Artificial surfaces	21,441,078,675	4,254,484,942	793,657,354	6,171,169	73,045,450	26,568,437,589	80.70
	Agricultural areas	3,716,604,868	180,639,404,283	7,043,498,184	259,278,878	409,603,827	192,068,390,039	94.05
	Forests and semi-natural areas	1,189,987,853	4,731,908,701	101,427,709,247	229,566,613	353,471,973	107,932,644,387	93.97
	Wetlands	19,480,538	265,227,684	94,844,910	1,091,298,028	87,228,386	1,558,079,545	70.04
	Water bodies	205,020,610	366,558,979	288,493,270	2,367,104,343	26,428,518,469	29,655,695,672	89.12
	Total	26,572,172,545	190,257,584,588	109,648,202,965	3,953,419,031	27,351,868,103	357,783,247,232	-
	User's accuracy	80.69	94.94	92.50	27.60	96.62	-	
	Overall accuracy	92.52						
	Kappa index (%)	87.74						

Figure 3 presents a comparison of the areas for the different LC classes in CORINE and GlobeLand30. Most areas were similar in size when compared although the difference for wetlands is comparatively higher, which confirms the accuracy findings above for this class.

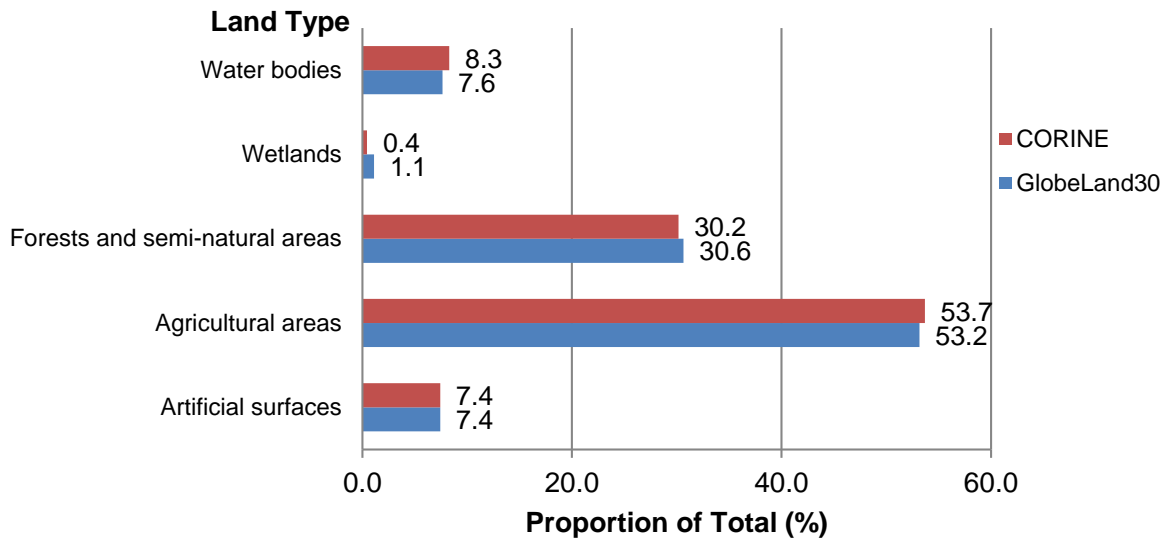


Figure 3: Area comparison across different land types between CORINE and GlobeLand30

Figure 4 displays the spatial distribution of agreement/disagreement in three areas of Germany. While there are very minor disagreements that are speckled across the country, three detailed views are provided to examine different types of mismatches. Case 1 shows a region close to the border with the Czech Republic on the eastern side of the country. Shown in transparent orange is an area that GlobeLand30 reports as wetlands, while CORINE reports these same areas as a mixture of agricultural areas and water bodies.

Case 2 shows an area on the western side of the country, where a mineral extraction site began development before 2006 (the time of the CORINE dataset) and was then further developed in 2010 (the date of GlobeLand30). While this area in CORINE is marked as artificial surfaces, it is mapped as forests and semi-natural areas in GlobeLand30. By looking at Google Earth historical images, the CORINE classification can be confirmed. Case 3 shows a delta in the north-west coast of Germany, which is classified as wetlands in CORINE and water bodies in GlobeLand30. Determining which one is correct is difficult to judge from the imagery alone. Looking at photographs from Panoramio, a water body can be seen, partly mixed with areas of clay and sand while from the satellite image it looks like land. Using the photographs as verification, it is most likely a water body.

Both cases 1 and 3 show examples of wetland disagreement, which is the one class that was poorly classified. Thus the disagreements could be the result of land change due to the time difference between CORINE and GlobeLand30, misclassification or disagreements on the edge of land polygons/pixels as outlined by Diaz-Pacheco and Gutiérrez (2013).

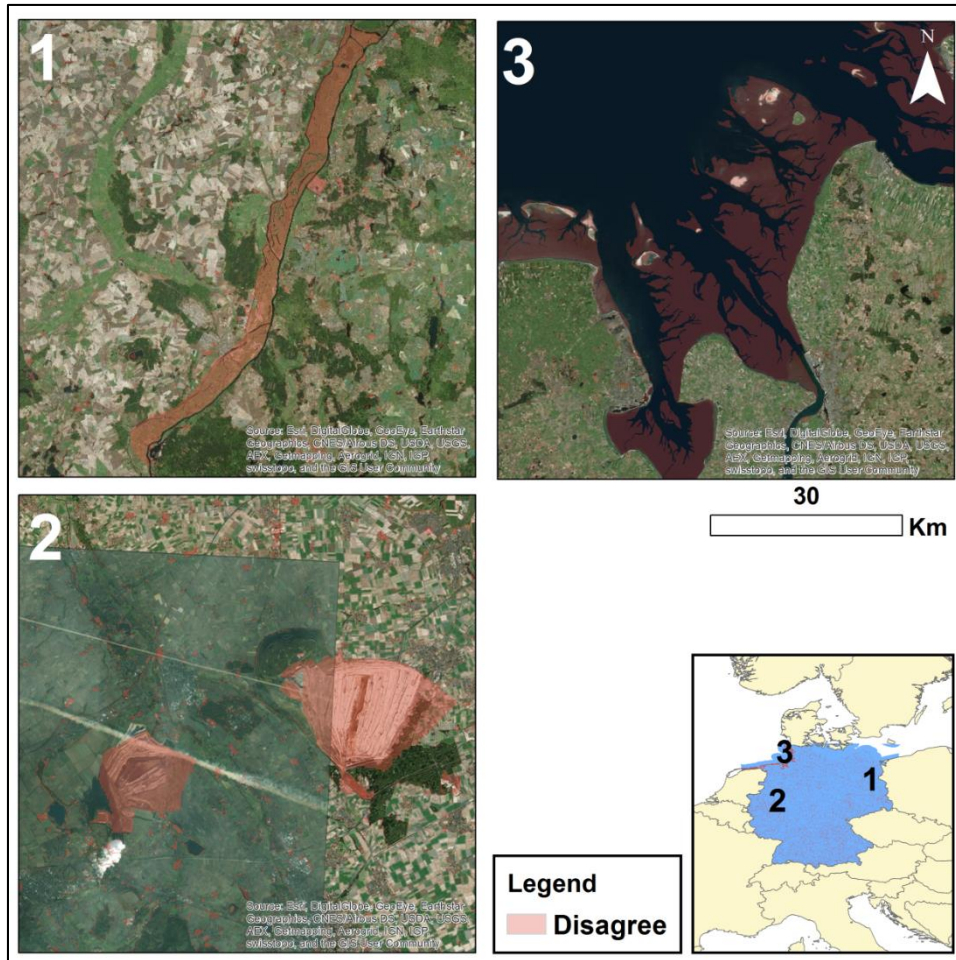


Figure 4: The spatial distribution of disagreement between GlobeLand30 and CORINE in three different places across Germany.

4.2 GlobeLand30 compared with Urban Atlas

In this section, we compare the UA with the GlobeLand30 product where the UA has the same nomenclature as CORINE so any issues related to harmonization may also apply here. Unlike CORINE, the MMU of the UA (0.0025-0.01 km²) is higher than GlobeLand30, which is 0.008-0.09 km². Table 3 provides the measures of agreement. The overall agreement is still high at 85.4% with a Kappa index of 75.0% but considerably lower than CORINE. User's accuracies are high for all land cover classes except for wetlands as in the previous comparison.

Table 3: Agreement derived from the confusion matrix between the Urban Atlas (as reference) and GlobeLand30 (as source)

Land Type	Producer's accuracy %	User's accuracy %	Overall agreement %	Kappa index%
Artificial surfaces	63.13	87.03	85.43	75.02
Agricultural areas	94.15	84.74		
Forests and semi-natural areas	82.98	87.15		
Wetlands	0.00	0.00		
Water bodies	71.28	88.21		

A visual examination with Google Earth imagery is used to better understand the issues related to the wetland class. Figure 5 (left) displays a snapshot from Google_Earth (year 2011), 6 geotagged photographs from Panoramio, which represent the reality from an oblique view angle, and the disputed area for this case (Figure 5 right). From our viewpoint, this disagreement is caused by different interpretations of the wetland class. While this area in UA is mapped as agricultural areas with some water bodies inside them, GlobeLand30 maps it as a wetland. This is a protected area in the north-west of Germany called Rehdener Geestmoor.

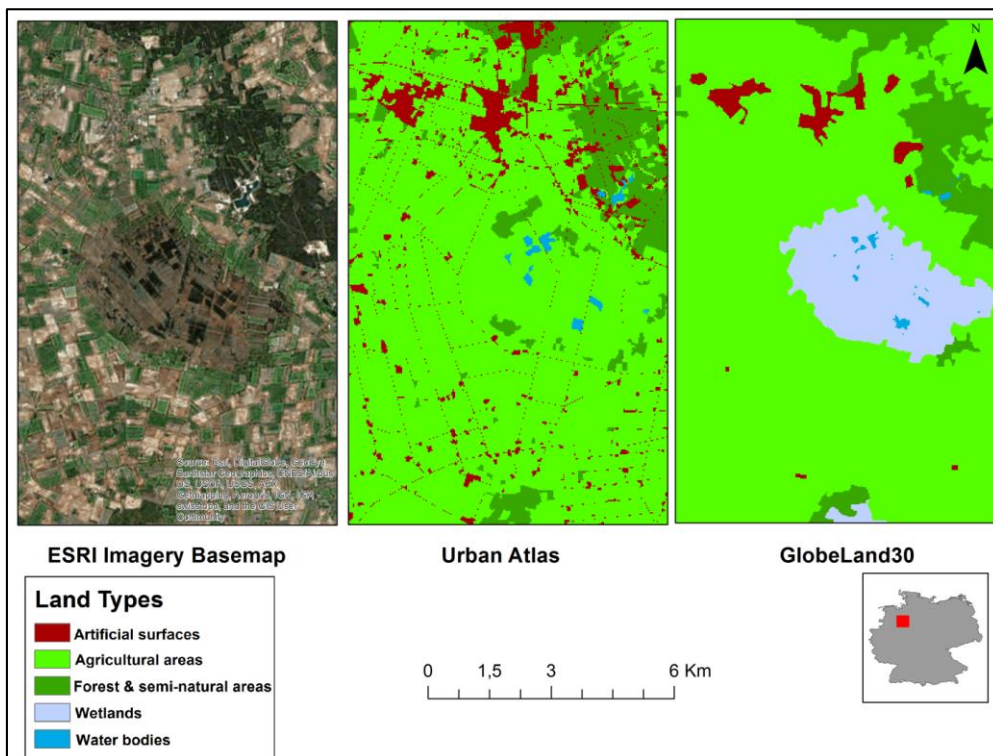
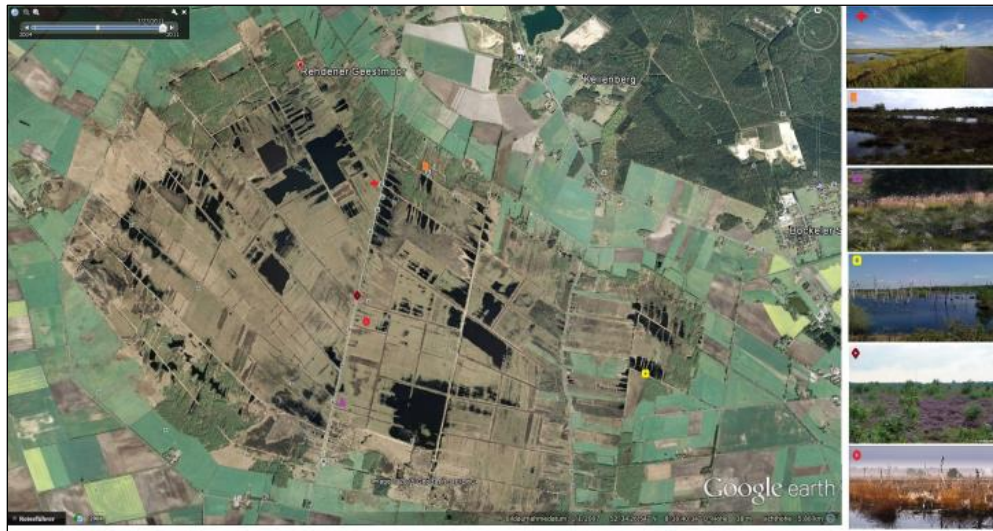


Figure 5: A snapshot of an area of disagreement from different sources and comparison with very high resolution satellite imagery (bottom and zoomed in at the top)

4.3 GlobeLand30 versus OSM

OSM features cover around 56% of the country with both urban and rural landscapes. As a dataset, OSM has a number of advantages including: a) a large range of land types, b) free form tagging, allowing users to add “user-defined” classes, and c) a high MMU although there are no official figures. Disadvantages include a) the need to manually translate features to global nomenclatures, which is difficult as no universal documentation is provided and b) in some places there are geometrical inconsistencies that need to be handled.

The quality analysis (Table 4) reveals an overall agreement of 74.2% and a Kappa index of 51.7%, which are lower values than the previous comparisons but still show substantial agreement (Landis & Koch 1977). However, the analysis at the class level exposes large disagreements except for agricultural areas and artificial surfaces with good to high producer’s and user’s accuracies.

Table 4: Accuracy values derived from confusion matrix between OpenStreetMap (as reference) and GlobeLand30 (as source)

Land Type	Producer’s accuracy %	User’s accuracy %	Overall accuracy %	Kappa index%
Artificial surfaces	94.12	63.53	74.24	51.70
Agricultural areas	72.63	96.62		
Forests and semi-natural areas	23.65	10.14		
Wetlands	0.00	0.00		
Water bodies	14.27	41.80		

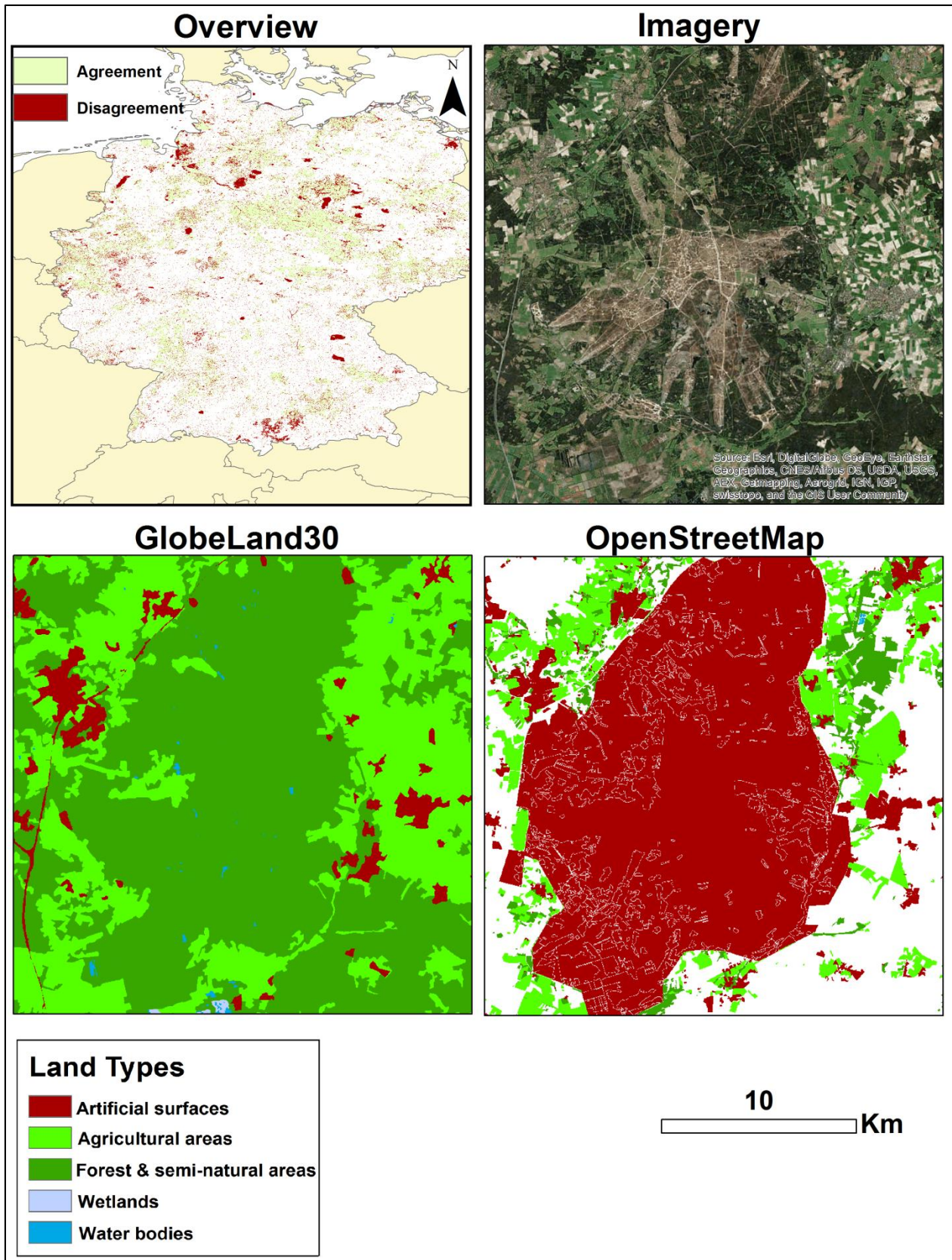


Figure 6: A visual comparison of GlobeLand30 and OSM

Figure 6 shows the agreement and disagreement between these two datasets in the top left hand corner. One polygon was then selected to illustrate the disagreement and is displayed in the bottom half of the figure, with a very high resolution satellite image for reference in the top right hand side. While OSM reports this polygon as the class artificial surfaces, GlobeLand30 reports it as forest and semi-natural areas with some agricultural areas. By looking at very high resolution satellite imagery on Google Earth and Panoramio photos, we found that this area is a military training base in which a forest area with a massive unpaved road network can be detected. This serves to illustrate the disagreement that arises from a mixture between LU and LC concepts. From a LC perspective, this area is a forest but when LU is considered, i.e. a military base, then artificial surfaces are more appropriate.

4.4 GlobeLand30 versus ATKIS

In this section, ATKIS, as the ultimate gold standard dataset, is used to evaluate the accuracy of GlobeLand30. The analysis was carried out only for the Rhine-Neckar region for which ATKIS data were available. In terms of nomenclature, ATKIS includes a wide variety of more than 35 land types, which is far more detailed than GlobeLand30. Furthermore, the MMU of ATKIS is 0.001-0.01 km², which is finer than GlobeLand30 and therefore allows smaller objects to be captured.

Overall accuracy and Kappa index between the two datasets are 85.2% and 76.8%, respectively (Table 5). While classes like artificial surfaces, agricultural areas, and forests and semi-natural areas show good agreement, water bodies show lower agreement. Once again the wetlands class is problematic and there is no agreement between the two datasets, similar to what occurred in the comparison with the UA.

Table 5: Agreement derived from the confusion matrix between ATKIS (as reference) and GlobeLand30 (as source)

Land Type	Producer's accuracy %	User's accuracy %	Overall accuracy %	Kappa index%
Artificial surfaces	71.93	83.81	85.23	76.82
Agricultural areas	92.32	81.30		
Forests and semi-natural areas	84.63	92.28		
Wetlands	0.00	0.00		
Water bodies	64.57	68.42		

The disagreement map shown in Figure 7 demonstrates that although disagreement areas are spread over the entire region, they are mostly concentrated along the roads and water bodies. This is due to the different MMUs of the two datasets. In Heidelberg, water bodies are not correctly reflected in GlobeLand30, which is due to the coarser spatial resolution of 30 m from Landsat images so is limited in delineating water bodies with smaller areas. Furthermore, some scattered artificial surfaces around the city are not reflected in GlobeLand30. The bottom three images in Figure 7 are from an area on the east side of Mannheim, which shows the same patterns of disagreement between the two sources. Furthermore, while a spot in the west side of the map shows forest and semi-natural areas, this area is, in reality, an artificial surface. A double check with the very high resolution image in the third column of Figure 7 confirms this.

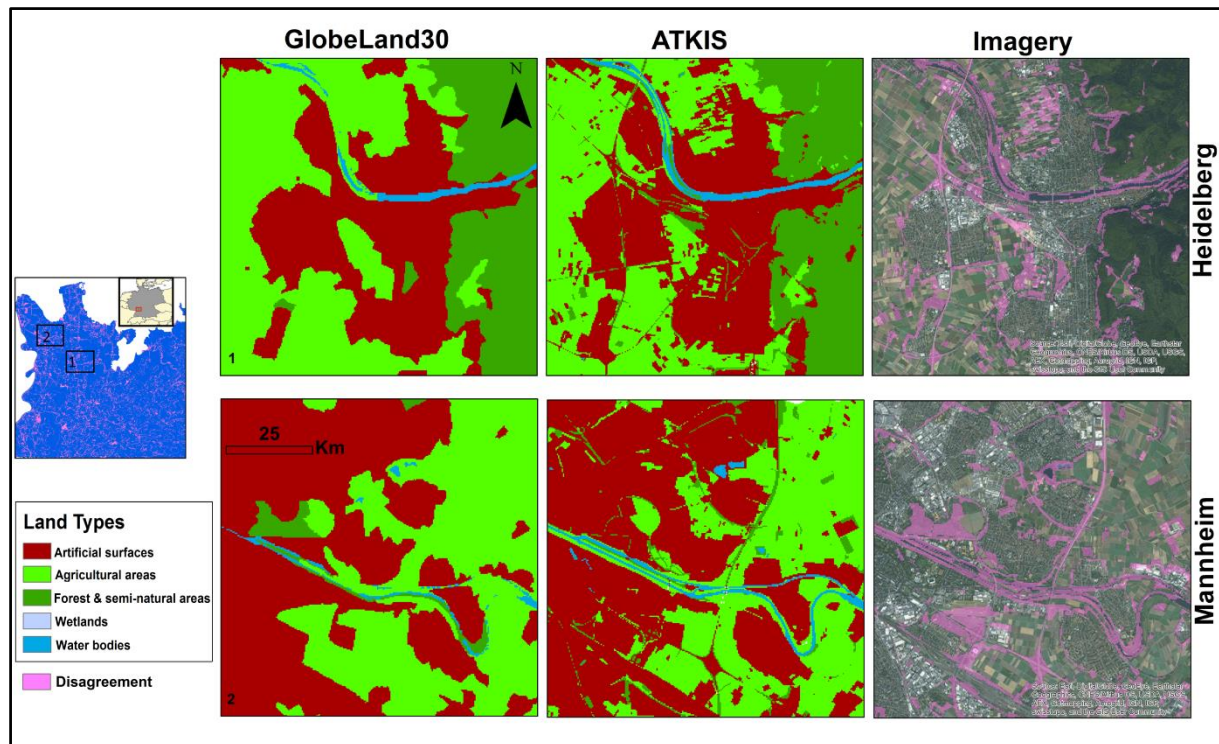


Figure 7: A comparison of GlobeLand30 and ATKIS, with more detailed examples for Mannheim (bottom) and Heidelberg (top)

5 Conclusions

GlobeLand30 is a recently produced global LC map at a resolution of 30m. Although the map has been validated by the NGCC as part of the map production process, independent validation against other regional and national sources of LULC has not yet been undertaken to any considerable extent. Thus, this study was an attempt to compare GlobeLand30 with existing datasets in Germany (i.e., CORINE, Urban Atlas, OpenStreetMap, and ATKIS).

In general, the reported overall agreements confirm good correspondence between the different datasets and GlobeLand30, ranging from 74% for OSM to 92% for CORINE. Among these datasets, ATKIS and UA, which have the highest MMU and positional accuracy and are the best reference sources to confirm the quality of GlobeLand30, had agreements >85%. Therefore, GlobeLand30 is a potentially useful product for capturing LC in Germany. The visualization of areas of agreement/disagreement reveals that the disagreements are spread across the entire region and are not concentrated around specific areas. This highlights the importance of looking at per class accuracies, i.e., user's accuracies. The per class analysis shows considerable disagreements for the wetlands class in all cases. Furthermore, agreement for the water bodies class is lower at 68%, which falls into "substantial" agreement according to Landis & Koch (1977). While water bodies are relatively simpler objects to map, differentiating between water bodies and wetland areas has resulted in areas of disagreement. The comparison between GlobeLand30 and OSM reports that only agricultural areas match perfectly, while the other classes do not show high agreement apart from a slightly better value for the class artificial surfaces. Although OSM is a unique dataset and using it as a reference dataset can be problematic, it would be interesting to better understand why these disagreements are occurring. To be more specific, are the

disagreements a) a result of recent land changes, which are not recorded in our databases, because OSM updating is continuous, or b) due to incorrect OSM contributions, or c) due to misinterpretation of land types by GlobeLand30, e.g. see Figure 6, or d) due to wrong translations of land types into the CLC nomenclature, since this step is fraught with difficulties. These issues need to be further explored within future studies as it is not the intention of this study to explore these here.

As mentioned earlier, the quality of each dataset can be externally explored by addressing the fitness of each product for a specific usage. Based on our findings for Germany, GlobeLand30 represents artificial surfaces, farming areas, forest and semi-natural areas, and water bodies well. Therefore, this product can be recommended for applications that require characterization of land cover and land cover analyses such as calculation of landscape metrics. Several issues might have contributed to the disagreements as follows: a) it can likely be due to the time differences between the datasets, b) it can be due to the different LC nomenclatures and problematic understanding/interpretation of them, c) uncertainty within each dataset, d) it can be caused by different data characteristics e.g., MMU, scale, resolution, and e) it could be due to data conversion as described earlier and also outlined by Brovelli et al. (2015). Each of these issues plays some role in data uncertainty.

Our conclusions draw attention to the value of GlobeLand30 as a useful product for comparison with other global LC maps, e.g. to highlight areas of disagreement and for integration into hybrid products, e.g. to continue improving global cropland maps (Fritz et al. 2015). This encouraging result paves the way for exploring the competence of this product in other continents. Once verified across multiple areas and climatic zones, it may be possible to use GlobeLand30 with confidence, particularly in developing countries where good land cover maps are difficult to find.

Acknowledgements

This research was partly supported by the ERC CrowdLand (No. 612755) project. The authors would also like to acknowledge the support and contribution of COST Action TD1202 'Mapping and the Citizen Sensor' as well as COST Action IC1203 'European Network Exploring Research into Geospatial Information Crowdsourcing' (ENERGIC).

References

- Arino, O. et al., 2012. Global Land Cover Map for 2009 (GlobCover 2009). Available at: <http://doi.pangaea.de/10.1594/PANGAEA.787668>.
- Arnold, S., 2009. Digital landscape model DLM-DE – deriving land cover information by integration of topographic reference data with remote sensing data. In *ISPRS Hannover Workshop 2009-4th Workshop on Land Use & Land Cover*. Available at: http://www.isprs.org/proceedings/XXXVIII-1-4-7_W5/paper/Arnold-167.pdf.
- Ban, Y. et al., 2015. Validation of the water layers of global land cover products using Geo-Wiki and national land cover maps. In *35th EARSeL Symposium 2015*. Stockholm, Sweden, p. Pending.
- Batisani, N. & Yarnal, B., 2009. Uncertainty awareness in urban sprawl simulations: Lessons from a small US metropolitan region. *Land Use Policy*, 26(2), pp.178–185. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0264837708000318> [Accessed March 9, 2012].

- Becker, C. et al., 2010. Semiautomatic Quality Control of Topographic Reference Datasets. *Isprs Tc Iv*, pp.2–7. Available at: http://ftp.tnt.uni-hannover.de/papers/data/871/2010_Helmholz_et_al_SemiautomaticQualityControlOfTopographicReferenceDatasets_01.pdf.
- Brovelli, M. et al., 2015. The First Comprehensive Accuracy Assessment of GlobeLand30 at a National Level: Methodology and Results. *Remote Sensing*, 7(4), pp.4191–4212. Available at: <http://www.mdpi.com/2072-4292/7/4/4191/>.
- Büttner, G., Feranec, J. & Gabriel, J., 2002. *Corine land cover update 2000*, Available at: <http://www.fomento.es/nr/rdonlyres/de88aec2-d5fa-4e95-9694-939a326cd026/3128/021219techrep89.pdf>.
- Castelein, W. et al., 2010. A characterization of Volunteered Geographic Information. In *13th AGILE International Conference on Geographic Information Science 2010, Guimarães, Portugal*. Guimarães, Portugal, pp. 1–10. Available at: http://agile2010.dsi.uminho.pt/pen/ShortPapers_PDF\106_DOC.pdf.
- Chen, J. et al., 2015. Global land cover mapping at 30m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 103, pp.7–27. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0924271614002275>.
- Congalton, R. et al., 2014. Global Land Cover Mapping: A Review and Uncertainty Analysis. *Remote Sensing*, 6(12), pp.12070–12093. Available at: <http://www.mdpi.com/2072-4292/6/12/12070/htm> [Accessed November 4, 2015].
- Diaz-Pacheco, J. & Gutiérrez, J., 2013. Exploring the limitations of CORINE Land Cover for monitoring urban land-use dynamics in metropolitan areas. *Journal of Land Use Science*, (February 2013), pp.1–17. Available at: <http://www.tandfonline.com/doi/abs/10.1080/1747423X.2012.761736>.
- Estima, J. & Painho, M., 2015. Investigating the Potential of OpenStreetMap for Land Use/Land Cover Production: A Case Study for Continental Portugal. In J. Jokar Arsanjani et al., eds. *OpenStreetMap in GIScience SE - 14*. Lecture Notes in Geoinformation and Cartography. Springer International Publishing, pp. 273–293. Available at: http://dx.doi.org/10.1007/978-3-319-14280-7_14.
- European Union, 2011. *Mapping Guide for a European Urban Atlas*,
- Feranec, J. et al., 2007. Corine land cover change detection in Europe (case studies of the Netherlands and Slovakia). *Land Use Policy*, 24(1), pp.234–247.
- Feranec, J. et al., 2010. Determining changes and flows in European landscapes 1990 – 2000 using CORINE land cover data. *Applied Geography*, 30(1), pp.19–35. Available at: <http://dx.doi.org/10.1016/j.apgeog.2009.07.003>.
- Fowler, A. et al., 2013. How Reliable are Citizen-Derived Scientific Data? Assessing the Quality of Contrail Observations Made by the General Public. *Transactions in GIS*, 17(4), pp.488–506. Available at: <http://doi.wiley.com/10.1111/tgis.12034> [Accessed October 9, 2013].

- Friedl, M.A. et al., 2002. Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, 83(1), pp.287–302.
- Fritz, S. et al., 2012. Geo-Wiki: An online platform for improving global land cover. *Environmental Modelling & Software*, 31(0), pp.110–123.
- Fritz, S. et al., 2003. *Harmonisation, mosaicing and production of the Global Land Cover 2000 database (Beta Version)*, Office for Official Publications of the European Communities Luxembourg.
- Fritz, S. et al., 2011. Highlighting continued uncertainty in global land cover maps for the user community. *Environmental Research Letters*, 6(4), p.044005.
- Fritz, S. et al., 2015. Mapping global cropland and field size. *Global change biology*, 21(5), pp.1980–1992.
- Gallego, F., 2010. A population density grid of the European Union. *Population and Environment*, 31(6), pp.460–473. Available at: <http://dx.doi.org/10.1007/s11111-010-0108-y>.
- Giri, C. et al., 2013. Next generation of global land cover characterization, mapping, and monitoring. *International Journal of Applied Earth Observation and Geoinformation*, 25(1), pp.30–37. Available at: <http://dx.doi.org/10.1016/j.jag.2013.03.005>.
- Goodchild, M.F., 2007. Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4), pp.211–221. Available at: <http://link.springer.com/10.1007/s10708-007-9111-y> [Accessed July 14, 2014].
- Grekousis, G., Mountrakis, G. & Kavouras, M., 2015. An overview of 21 global and 43 regional land-cover mapping products. *International Journal of Remote Sensing*, 36(21), pp.5309–5335. Available at: <http://dx.doi.org/10.1080/01431161.2015.1093195>.
- Haase, D. & Nuissl, H., 2007. Does urban sprawl drive changes in the water balance and policy? *Landscape and Urban Planning*, 80(1-2), pp.1–13. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0169204606001022> [Accessed June 16, 2012].
- Han, G. et al., 2014. A web-based system for supporting global land cover data production. *ISPRS Journal of Photogrammetry and Remote Sensing*, 103, pp.66–80. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0924271614001956>.
- Heipke, C., 2010. Crowdsourcing geospatial data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(6), pp.550–557. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0924271610000602> [Accessed March 16, 2012].
- Janssen, S. et al., 2008. Spatial interpolation of air pollution measurements using CORINE land cover data. *Atmospheric Environment*, 42(20), pp.4884–4903.
- Jokar Arsanjani, J., Mooney, P., Helbich, M., et al., 2015. An exploration of future patterns of the contributions to OpenStreetMap and development of a Contribution Index. *Transactions in GIS*, p.pending.

- Jokar Arsanjani, J., Zipf, A., Mooney, P., et al., 2015a. An Introduction to OpenStreetMap in Geographic Information Science: Experiences, Research, and Applications. In J. Jokar Arsanjani et al., eds. *OpenStreetMap in GIScience SE - 1*. Lecture Notes in Geoinformation and Cartography. Springer International Publishing, pp. 1–15. Available at: http://dx.doi.org/10.1007/978-3-319-14280-7_1.
- Jokar Arsanjani, J., Zipf, A., Mooney, P., et al., 2015b. An introduction to OpenStreetMap in GIScience: Experiences, Research, Applications. In J. Jokar Arsanjani et al., eds. *OpenStreetMap in GIScience: Experiences, Research, Applications*. Springer, pp. 1–20.
- Jokar Arsanjani, J., Mooney, P., Zipf, A., et al., 2015. Quality assessment of the contributed land use information from OpenStreetMap versus authoritative datasets. In J. Jokar Arsanjani et al., eds. *OpenStreetMap in GIScience: experiences, research, applications*. Springer.
- Jokar Arsanjani, J., Helbich, M., Bakillah, M., et al., 2013. Toward mapping land-use patterns from volunteered geographic information. *International Journal of Geographical Information Science*, 27(12), pp.2264–2278.
- Jokar Arsanjani, J., Helbich, M. & Mousivand, A.J., 2014. A Morphological Approach to Predicting Urban Expansion. *Transactions in GIS*, 18(2), pp.219–233. Available at: <http://doi.wiley.com/10.1111/tgis.12031> [Accessed October 16, 2014].
- Jokar Arsanjani, J., Helbich, M. & Vaz, E., 2013. Spatiotemporal simulation of urban growth patterns using agent-based modeling: The case of Tehran. *Cities*, 32, pp.33–42.
- Jokar Arsanjani, J. & Vaz, E., 2015. An assessment of a collaborative mapping approach for exploring land use patterns for several European metropolises. *International Journal of Applied Earth Observation and Geoinformation*, 35(0), pp.329–337.
- Kalantari, M. & La, V., 2015. Assessing OpenStreetMap as an Open Property Map. In J. Jokar Arsanjani et al., eds. *OpenStreetMap in GIScience SE - 13*. Lecture Notes in Geoinformation and Cartography. Springer International Publishing, pp. 255–272. Available at: http://dx.doi.org/10.1007/978-3-319-14280-7_13.
- Koschke, L. et al., 2012. A multi-criteria approach for an integrated land-cover-based assessment of ecosystem services provision to support landscape planning. *Ecological Indicators*, 21, pp.54–66.
- Landis, J.R. & Koch, G.G., 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33(1), pp.159–174. Available at: <http://europemc.org/abstract/MED/843571>.
- Latham, J. et al., 2014. *Global Land Cover SHARE (GLC-SHARE) database Beta-Release Version 1.0-2014*,
- Loveland, T.R. et al., 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, 21(6-7), pp.1303–1330. Available at: <http://dx.doi.org/10.1080/014311600210191>.
- Mas, J., 2004. Modelling deforestation using GIS and artificial neural networks. *Environmental Modelling & Software*, 19(5), pp.461–471. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S1364815203001610> [Accessed March 5, 2012].

- Mayaux, P. et al., 2006. Validation of the Global Land Cover 2000 Map. , 44(7), pp.1728–1739.
- Meehan, T.D. et al., 2013. Ecosystem-Service Tradeoffs Associated with Switching from Annual to Perennial Energy Crops in Riparian Zones of the US Midwest. *PLoS ONE*, 8(11), p.e80093. Available at: <http://dx.doi.org/10.1371/journal.pone.0080093>.
- Neumann, K. et al., 2007. Comparative assessment of CORINE2000 and GLC2000 : Spatial analysis of land cover data for Europe. , 9, pp.425–437.
- Pekkarinen, A., Reithmaier, L. & Strobl, P., 2009. ISPRS Journal of Photogrammetry and Remote Sensing Pan-European forest / non-forest mapping with Landsat ETM + and CORINE Land Cover 2000 data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(2), pp.171–183. Available at: <http://dx.doi.org/10.1016/j.isprsjprs.2008.09.004>.
- Pérez-hoyos, A., García-haro, F.J. & San-miguel-ayanz, J., 2012. ISPRS Journal of Photogrammetry and Remote Sensing Conventional and fuzzy comparisons of large scale land cover products : Application to CORINE , GLC2000 , MODIS and GlobCover in Europe. , 74, pp.185–201.
- Pineda Jaimes, N.B. et al., 2010. Exploring the driving forces behind deforestation in the state of Mexico (Mexico) using geographically weighted regression. *Applied Geography*, 30(4), pp.576–591. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0143622810000561> [Accessed May 21, 2013].
- Rouse, L.J., Bergeron, S.J. & Harris, T.M., 2007. Participating in the Geospatial Web: Collaborative Mapping, Social Networks and Participatory GIS. In A. Scharl & K. Tochtermann, eds. *The Geospatial Web*. Advanced Information and Knowledge Processing. Springer London, pp. 153–158. Available at: http://dx.doi.org/10.1007/978-1-84628-827-2_14.
- Schlamadinger, B. et al., 2007. A synopsis of land use, land-use change and forestry (LULUCF) under the Kyoto Protocol and Marrakech Accords. *Environmental Science & Policy*, 10(4), pp.271–282. Available at: <http://www.sciencedirect.com/science/article/pii/S1462901107000184> [Accessed October 19, 2015].
- See, L. et al., 2015. Harnessing the power of volunteers, the internet and Google Earth to collect and validate global spatial information using Geo-Wiki. *Technological Forecasting and Social Change*, 98, pp.324–335. Available at: <http://www.sciencedirect.com/science/article/pii/S004016251500061X>.
- Seifert, F., 2009. Improving Urban Monitoring toward a European Urban Atlas. In *Global Mapping of Human Settlement*. Remote Sensing Applications Series. CRC Press. Available at: <http://dx.doi.org/10.1201/9781420083408-c11>.
- Sester, M. et al., 2014. Integrating and Generalising Volunteered Geographic Information. In D. Burghardt, C. Duchêne, & W. Mackaness, eds. *Abstracting Geographic Information in a Data Rich World SE - 5*. Lecture Notes in Geoinformation and Cartography. Springer International Publishing, pp. 119–155. Available at: http://dx.doi.org/10.1007/978-3-319-00203-3_5.
- Sexton, J.O. et al., 2013. Long-term land cover dynamics by multi-temporal classification across the

- Landsat-5 record. *Remote Sensing of Environment*, 128(0), pp.246–258. Available at: <http://www.sciencedirect.com/science/article/pii/S0034425712003926>.
- Tayyebi, A. et al., 2013. Hierarchical modeling of urban growth across the conterminous USA: developing meso-scale quantity drivers for the Land Transformation Model. *Journal of Land Use Science*, 8(4), pp.422–442. Available at: <http://www.tandfonline.com/doi/abs/10.1080/1747423X.2012.675364> [Accessed November 21, 2013].
- Tayyebi, A., Pijanowski, B.C. & Tayyebi, A.H., 2011. An urban growth boundary model using neural networks, GIS and radial parameterization: An application to Tehran, Iran. *Landscape and Urban Planning*, 100(1-2), pp.35–44. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0169204610002896> [Accessed March 4, 2012].
- The Economist. 2015a. “How to Shrink a City: Many Cities are Losing Inhabitants.” Better to manage decline than try to stop it. The Economist. <http://www.economist.com/news/leaders/21652343-many-cities-are-losing-inhabitantsbetter-manage-decline-try-stop-it-how-shrink>.
- The Economist. 2015b. “Rus in Urbe Redux: A Growing Number of Cities will have to Plan for Drastically Smaller Populations.” The Economist. <http://www.economist.com/news/international/21652314-growing-number-citieswill-have-plan-drastically-smaller-populations-rus>.
- The Federal Statistical Office. 2011. Census 2011: Population Statistics of Germany. https://www.destatis.de/DE/PresseService/Presse/Pressekonferenzen/2013/Zensus2011/bevoelkerung_zensus2011.pdf?__blob=publicationFile.
- Tian, G. et al., 2011. Simulating spatiotemporal dynamics of urbanization with multi-agent systems—A case study of the Phoenix metropolitan region, USA. *Ecological Modelling*, 222(5), pp.1129–1138. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0304380010006848> [Accessed March 2, 2012].
- Touya, G., and A. Reimer. 2015. “Inferring the Scale of OpenStreetMap Features.” In *OpenStreetMap in GIScience SE-5. Lecture Notes in Geoinformation and Cartography*, edited by J. J. Arsanjani, A. Zipf, P. Mooney, and M. Helbich, 81–99. Springer International Publishing. doi:10.1007/978-3-319-14280-7_5.
- United Nations, Department of Economic and Social Affairs, P.D. 2014. World Urbanization Prospects: The 2014 Revision, Highlights. <http://esa.un.org/unpd/wup/Highlights/WUP2014-Highlights.pdf>.
- Vaz, E., 2016. The future of landscapes and habitats: The regional science contribution to the understanding of geographical space. *Habitat International*, 51, pp.70–78. Available at: <http://www.sciencedirect.com/science/article/pii/S0197397515002106>.
- Vaz, E. & Jokar Arsanjani, J., 2015. Crowdsourced mapping of land use in urban dense environments: An assessment of Toronto. *The Canadian Geographer / Le Géographe canadien*, 0(0), p.n/a–n/a. Available at: <http://dx.doi.org/10.1111/cag.12170>.