

Technical Note

LACO-Wiki: A New Online Land Cover Validation Tool Demonstrated Using GlobeLand30 for Kenya

Linda See ^{1,*} , Juan Carlos Laso Bayas ¹ , Dmitry Schepaschenko ¹ , Christoph Perger ¹, Christopher Dresel ¹, Victor Maus ¹ , Carl Salk ^{1,2}, Juergen Weichselbaum ³, Myroslava Lesiv ¹ , Ian McCallum ¹, Inian Moorthy ¹ and Steffen Fritz ¹

- ¹ Ecosystems Services and Management Program, International Institute for Applied Systems Analysis (IIASA), A-2361 Laxenburg, Austria; lasobaya@iiasa.ac.at (J.C.L.B.); schepd@iiasa.ac.at (D.S.); pergerch@iiasa.ac.at (C.P.); dresel@iiasa.ac.at (C.D.); maus@iiasa.ac.at (V.M.); salk@iiasa.ac.at (C.S.); lesiv@iiasa.ac.at (M.L.); mccallum@iiasa.ac.at (I.M.); moorthy@iiasa.ac.at (I.M.); fritz@iiasa.ac.at (S.F.)
- ² Southern Swedish Forest Research Centre, Swedish University of Agricultural Sciences, SE-23053 Alnarp, Sweden; carl.salk@slu.se
- ³ GeoVille Information Systems GmbH, A-6020 Innsbruck, Austria; weichselbaum@geoville.com
- * Correspondence: see@iiasa.ac.at; Tel.: +43-2236-807-423

Academic Editors: Parth Sarathi Roy, James Campbell and Prasad S. Thenkabail
Received: 6 May 2017; Accepted: 19 July 2017; Published: 22 July 2017

Abstract: Accuracy assessment, also referred to as validation, is a key process in the workflow of developing a land cover map. To make this process open and transparent, we have developed a new online tool called LACO-Wiki, which encapsulates this process into a set of four simple steps including uploading a land cover map, creating a sample from the map, interpreting the sample with very high resolution satellite imagery and generating a report with accuracy measures. The aim of this paper is to present the main features of this new tool followed by an example of how it can be used for accuracy assessment of a land cover map. For the purpose of illustration, we have chosen GlobeLand30 for Kenya. Two different samples were interpreted by three individuals: one sample was provided by the GlobeLand30 team as part of their international efforts in validating GlobeLand30 with GEO (Group on Earth Observation) member states while a second sample was generated using LACO-Wiki. Using satellite imagery from Google Maps, Bing and Google Earth, the results show overall accuracies between 53% to 61%, which is lower than the global accuracy assessment of GlobeLand30 but may be reasonable given the complex landscapes found in Kenya. Statistical models were then fit to the data to determine what factors affect the agreement between the three interpreters such as the land cover class, the presence of very high resolution satellite imagery and the age of the image in relation to the baseline year for GlobeLand30 (2010). The results showed that all factors had a significant effect on the agreement.

Keywords: land cover; accuracy assessment; validation; statistical inference; GlobeLand30; sampling design; Kenya

1. Introduction

The relevance of global scale land cover is reflected in its status as an Essential Climate Variable (ECV) [1], providing Plant Functional Types (PFTs) to global climate models [2]. Land cover is also used in applications of ecosystem accounting, conservation, forest and water management, natural hazard prevention and mitigation, monitoring of agricultural policies and economic land use modeling [3–7]. Consequently, many global land cover maps have been produced over the last three decades, initially at a coarse resolution of 8 km [8] to medium resolutions of 300 m to 1 km [9–11] and more recently at a 30 m Landsat resolution [12,13], facilitated by the opening up of the Landsat archive [14] and

improvements in data processing and storage [15]. Hence, there has been an increasing trend in the development of higher resolution land cover and land use maps [16]. Taking advantage of the new Sentinel satellites in a multi-sensor approach, even higher resolution global products are starting to appear, e.g., the Global Urban Footprint layer, produced by the German Aerospace Agency (DLR) at the highest resolution of 12 m [17].

Even more important is the monitoring of land cover change over time, which is one of the largest drivers of global environmental change [18]. For example, the agriculture, forestry and other land use sector (ALOFU) contributes 24% to global greenhouse gas emissions [19]. Land cover change monitoring is also a key input to a number of the Sustainable Development Goals (SDGs) [20]. New land cover products have appeared recently that capture this temporal dimension, e.g., the ESA-CCI annual land cover products for 1992–2015 [21], the forest loss and gain maps of Hansen [22], the high resolution global surface water layers covering the period 1984–2015 [23] and the Global Human Settlement Layer for 1975, 1990, 2000 and 2014 for monitoring urbanization [24].

As part of the development chain of land cover and land cover change products, accuracy assessment is a key process that consists of three main steps: response design; sampling design; and analysis [25]. The response design outlines the details of what information is recorded, e.g., the land cover classes; the type of data to be collected, e.g., vector or raster; and the scale. The sampling design outlines the method by which the sample units are selected while analysis usually involves the estimation of a confusion matrix and a number of evaluation measures that are accompanied by confidence intervals, estimated using statistical inference [26]. Note that in this paper we use the terms accuracy assessment and validation interchangeably as these terms are used by GOF-C-GOLD (Global Observation for Forest Cover and Land Dynamics) and the CEOS Calibration/Validation Working Group (CEOS-CVWG), who have been key proponents in establishing good practice in the validation of global land cover maps as agreed by the international community [27]. Moreover, both accuracy assessment (e.g., [28–30]) and validation (e.g., [31–33]) have been used in the recent literature.

In general, the accuracy assessment of land cover products is undertaken using commercial tools or in-house technologies. Increasingly, there are open source tools becoming available, e.g., the RSToolbox package in R [34], a QGIS plug-in called “Validation Tool” [35] and spreadsheets that calculate accuracy measures based on a confusion matrix, e.g., [36]. There are also new open source remote sensing tools, e.g., SAGA [37], GRASS [38] and ILWIS [39], among others [40]. However, the use of these tools, and R packages for accuracy assessment, in particular, still requires a reasonable level of technical expertise, or in some cases they do not include the full validation workflow. Hence there is a clear need for a simple-to-use, open and online validation tool.

Another key recommendation from the GOF-C-GOLD CEOS-CVWG is the need to archive the reference data, making them available to the scientific community. This recommendation is also directly in line with the Group on Earth Observation’s (GEO) Quality Assurance Framework for Earth Observation (QA4EO), which includes transparency as a key principle [41]. Although confusion matrices are sometimes published, e.g., [9,10,42], the reference data themselves are often not shared. There are exceptions, which include: (i) the GOF-C-GOLD validation portal [43]; (ii) the global reference data set from USGS [44,45], designed by Boston University and GOF-C-GOLD [46]; (iii) a global validation data set developed by Peng Gong’s group at Tsinghua University [47]; and (iv) data collected through Geo-Wiki campaigns [48]. However, the amount of reference data that is not shared far outweighs the data currently being shared. If more data were shared, they could be reused for both calibration and validation purposes, which is one of the recommendations in the GOF-C-GOLD CEOS-CVWG guidelines [27]. Although there are issues related to inclusion of existing validation samples within a probability-based design and in terms of harmonization of nomenclatures, a recent study by Tsendbazar et al. [49] showed that existing global reference data sets do have some reuse potential.

The LACO-Wiki online tool is intended to fill both of these identified gaps, i.e., to provide an online platform for undertaking accuracy assessment and to share both the land cover maps and

the reference data sets generated as part of the accuracy assessment process. Hence, the aim of this paper is to present the LACO-Wiki tool and to demonstrate its use in the accuracy assessment of Globeland30 [13] for Kenya. During this process, we highlight the functionality of LACO-Wiki and provide lessons learned in land cover map validation using visual interpretation of satellite imagery.

2. The LACO-Wiki Tool

LACO-Wiki has been designed as a distributed, scalable system consisting of two main components (Figure 1): (i) the online portal, where the user interaction and initial communications take place; and (ii) the storage servers, where the data are handled, analyzed and published through a Web Map Service (WMS), among other functions. The communication between the portal and the storage servers is implemented using RESTful Web Services. The portal has been built using ASP.NET MVC while the storage is comprised of Windows services written in C#. OpenLayers is used as the mapping framework for displaying the base layers (i.e., Google, Bing, OpenStreetMap), the uploaded maps, the samples, and additional layers. In addition, PostgreSQL is used for the database on each of the components, GeoServer and its REST API are used for publishing the data WMS layers and styling them, the GDAL/OGR library is employed for raster and vector data access, and Hangfire handles job creation and background processing.

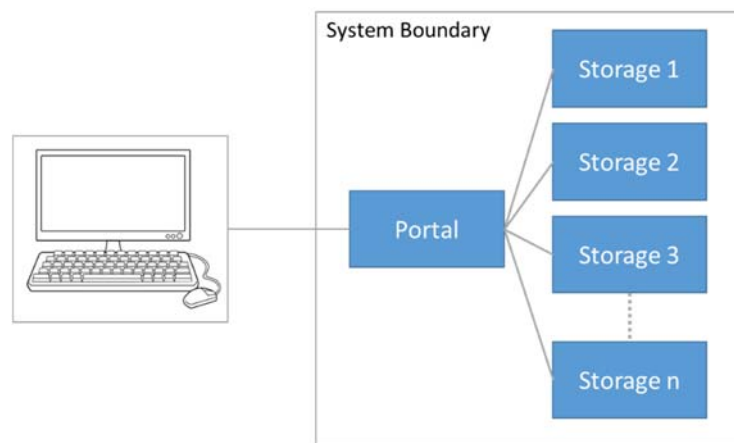


Figure 1. The LACO-Wiki system architecture. The client/user accesses the LACO-Wiki portal, which can make use of multiple storage servers to distribute the data and the processing tasks.

An important feature of LACO-Wiki is the data security and the access rights of the users. Since the WMS is not exposed to the public, this means that any maps and sample sets are stored securely in the system. Users can choose to share their data sets (i.e., the maps) and sample sets with individual users or openly to all users. However, downloads of the data sets, sample sets and the accuracy reports are only possible for those with proper access rights.

Note that these sharing features relate to the maps uploaded to LACO-Wiki and the intermediate products created in the system, e.g., a validation session, which can be shared with others to aid in the visual interpretation process. This is separate to the user agreement that governs the sharing of interpreted samples in the system. By using the platform, users agree to the terms of use (visible on the “About” page—<https://laco-wiki.net/en/About>), which state that as part of using LACO-Wiki, they agree to share their reference data. At present there is not much reference data available in the system, but in the future, any user will be able to access the data through a planned LACO-Wiki API (Application Programming Interface).

The LACO-Wiki portal can be accessed via <https://www.laco-wiki.net>. To log in, a user can either use their Geo-Wiki account after registering at <https://www.geo-wiki.org> or their Google or Facebook accounts, where the access has been implemented through Open Authentication (OAUTH2)

technology. The default language for LACO-Wiki is English but it is possible to change to one of 11 other languages: German, French, Spanish, Italian, Portuguese, Greek, Czech, Russian, Ukrainian, Turkish and Bulgarian.

2.1. The Validation Workflow

There are four main components to the validation workflow in LACO-Wiki (Figure 2). Each component is outlined in more detail below. These components are shown as menu items at the top of the LACO-Wiki screen (Figure 3).

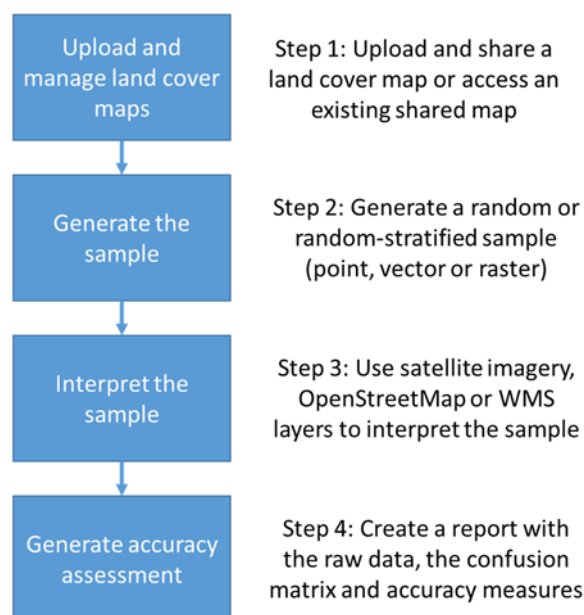


Figure 2. The components of the LACO-Wiki workflow.



Figure 3. The LACO-Wiki interface (<http://www.laco-wiki.net>) showing the four components of the validation workflow as menu items at the top of the screen.

The section that follows describes the validation workflow encapsulated in LACO-Wiki. This is followed by a description of how the tool can be used in different use cases.

2.1.1. Create and Manage Data

The first step in the workflow is to upload a data set, i.e., a land cover map, to LACO-Wiki (Figure 4). By clicking on the Data menu option (Figure 3), the “Manage your Datasets” screen will appear. All the data sets that the user has uploaded or which are shared with the user will appear in their list. The “Upload a new Dataset” option appears at the bottom of this screen. This will display the “Create a new Dataset” screen (Figure 4). Here, the user must enter a name for the data set and then indicate whether it is categorical, i.e., land cover classes, continuous, e.g., 0–100%, or a background layer that can be used additionally in the validation process such as an aerial photograph of the area.

The user must then enter a data set description. Finally, the user can select a raster or vector file in geotiff or shapefile format, respectively. At present the map must have a WGS 84 Platte Carrée projection EPSG: 32662 or the European projection ETRS89/ETRS-LAEA EPSG: 3035. The file size that can be uploaded is currently limited to 1 GB. More information on how to handle larger file sizes is presented in Section 2.2. The data set is then created and stored in the LACO-Wiki system.

Create a new Dataset

Here you can create a new dataset. Define a name, the land cover type and choose the corresponding files.

Dataset Information

Dataset Name: ✓

Land Cover Type: ✓

Categorical:
Categorical data, also called thematic, discrete, or classified data, are used both for vector and raster data to represent discrete information. Examples of categorical data are land cover or land use maps like pan-European CORINE land cover, Urban Atlas or regional LISA Land Cover as well as global CCI Land Cover, GlobCover etc.

Continuous:
Continuous data, representing phenomena such as percentage, elevation data or density such as the Copernicus high resolution layers of imperviousness degree and forest density or population (density) maps.

Background:
Background layers can be used in addition to Google and Bing maps base imagery in the validation step. Examples of such datasets would be an orthophoto of the study area.

Dataset Description: ✓

Files

Drag and Drop Files here
or

No files selected...

Supported file types: Shape (.shp, .shx, .dbf, .prj), GeoTiff (.tif, .tiff)
Supported SRS: WGS84, ETRS89 LAEA

Figure 4. Uploading a land cover map in Step 1 of the LACO-Wiki validation workflow.

By clicking on the data set that was just uploaded, basic metadata about the land cover map will be shown. At this stage the user must indicate the name of the column in the data set that contains the land cover classes or the continuous data. The legend designer is then used to change the default colors assigned to the classes. Some pre-defined legends from different land cover maps are also available to the user. Once the final colors have been chosen, the “Preview Image” window on the “Dataset Details” screen will show the map using these colors and the legend will be displayed below the map (Figure 5). Finally, the user can share the data set with other individual users, share the data with everyone or the data set can remain private to a user.

Dataset Details

[← Go to Dataset Overview](#)

Basic Information

Owner: juanlaso
 Dataset Name: Kenya Globcover30
 Dataset Type: Raster
 Land Cover Type: Categorical
 Uploaded: Friday, March 10, 2017 12:48 PM

Dataset Description

For training L.See

Raster Details

Raster Type: Integer
 Extent: [33.9204 -4.7207, 41.8987 4.6338]
 Projection Name: WGS84
 Extent WGS84: [33.9204 -4.7207, 41.8987 4.6338]
 Pixel Size X / Y: 0.00027778 / 0.00027778
 Rows X Columns: 33676 / 28722
 Min / Max: 10 / 90

Validation Samples based on this dataset

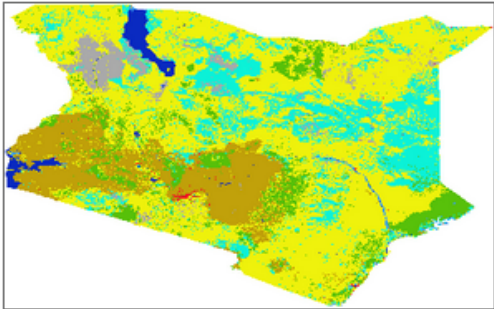
You haven't created any validation samples based on this dataset yet. You can define a new validation sample [here](#).

Sharing

Shared data sets can be used by other users to create a validation sample.

Everyone (Group)
 geolms (User)

Preview Image



Legend

- Cultivated land (10)
- Forest (20)
- Grassland (30)
- Shrubland (40)
- Wetland (50)
- Water bodies (60)
- Artificial surfaces (80)
- Bareland (90)
- Permanent snow and ice (100)

Figure 5. Details of a data set in LACO-Wiki including options to share and preview the data set and to generate a sample.

2.1.2. Generate a Sample

The second step in the workflow (Figure 2) is to generate a sample from a data set. There are two main types of sampling design available: random and stratified. Systematic sampling will be added in the future. In the case of a vector data set, this can be a point sample or a vector sample where there are two main choices by which objects are chosen, i.e., based on the number of objects or based on their area. In the case of a raster data set, this can be a point sample or a pixel sample. For a random sample, the user indicates the total sample size desired. For stratified, the number of sample units per class is entered. A sample calculator is currently being implemented that will allow users to determine the minimum number of samples needed. As with the data sets, samples can remain private or they can be shared. A sample can also be downloaded as a KML file for viewing in Google Earth or in a GIS package.

2.1.3. Generate a Validation Session and Then Interpret the Sample Using Imagery

Once a sample is generated, the third step in the workflow (Figure 2) is to create a validation session and then interpret the sample. Clicking on the Sample menu option at the top of the LACO-Wiki screen will bring up the “Manage your Samples” screen. This will list all sample sets created by the user or shared by other users. Clicking on one of the sample sets will select it and then basic metadata about that sample will be displayed (see Figure 6).

At the bottom of the screen there will be the option to create a validation session as shown in Figure 6. Figure 7 shows the screen for creating a validation session. In addition to the name and description, the user must select blind, plausibility or enhanced plausibility as the Validation Method. In blind validation, the interpreter does not know the value associated with the layer that they are validating and they choose one class from the legend when interpreting the satellite imagery. In plausibility validation, the interpreter is provided with the land cover class and they must agree or disagree with the value while enhanced plausibility validation allows the interpreter to provide a corrected class or value.

Sample Details

Go to Sample Overview

Basic Information

Owner: juanlaso
 Sample Name: Main sample Kenya GL30
 Associated Dataset: Kenya Globcover30
 Created: Tuesday, March 14, 2017 11:12 AM

Sample Description

Main sample for training at Kenya. L. See.

Composition Details

500x Stratified Pixel

Validation Sessions based on this sample

GLC30 Full Sample for Dima ([validate this validation session](#))
 GLC30 Full Sample for Juan ([validate this validation session](#))
 GLC30 Full Sample for Linda ([validate this validation session](#))
 Sample for Comparison ([validate this validation session](#))

[Add another validation session based on this sample here.](#)

Sharing

Shared samples can be used by other users to create a validation session.

Everyone (Group)
 geolms (User)

Sample Download

Download your sample locations.

ESRI Shapefile Sample as KMZ

Figure 6. Details about a sample in LACO-Wiki. This sample can be shared, downloaded or used to create a validation session.

Create your Validation exercise

Here you can define a validation exercise for your previously created samples.

Basic Settings:

Name:

Description:

Validation Method: Blind ▼

Blind:
The user has no information about the classification of the sample and therefore needs to provide a new interpretation based on a pre-defined list of classes.

Plausibility:
The user has information about the classification of the sample and evaluates the plausibility of the thematic classification with yes or no.

Enhanced Plausibility:
The user has information about the classification of the sample and evaluates the plausibility of the thematic classification with yes or no. In case the classification is incorrect, the user provides a new interpretation based on a pre-defined list of classes.

Validation Platform: Web ▼

Web:
Web platform validation sessions can be validated by the LACO-Wiki web application. This is the default use case.

Mobile:
Mobile platform validation sessions can be validated by the LACO-Wiki Mobile app. The LACO-Wiki Mobile app will be made public in the near future.

Validation Settings:

Comment field:

Geometry validation:

Geometry Validation:
Allows the validator to enter additional information about the quality of the geometry.

Min. Zoom: 6 (default) ▼

Max. Zoom: 18 (default) ▼

Sample Navigation: Feature Extent ▼

Layer Configuration:

Base layers:

Google Maps ▼ ✓

Overlays:

Validation Dataset ▼ Ⓞ

Figure 7. Creating a validation session in LACO-Wiki based on an existing sample. This involves providing a name, a description, choosing the type of validation (blind, plausibility, enhanced plausibility), whether this is for an online or mobile session, the validation settings, which include additional fields such as a comment box or whether the user should judge the positional accuracy, and what base layers should appear in the validation session, e.g., Google Maps, Bing Maps, etc.

The user must then indicate if the validation session is for the web, i.e., online through the LACO-Wiki interface, or for a mobile phone, i.e., through the LACO-Wiki mobile application for collection of data on the ground, which can complement the visual interpretation of the samples; geotagged photographs can also be collected. For the purpose of this paper, the validation session takes place on the web.

The user then chooses the background layers that will be available for the validation session. Google Earth imagery is always added as a default, but satellite imagery from Bing and an OpenStreetMap layer can also be added. Finally, other external layers can be added via a WMS, e.g., Sentinel 2 layers could be added from Sentinel Hub or any other WMS could be added, e.g., to display aerial photography.

The validation session is created and the same sharing options are available, i.e., private, shared with individuals or shared with everyone. If an individual validation session is shared with others, then the task can be divided among multiple persons, which is useful when the process is done in teams (e.g., with a group of students). Similarly, multiple copies of the validation session can be generated and shared with different individuals so that multiple interpretations are collected at the same location. This is useful for quality control purposes. Once the validation session has been created, a progress bar will be displayed. The validation session can then be started whereby a user would step through the individual sample units to complete the interpretation. Figure 8 shows an example of a pixel and the underlying imagery from Google Maps, which can be changed to Bing by clicking on the layer icon located on the top right of the image.

Interpret your Samples

Go to Validation Session Details

Sample Unit #: 500

Classify the object:

■	Cultivated land (10)
■	Forest (20)
■	Grassland (30)
■	Shrubland (40)
■	Wetland (50)
■	Water bodies (60)
■	Artificial surfaces (80)
■	Bareland (90)
■	Permanent snow and ice (100)

Sample Information

Sample Item ID 106036
Validated by
Timestamp -
Bing imagery date 02 Mar 2013 GMT

Download sample (kmz)

Progress in Total: 99.8 % (499 / 500)

Figure 8. Example of visual interpretation in LACO-Wiki, showing the final pixel in Sample 2 (see Section 3.3). The imagery can be changed to Bing or the pixel can be viewed directly in Google Earth by pressing the “Download sample (kmz)” button.

The dates of the Bing imagery are listed but to obtain the dates for Google Earth imagery, there is a “Download sample (kmz)” button that automatically displays the location on Google Earth if installed locally on the user’s computer. The validation session is completed once all sample units have been interpreted.

2.1.4. Generate a Report with the Accuracy Assessment

A report is produced in Excel format, which contains options for including the raw data, a confusion matrix and a series of accuracy measures. These accuracy measures are calculated using R routines, which are called by the LACO-Wiki application. At the most basic level, these include the overall accuracy and the class-specific user’s and producer’s accuracy [50]. Overall accuracy is the number of correctly identified classes, which is the diagonal of the confusion matrix, divided by the total of all elements in the confusion matrix. The user’s accuracy, also referred to as the error of commission, estimates the proportion of reference data that are correctly classified and hence indicates how well the map represents the truth. The producer’s accuracy, or error of omission, estimates the proportion of correctly classified reference locations in the map. Users can also choose to select Kappa, which was introduced to account for chance agreement and hence is generally lower than overall accuracy [50]. Note that we are not recommending this as a measure of accuracy, in line with the recommendations of the GOF-C-GOLD CEOS-CVWG [27] and other research [36], but we also recognize that many people will generate this accuracy measure regardless. As one of the use cases for LACO-Wiki is educational (see Section 2.3.2), learning about the advantages and pitfalls of different accuracy measures could be one use of this tool.

Finally, we offer other measures that have appeared in the literature, i.e., Average Mutual Information (AMI) [51], which measures consistency rather than correctness, and should therefore be used in combination with measures such as overall accuracy. It reflects the amount of information shared between the reference data and the map. Allocation and quantity disagreement [36] are two components of the total disagreement. Allocation disagreement represents the amount of difference between the reference data and the map that reflects a sub-optimal spatial allocation of the classes given the proportion of classes in the reference data set and the map while quantity disagreement captures the less than perfect match between the reference data and the map in terms of the overall quantity. Finally, Portmanteau accuracy includes both the presence and absence of a class in the evaluation and hence provides a measure of the probability that the reference data and the map are the same [52]. The reader is referred to the above references for further information on how these measures are calculated. Additional measures required by the user can be calculated from the confusion matrix. Alternatively, a request can be made to add accuracy measures to LACO-Wiki (contact info@laco-wiki.net).

2.2. Additional Features in LACO-Wiki

In addition to generating a sample set from a data set within LACO-Wiki, it is possible to generate it outside of LACO-Wiki, e.g., in a GIS package or using a programming language such as R. This is particularly useful in those situations where the land cover map is too large, i.e., >1 GB at the current limit, and can thus be generated externally. Once the sample set is generated, it can be uploaded to LACO-Wiki and it will then appear in the “Manage my Samples” list. A validation session can then be generated and interpreted as per Section 2.1.3. Users can also transform 32 bit rasters to 8 bit if they exceed the 1 GB limit or the area could be divided into regions prior to uploading as another way to handle the current file size limitation.

Another useful feature built-into LACO-Wiki is the sample calculator, although it is currently still being tested before release. This option will appear in Step 2 of the workflow, i.e., sample generation. This option allows users to determine the minimum number of samples needed given a specified precision.

2.3. Use Cases for LACO-Wiki

2.3.1. Scientific Research

There are many land cover maps created and published in the scientific literature every year at a range of scales from local to global. For example, in the study by Yu et al. [53], they assembled a database of 6771 research papers published before 2013 on land cover maps. LACO-Wiki could provide a research tool for accuracy assessment of land cover maps developed as part of ongoing research. As part of the research process, both the maps and the reference data can be shared, creating a potentially vast reference data repository. This in turn could stimulate further research into how these shared reference data could be reused in the calibration and validation of new land cover products. This also means that all data sets will have to comply with basic metadata standards.

2.3.2. Education

One of the main aims behind the development of LACO-Wiki has been its potential use as an educational tool. For example, it can be used within undergraduate remote sensing courses to provide an easy-to-use application for illustrating accuracy assessment as part of lessons on land cover classification. Students would most likely follow the complete workflow, i.e., upload a land cover map that they have created as part of learning about classification algorithms, generate a stratified-random sample, interpret the sample with imagery and then generate the accuracy measures. Lessons might also be focused on the accuracy assessment of a particular land cover product in which case the students might only be assigned a validation session followed by generation of the report on accuracy measures. This might provide an interesting classroom experiment on comparing the results achieved across the class.

2.3.3. Map Production

Map producers around the world could use the tool as one part of their production chain, which could vastly increase the pool of reference data in locations where the data are currently sparse. The use of local knowledge for image interpretation as well as in-house base layers such as orthophotos, which can be added to LACO-Wiki via a WMS, may result in higher quality reference data in these locations, which could, in turn, benefit other map producers when reusing the data.

2.3.4. Accuracy Assessment

LACO-Wiki is currently one of the tools recommended by the European Environment Agency (EEA) for use by EEA member countries in undertaking an accuracy assessment of the local component layers, e.g., the Urban Atlas. Starting in the summer of 2017, LACO-Wiki will be used for this purpose. The private sector could also use LACO-Wiki as part of land cover validation contracts. The sharing features built into LACO-Wiki facilitate completion of validation tasks by teams or can be used to gather multiple interpretations at the same location for quality control. Map users could also carry out independent accuracy assessments of existing products, geared towards their specific application or location.

3. Application of the LACO-Wiki Tool to the Accuracy Assessment of GlobeLand30 for Kenya

In this section, we use LACO-Wiki to demonstrate how a global land cover product such as GlobeLand30 can be validated. Here we focus only on one country for the purpose of the demonstration: Kenya. Two different sample sets were interpreted: one was generated externally by the GlobeLand30 team [13] as part of their international validation efforts while the second was generated using the sampling design tools embedded with the LACO-Wiki workflow.

3.1. GlobeLand30

Produced by the National Geomatic Center of China (NGCC), GlobeLand30 is one of the first global 30-m land cover products produced using Landsat imagery in combination with additional imagery from the Chinese Environmental and Disaster satellite (HJ-1) [13]. Land cover maps were created for the years 2000 and 2010 from more than 10,000 Landsat images, which are downloadable from the GlobeLand30 website. Table 1 lists the 10 land cover classes. These classes were extracted in a hierarchical manner (beginning with water bodies first and ending with the tundra class) via a combination of pixel- and object-based classification approaches. Other layers were then used to improve the classification including other land cover products from national to global scales, OpenStreetMap, satellite imagery from Google Earth, etc. An online tool was developed for labelling the objects and verifying the results.

Table 1. The classes in GlobeLand30 [13].

Code	Class Name	Class Description
10	Cultivated land	Arable land (cropland): dry land, paddy field, land for greenhouses, vegetable fields, artificial tame pastures, economic cropland in which shrub crops or herbaceous crops are planted, and land abandoned with the reclamation of arable land
20	Forest	Broadleaved deciduous forest, evergreen broad-leaf forest, deciduous coniferous forest, evergreen coniferous forest, mixed broadleaf-conifer forest
30	Grassland	Typical grassland, meadow grassland, alpine grassland, desert grassland, grass
40	Shrubland	Desert scrub, mountain scrub, deciduous and evergreen shrubs
50	Wetland	Lake swamp, river flooding wetlands, seamarsh, shrub/forest wetlands, mangrove forest, tidal flats/salt marshes
60	Water bodies	Open water, i.e., lakes, reservoirs/fishponds, rivers
70	Tundra	Brush tundra, poaceae tundra, wet tundra, bare tundra, mixed tundra
80	Artificial surfaces	Settlement place, industrial and mining area, traffic facilities
90	Bareland	Saline-alkali land, sand, gravel, rock, microbiotic crust
100	Permanent snow/ice	Permanent snow, ice sheet and glaciers

The overall accuracy of the GlobeLand30 product for 2010 is greater than 80% [13]. The accuracy of artificial surfaces was tested in eight areas around the world, which ranged from 79% to 97%, showing improvements when compared with CORINE Land Cover [54] and the FROM-GLC product [55]. A more recent validation of GlobeLand30 for Dar Es Salaam, Tanzania, and Kathmandu, Nepal, yielded overall accuracies of 61% and 54%, respectively [56]. Other independent validations of GlobeLand30 include comparisons with authoritative products in Italy [57], Germany [58], Iran [59], and water bodies in Scandinavian countries [60], with agreements of greater than 78%. Hence, there are few examples of published accuracy assessments of GlobeLand30 outside of Europe. Note that the focus of this paper is to assess the accuracy of the 2010 product for Kenya and not 2000, particularly since the imagery in Google Earth and Bing is more recent so this would not be possible. For the same reason, we have not undertaken an accuracy assessment of the change between 2000 and 2010; moreover, LACO-Wiki would require other modifications to do this, e.g., displaying pairs of images, even if the imagery dates were not the limiting factor.

The NGCC, in collaboration with GEO, has invited GEO members to take part in an international validation exercise of GlobeLand30 for 2010. The GlobeLand30 team has produced a global sample using a methodology outlined in their guidance document for validation [61]. A two stage sampling design was chosen. The first stage was at the global level and involved selecting map sheets. In the second stage the sample size of each map sheet was calculated and samples were then generated using a stratified random approach where the strata were the land cover types. The objective of this

approach was to optimize the sample size while allowing for calculation of overall and class accuracies using a probability-based estimator.

The guidance document for validation also outlines the validation protocol as follows [61]: validation should take place at the level of the 30 m pixel, choosing the dominant land cover type when there are mixed pixels. The exceptions to this rule are the forest and shrubland classes, which require only 30% coverage to be assigned to these classes. The guidance document [61] also recommends that each sample is interpreted by at least three different people.

3.2. Using LACO-Wiki with an External Sample Uploaded to the System: Sample 1

LACO-Wiki can be used to interpret samples that have been generated outside of the application. The sample centroids for Kenya were obtained from the GlobeLand30 team for Kenya, which is part of this larger global sample as mentioned previously. There were 179 samples in the sample set and hereafter we refer to this as “Sample 1”. The sample set was then uploaded to LACO-Wiki and three validation sessions were created. Three different individuals (J.C.L.B., L.S. and D.S.) then interpreted the sample set independently using blind validation, producing three visual interpretations of GlobeLand30 classes at each sample location. The results from the three interpreters were then compared; there was complete agreement in 77 samples (43.0%), two out of three agreed in 89 samples (49.7%) and there was complete disagreement in 13 samples (7.3%). Rather than calculate the accuracy measures for different levels of degrees of trust, as set out in the validation guidance document for GlobeLand30 [61], we decided to produce a single consolidated reference data set for Sample 1 by reaching consensus on those samples where there was not full agreement. This involved reviewing these samples together, using Google Earth and the NDVI tool in Geo-Wiki to come to an agreement between interpreters. Google Earth is directly available from LACO-Wiki while the NDVI tool will eventually be incorporated into LACO-Wiki. The consolidated Sample 1 set was then revisited by one interpreter (L.S.) to see whether the class changed if a 3×3 neighborhood of pixels was considered, taking into account some uncertainty related to geometric accuracy; any changes in class were then recorded. Finally, three other pieces of information were recorded: (i) the date of the Bing imagery, which is displayed in LACO-Wiki; (ii) the date closest to 2010 that was available in the historical imagery in Google Earth; and (iii) whether only Landsat (or similar resolution) imagery versus very high resolution imagery was available. The confusion matrices and accuracy measures were then estimated taking the areas of the classes in Kenya into account.

3.3. Using LACO-Wiki for the Complete Validation Workflow: Sample 2

The GlobeLand30 map for Kenya was uploaded to LACO-Wiki using the procedure described in Section 3.1. A random sample of 400 30-m pixels was first generated and then interpreted by two individuals (J.C.L.B. and L.S.), where the size was chosen because it was considered manageable in terms of the time needed to complete the validation using LACO-Wiki. The reason for this initial sample was to obtain a variance estimate (σ^2):

$$\sigma^2 = (p(1 - p))^2 \quad (1)$$

where p is the lowest producer’s accuracy across all classes. This was then used to determine the required sample size (n):

$$n = (1.96^2) \left(\sigma^2 \right) (HW^{-2}) \quad (2)$$

where 1.96 is the critical t value at 95% confidence for a large enough sample (>400) and HW is the half width of the expected confidence interval (expected precision, %).

With a confidence of 95% and an initial expected precision of 5% ($HW = 2.5\%$), the estimated sample size was $n = 373$. To ensure that enough samples were interpreted, a total of 500 samples were generated. The sample design used for obtaining the sample data was stratified-random sampling, where the strata are the GlobeLand30 classes. The objective of this design was to estimate the overall

accuracy and the accuracy by class. To avoid under-sampling of small classes, a minimum of 30 samples was allocated to classes with less than 5% of map coverage, with the exception of the class Permanent ice and snow, due to its almost negligible size in Kenya (Table 2). Hereafter we refer to this sample set as “Sample 2”. The area weighted overall accuracy and weighted producer’s accuracies were then calculated using the dtwSat (Time-Weighted Dynamic Time Warping for Satellite Image Time Series Analysis) package in R, version 0.2.2, [62], which follows recommended good practices for accuracy estimation [30] and is a probability-based estimator.

Table 2. Distribution and reallocation of samples across GlobeLand30 classes present in Kenya, sorted according to the proportion of each class in the map in descending order by area.

Class	Proportion in Map (%)	Allocated Samples	Redistributed Samples
Grassland	52.8729	264	205
Shrubland	17.1389	86	72
Cultivated land	14.3653	72	65
Forest	9.1390	46	35
Bareland	3.8734	19	30
Water bodies	2.0173	10	30
Wetland	0.3907	2	30
Artificial surfaces	0.2023	1	30
Permanent snow and ice	0.0001	0	3

The same procedure as described for Sample 1 (Section 3.3) was used for interpretation of the samples, i.e., blind validation was employed. This time full agreement was reached in 273 of the samples (54.6%), two out of three interpreters agreed on 205 samples (41.0%) while complete disagreement occurred in only 22 samples (4.4%). Once again, consensus was reached to produce a final reference data set of 500 sample units that comprise Sample 2. The confusion matrices and accuracy measures were then estimated, weighted by area to account for classes with small numbers and to account for the sample reallocation (Table 2).

3.4. Models to Examine Drivers of Agreement between Interpreters

Since all three individuals interpreted each sample set individually, the agreement between the interpreters could be calculated and the possible drivers affecting the agreement could be investigated. For this purpose, a multivariate generalized linear model with a binomial distribution and a logit link function was employed. The agreement between the interpreters was translated into a binomial variable (fully agree = 1, disagree = 0) and used as the response variable, *ag*. Equation (3) lists the factors that may affect agreement:

$$ag = f(gl30, type, vhr, dif2010) \quad (3)$$

where *gl30* is the GlobeLand30 class for the sample being validated; *type* indicates if the data come from Sample 1 or Sample 2; *vhr* is a dummy variable indicating whether there was very high resolution imagery (Google or Bing) available to interpret a given sample unit; and *dif2010* is the difference in years between the available very high resolution imagery and the reference year for GlobeLand30, i.e., 2010. Initial tests for correlations between the predictor variables were undertaken before introducing the variables into the model to avoid multicollinearity.

Since the proportion of samples for each of the different land cover classes is different, these proportions were used as weights in the model. The model was run in SAS[®] using Proc Glimmix using a Laplace estimation method.

4. Results

4.1. Temporal and Spatial Resolution of the Imagery Used in the Interpretation

As outlined in Section 2.1.3, LACO-Wiki can display satellite imagery from both Google Maps and Bing through OpenLayers to undertake the interpretation. The date of the Bing imagery is provided in LACO-Wiki while the sample must be downloaded using the “Download sample (kmz)” feature and loaded into Google Earth so that we could find satellite imagery that was as close to 2010 as possible. The dates of both the Bing imagery and those closest to 2010 in Google Earth were recorded for each pixel in both Samples 1 and 2. When only Terramatics base imagery was available in Google Earth (i.e., panchromatically sharpened 15 m Landsat) or similar resolution imagery was available, e.g., from CNES/SPOT, and no images were available in Bing, this information was also recorded to differentiate between the presence and absence of very high resolution imagery. The difference between the image date and 2010 was then calculated. Figure 9 shows the distribution of the difference in years when considering only the availability of very high resolution imagery and the reference year (2010) for both Samples 1 and 2. For Sample 1, the percentage of very high resolution images within two years of 2010 is 83.7%, with a similar percentage for Sample 2, i.e., 85.1%. Thus, the majority of samples were interpreted with imagery within two years of the target date of 2010.

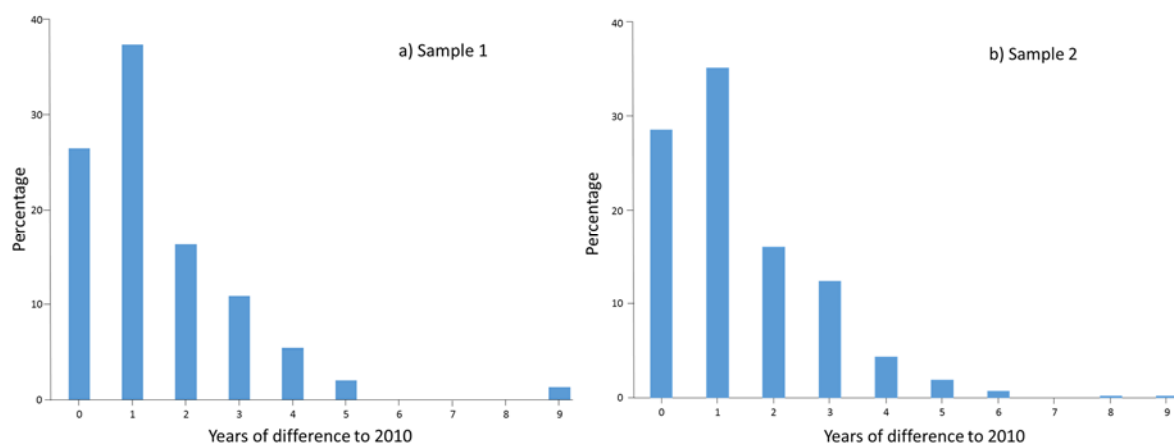


Figure 9. Distribution of the difference in years between the reference year for GlobeLand30 (2010) and the available high resolution imagery in: (a) Sample 1 ($n = 147$); and (b) Sample 2 ($n = 411$).

Figure 10 shows the agreement of the interpreters for Samples 1 and 2 together ($n = 679$) for each GlobeLand30 class. In addition, a relative importance index for very high resolution imagery was calculated, which is simply the number of very high resolution images available for a given class multiplied by the percentage of sample units for that class. The index highlights those classes for which the presence of very high resolution imagery is important for visual interpretation, in particular the Grassland class followed by the Shrubland and Cultivated land classes.

4.2. Accuracy Assessment of GlobeLand30 for Kenya

Table 3 contains the confusion matrix and the accuracy measures for Sample 1 based on the consensus between the three interpreters where both first and second choices for classes are allowed; this occurred in 12.8% of the sample. Table 4 is similar to Table 3 except that any of the following classes are allowed (Forest, Shrubland, Cultivated land, Grassland, Bareland, Wetland and Permanent Snow and Ice) when only Landsat or similar resolution imagery was the only imagery available. Table 5 takes the 3×3 pixel interpretations into account. Based on these three different variations (Tables 3–5), the overall accuracy varies from 0.53 to 0.61. However, the overall accuracies for Sample 1 are only

slightly larger than Sample 2 (varying from 0.56 to 0.60), where the confusion matrices are reported in Tables 6–8. Thus, both samples produced a similar picture in terms of overall accuracy.

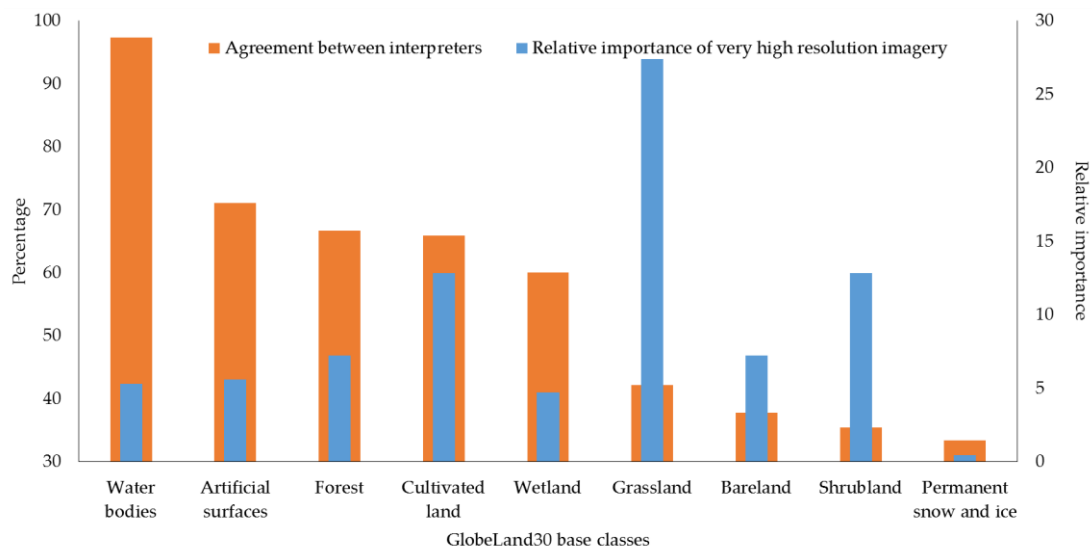


Figure 10. The agreement of interpreters and the relative importance of very high resolution imagery by class in GlobeLand30.

For Sample 1, the producer's accuracy is very high for Artificial Surfaces, Cultivated land, Water bodies and Wetland while Forest has the lowest producer's accuracy (Tables 3 and 4) except for Grassland when considering 3×3 pixels. For Sample 2, the producer's accuracy was very high for similar classes to Sample 1 except for Artificial Surfaces and Wetland. However, this is because these classes only have a very small area within Kenya. For example, the producer's accuracy for Artificial surfaces before weighting is 0.85 but since the weighted producer's accuracy also depends on the mapped areas of the other classes, the weighted producer's accuracy is low because the misclassified pixels occur in Cultivated land (Table 6), which is a class with a large area relative to Artificial surfaces. If the misclassification had occurred in a class such as Wetland, which is also a small class based on area, then the producer's accuracy would have been much higher. The weighted producer's accuracies for Artificial surfaces and Wetland also have large confidence intervals. Forest, Shrubland and Bareland have similar producer's accuracies to Sample 1 while Grassland is much higher in this second sample, i.e., ranging from 0.73 to 0.77.

For Sample 1, the user's accuracy for Forest and Water Bodies is very high followed by Artificial Surfaces and Bareland, while Wetland is the class with the lowest values (0.20–0.30). Some improvements in Grassland can be observed when multiple classes are considered, i.e., from 0.48 (Table 3) to 0.55 (Table 4), but it decreases to 0.38 when considering 3×3 pixels. Similar patterns can be found in Sample 2 for user's accuracy except that Grassland has the lowest values (0.43 to 0.51), while the Wetland class performs better (ignoring the Permanent snow and ice class due to small numbers).

Figure 11 shows the amount of area mapped by GlobeLand30 compared to the adjusted areas from Sample 2 (allowing for choice 1 or 2). The graph indicates that Forest, Shrubland and Bareland are underestimated in GlobeLand30 for Kenya while Cultivated land and Grassland are overestimated since the areas fall outside of the confidence intervals of the sample.

Table 3. Confusion matrix for Sample 1 ($n = 179$) considering either choice 1 or choice 2 when present (which occurred in 12.8% of the sample). User's and producer's accuracy with 95% confidence intervals shown. Overall weighted accuracy is 0.57 ± 0.091 . Cells in bold are the diagonal of the confusion matrix.

		Interpreters								User. Acc.	\pm CI
		Artificial Surfaces	Bareland	Cult. Land	Forest	Grassland	Shrubland	Water Bodies	Wetland		
GlobeLand30	Artificial surfaces	6	0	0	0	2	0	0	0	0.75	0.32
	Bareland	0	16	0	1	4	2	0	0	0.70	0.19
	Cult. land	0	0	13	4	5	1	0	0	0.57	0.21
	Forest	0	0	1	22	1	1	0	0	0.88	0.13
	Grassland	0	3	0	10	20	9	0	0	0.48	0.15
	Shrubland	0	5	0	8	3	25	0	0	0.61	0.15
	Water bodies	0	0	0	0	0	0	7	0	1.00	0.00
	Wetland	0	3	0	5	0	0	0	2	0.20	0.26
	Prod Acc.	1.00	0.31	0.96	0.30	0.82	0.45	1.00	1.00		
	\pm CI	0.00	0.17	0.08	0.09	0.09	0.15	0.00	0.00		

Table 4. Confusion matrix for Sample 1 ($n = 179$) using results that allow for multiple classes when Landsat or other similar resolution imagery is only available for validation. User's and producer's accuracy with 95% confidence intervals shown. Overall weighted accuracy is 0.61 ± 0.091 . Cells in bold are the diagonal of the confusion matrix.

		Interpreters								User. Acc.	\pm CI
		Artificial Surfaces	Bareland	Cult. Land	Forest	Grassland	Shrubland	Water Bodies	Wetland		
GlobeLand30	Artificial surfaces	6	0	0	0	2	0	0	0	0.75	0.32
	Bareland	0	16	0	1	4	2	0	0	0.70	0.19
	Cult. land	0	0	13	4	5	1	0	0	0.57	0.21
	Forest	0	0	1	22	1	1	0	0	0.88	0.13
	Grassland	0	3	0	9	23	7	0	0	0.55	0.15
	Shrubland	0	5	0	8	3	25	0	0	0.61	0.15
	Water bodies	0	0	0	0	0	0	7	0	1.00	0.00
	Wetland	0	3	0	4	0	0	0	3	0.30	0.30
	Prod Acc.	1.00	0.31	0.96	0.32	0.84	0.51	1.00	1.00		
	\pm CI	0.00	0.17	0.08	0.10	0.08	0.17	0.00	0.00		

Table 5. Confusion matrix for Sample 1 ($n = 179$) considering 3×3 pixel validations. User's and producer's accuracy with 95% confidence intervals shown. Overall weighted accuracy is 0.53 ± 0.090 . Cells in bold are the diagonal of the confusion matrix.

		Interpreters								User. Acc.	\pm CI
		Artificial Surfaces	Bareland	Cult. Land	Forest	Grassland	Shrubland	Water Bodies	Wetland		
GlobeLand30	Artificial surfaces	5	0	0	0	3	0	0	0	0.63	0.36
	Bareland	0	16	0	0	4	3	0	0	0.70	0.19
	Cult. land	0	0	15	3	5	0	0	0	0.65	0.20
	Forest	0	0	1	22	1	1	0	0	0.88	0.13
	Grassland	0	3	0	10	16	13	0	0	0.38	0.15
	Shrubland	0	4	0	5	8	24	0	0	0.59	0.15
	Water bodies	0	0	0	0	0	0	7	0	1.00	0.00
	Wetland	0	3	0	4	0	1	0	2	0.20	0.26
	Prod. Acc.	1.00	0.33	0.96	0.33	0.73	0.37	1.00	1.00		
	\pm CI	0.00	0.19	0.07	0.10	0.12	0.12	0.00	0.00		

Table 6. Confusion matrix for Sample 2 ($n = 500$) considering either choice 1 or choice 2 when present (which occurred in 12.2% of the sample). User's and producer's accuracy with 95% confidence intervals shown. Overall weighted accuracy is 0.59 ± 0.046 . Cells in bold are the diagonal of the confusion matrix.

		Interpreters									User Acc.	\pm CI
		Artificial Surfaces	Bareland	Cult. Land	Forest	Grassland	P. Snow and Ice	Shrubland	Water Bodies	Wetland		
GlobeLand30	Artificial surfaces	22	0	1	1	5	0	1	0	0	0.73	0.16
	Bareland	0	26	0	0	3	0	1	0	0	0.87	0.12
	Cult. land	2	0	45	3	9	0	5	1	0	0.69	0.11
	Forest	0	0	0	30	3	0	1	0	1	0.86	0.12
	Grassland	0	6	1	37	99	0	60	0	2	0.48	0.07
	P. snow and ice	0	1	0	0	0	2	0	0	0	0.67	0.65
	Shrubland	0	6	0	4	20	0	42	0	0	0.58	0.11
	Water bodies	0	0	0	0	0	0	0	30	0	1.00	0.00
	Wetland	0	0	0	7	1	0	0	4	18	0.60	0.18
	Prod. Acc.	0.25	0.52	0.97	0.41	0.76	1.00	0.37	0.88	0.24		
\pm CI	0.26	0.14	0.05	0.07	0.06	0.00	0.07	0.17	0.21			

Table 7. Confusion matrix for Sample 2 ($n = 500$) that allow for multiple classes when Landsat or other similar resolution imagery is only available for validation. User’s and producer’s accuracy with 95% confidence intervals shown. Overall weighted accuracy is 0.60 ± 0.046 . Cells in bold are the diagonal of the confusion matrix.

		Interpreters											
		Artificial Surfaces	Bareland	Cult. Land	Forest	Grassland	P. Snow and Ice	Shrubland	Water Bodies	Wetland	User Acc.	\pm CI	
GlobeLand30	Artificial surfaces	22	0	1	1	5	0	1	0	0	0.73	0.16	
	Bareland	0	26	0	0	3	0	1	0	0	0.87	0.12	
	Cult. land	2	0	45	3	9	0	5	1	0	0.69	0.11	
	Forest	0	0	0	30	3	0	1	0	1	0.86	0.12	
	Grassland	0	6	1	37	105	0	54	0	2	0.51	0.07	
	P. snow and ice	0	1	0	0	0	2	0	0	0	0.67	0.65	
	Shrubland	0	6	0	4	20	0	42	0	0	0.58	0.11	
	Water bodies	0	0	0	0	0	0	0	30	0	1.00	0.00	
	Wetland	0	0	0	6	1	0	0	4	19	0.63	0.18	
	Prod Acc.	0.25	0.52	0.97	0.41	0.77	1.00	0.40	0.88	0.25			
\pm CI	0.26	0.14	0.05	0.07	0.06	0.00	0.07	0.17	0.22				

Table 8. Confusion matrix for Sample 2 ($n = 500$) considering 3×3 pixel validations. User’s and producer’s accuracy with 95% confidence intervals shown. Overall weighted accuracy is 0.56 ± 0.046 . Cells in bold are the diagonal of the confusion matrix.

		Interpreters											
		Artificial Surfaces	Bareland	Cult. Land	Forest	Grassland	P. Snow and Ice	Shrubland	Water Bodies	Wetland	User Acc.	\pm CI	
GlobeLand30	Artificial surfaces	20	0	1	2	6	0	1	0	0	0.67	0.17	
	Bareland	0	24	0	0	2	0	4	0	0	0.80	0.15	
	Cult. land	0	0	44	2	12	0	6	1	0	0.68	0.11	
	Forest	0	0	0	30	3	0	1	0	1	0.86	0.12	
	Grassland	0	8	2	36	89	0	68	0	2	0.43	0.07	
	P. snow and ice	0	2	0	0	0	1	0	0	0	0.33	0.65	
	Shrubland	0	5	0	5	20	0	42	0	0	0.58	0.11	
	Water bodies	0	0	0	0	0	0	0	30	0	1.00	0.00	
	Wetland	0	0	0	6	1	0	1	3	19	0.63	0.18	
	Prod Acc.	1.00	0.47	0.95	0.42	0.73	1.00	0.34	0.88	0.25			
\pm CI	0.00	0.14	0.07	0.07	0.06	0.00	0.06	0.17	0.22				

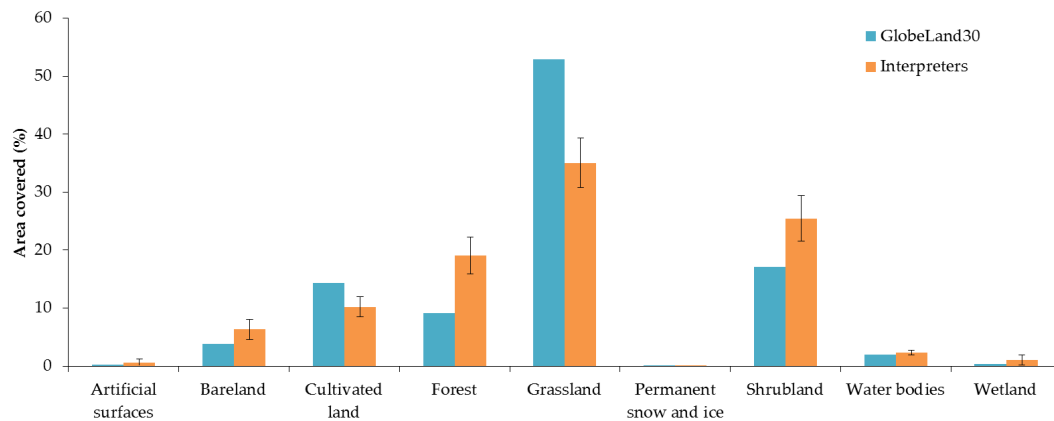


Figure 11. Comparison of the mapped area of GlobeLand30 compared to the adjusted areas based on Sample 2 (allowing for choices 1 or 2) showing 95% confidence intervals.

4.3. Drivers of Agreement

For the model described in Equation (3), all the predictors tested showed significant effects ($\alpha = 5\%$). Specific results show that interpreters were 1.8 times more likely to agree on Sample 2 than on Sample 1. This could be due to the fact that the interpreters discussed different examples after completing Sample 1 and landscapes would also have been more familiar by the time that Sample 2 was interpreted. Despite this difference in agreement, the overall accuracies on both sample sets are similar (see Section 4.2).

Another result from the model is that the type of land cover class in Globeland30 significantly affects the agreement between the interpreters. The magnitudes of agreement and their confidence intervals (95% confidence) per class are shown in Figure 12. If we disregard the class Permanent snow and ice because of its small sample size ($n = 3$) and large variability, the next three classes with the lowest agreement between interpreters were Shrubland, Bareland and Grassland as shown in Figure 12. The confusion matrices also show a large part of the disagreement to be between these classes.

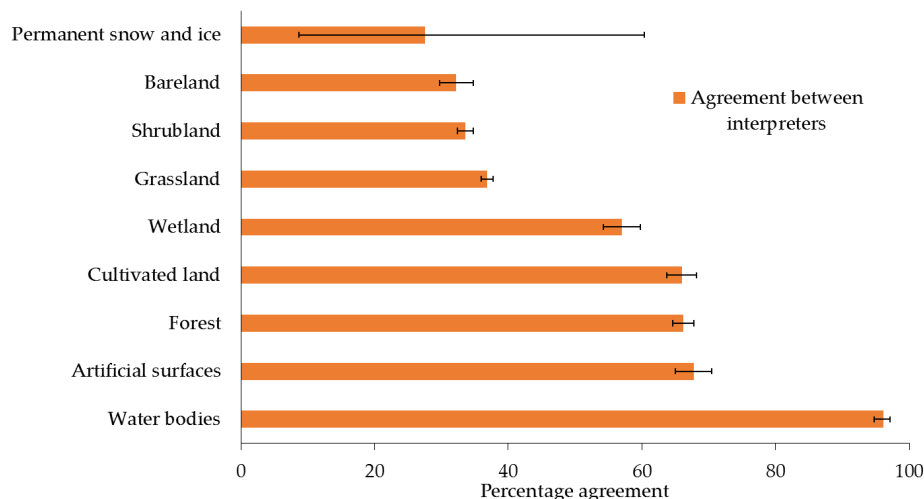


Figure 12. Agreement between interpreters showing adjusted medians and 95% confidence intervals ($n = 679$).

Additionally, when a sample location had very high resolution imagery available, the interpreters were 18% more likely to agree between them on any class. Furthermore, for every year of difference

between the image dates and 2010, the interpreters were 3.5% more likely to disagree between them on a given class, which is a significant but small effect.

5. Discussion

One of the criticisms that is often leveled at tools such as LACO-Wiki and Geo-Wiki is the problem related to the date of the imagery at the sample locations and the date of the land cover map. Hence if satellite imagery from Google Maps and Bing are the only sources of imagery used for sample interpretation, then it is important to record the dates of this imagery and to use the historical archive in Google Earth to support the accuracy assessment. As shown in the exercise undertaken here, more than 83% of the imagery used is within two years of 2010, which is the year to which GlobeLand30 corresponds. Moreover, there were no images used that were more than six years before or beyond 2010. Although we did not systematically record examples of where images before and after showed no change, there were a number of examples of where this was clearly the case, supporting the use of images that are not only available for 2010 or close to that year. Although the model showed a significant effect between agreement and the time difference in years between the imagery and 2010, implying that images further away in time from the base year 2010 had an effect on the agreement, this effect is very small. Figure 13 shows an example of the Google Maps image that appears in LACO-Wiki for 2017 (Figure 13a), the Bing image for 2012 (Figure 13b) and the image for 2010 that can be found in Google Earth (Figure 13c), which shows considerable change over time. If the Bing image had not been available, then the need to use the historical archive in Google becomes even more critical.



Figure 13. Imagery from: (a) Google Maps in LACO-Wiki for 2017; (b) Bing in LACO-Wiki for 2012; and (c) Google Earth for 2010.

To undertake an accuracy assessment of GlobeLand30, very high resolution imagery is ideally needed. However, this is not always available. Landsat or similar resolution imagery in Google Earth occurred in 17.8% of the time in both Samples 1 and 2. This situation is made worse when Bing very high resolution imagery is also not available, which occurred 6.1% of the time in Sample 1 and 4.2% of the time in Sample 2. In these latter situations, it is very hard to recognize the land cover types or differentiate between types such as Forest, Shrubland, Grassland, Bareland or Wetland. For this reason we allowed for multiple land cover types as acceptable interpretations when only Landsat or similar resolution imagery was available in Google Earth (Tables 4 and 7). These sample units should be considered as highly uncertain and ideally require another source of information such as ground-based

data, geotagged photographs or orthophotos. Although we checked Google Earth for the presence of Panoramio photographs, there were rarely any present in the locations being interpreted. Note that we manually took the imagery types and presence/absence of Bing imagery into account but ideally this functionality would be added to LACO-Wiki in the future.

The results of the accuracy assessment show that certain classes are more problematic than others in terms of confusion between classes. One clear area of confusion is between the Grassland and Bareland classes. It was sometimes difficult to determine from the imagery alone whether surfaces were vegetated or bare. For this reason we used the NDVI tool in Geo-Wiki to examine the mean profiles of Landsat 7 and 8 to distinguish between the Grassland and Bareland classes. For this reason we would recommend that the international validation of GlobeLand30 include such tools to aid the accuracy assessment process.

Another source of confusion was between Forest and Shrubland classes, since it was not always possible to see the heights of the trees or shrubs from the images. Where possible we looked at the size of crowns and for the presence of shadows to denote trees but this task was still difficult. The use of a 30% percent threshold for these two classes also made judgment difficult when the vegetation was sparse and near this threshold. Hence there was additional confusion between these classes and Grassland/Bareland. In some cases the use of a 3×3 pixel took care of this issue but not always. We would suggest that an online gallery containing many different examples of land cover types be developed by the GlobeLand30 team to aid their validation process. We have employed such an approach during our recent SIGMA campaign to crowd source cropland data [63].

Another source of error may have arisen from lack of local knowledge. The three interpreters have backgrounds in physical geography (L.S.), agronomy (J.C.L.B) and forestry (D.S.) although all have experience in remote sensing and visual image interpretation. However, none of these interpreters are from Kenya but rather from different locations around the world, i.e., Canada, Ecuador and Russia. These two samples should now be interpreted by local Kenyans to see whether local knowledge affects the agreement with GlobeLand30. This could have clear implications for needing the active involvement of GEO members in validating GlobeLand30 for their own countries. The results would also be of interest within the field of visual interpretation.

6. Conclusions

This paper outlined a new online tool called LACO-Wiki, which provides an easy-to-use application for accuracy assessment, capturing the full workflow from uploading a land cover map to generating accuracy measures. Such a tool is intended to make the accuracy assessment easier but is also a way of sharing reference data and land cover maps. We then demonstrated the use of the tool for the accuracy assessment of GlobeLand30 for Kenya. In addition to outlining the main functionality of the tool, the exercise undertaken here represents a contribution to the international efforts of NGCC in validating their 30 m global land cover product. We will share the two samples with the GlobeLand30 team and encourage them to review the sample sets and work with local Kenyans to repeat the exercise. Through this process, a definitive reference data set for Kenya could be developed and shared as part of the larger GlobeLand30 validation effort. Although the overall accuracy was only 53–61% compared to the overall accuracy of >80% achieved for the global product [13], this is not surprising since Kenya has many complex savannah landscapes that can be very difficult to interpret even using very high resolution imagery. Moreover, 18% of each of the two samples contained only Landsat or similar imagery in Google Earth, which further complicated the visual interpretation. Despite these issues, LACO-Wiki represents a tool that can be used to aid in the accuracy assessment of global products such as GlobeLand30.

Acknowledgments: The authors would like to acknowledge support from the Austrian Agency for the Promotion of Science (FFG) under ASAP-10 in the LACO-Wiki project (No. 844385), the LACO-Wiki extension funded by the EEA, the EU-funded ERC CrowdLand project (No. 617754) and the Horizon2020 LandSense project (No. 689812).

Author Contributions: L.S. and J.C.L.B. wrote the paper. L.S., J.C.L.B. and D.S. did the validation of GlobeLand30 with LACO-Wiki. J.C.L.B. did the analysis of the data and fit the statistical models. C.P. and C.D. developed LACO-Wiki. C.S. and V.M. wrote the R routines for the accuracy assessment used in LACO-Wiki. V.M. provided the R code for area weighted accuracy assessment and area estimates. J.W. was involved in the testing of LACO-Wiki during its development. S.F., M.L., I.McC. and I.M. have contributed ideas for and suggestions to the LACO-Wiki tools and provided comments and edited the paper.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AMI	Average Mutual Information
CEOS-CVWG	CEOS Calibration/Validation Working Group
DLR	German Aerospace Agency
EEA	European Environment Agency
EO	Earth Observation
FROM-GLC	Finer Resolution Mapping of Global Land Cover
GEO	Group on Earth Observations
GEO QA4EO	GEO's Quality Assurance Framework for Earth Observation
GOFC-GOLD	Global Observation for Forest Cover and Land Dynamics
KML	Keyhole Markup Language
NGCC	National Geomatics Center of China
NDVI	Normalized Difference Vegetation Index
OATH2	Open Authentication
OSM	OpenStreetMap
QA4EO	GEO's Quality Assurance Framework for Earth Observation
SDGs	Sustainable Development Goals
WMS	Web Map Service

References

- Hollmann, R.; Merchant, C.J.; Saunders, R.; Downy, C.; Buchwitz, M.; Cazenave, A.; Chuvieco, E.; Defourny, P.; de Leeuw, G.; Forsberg, R.; et al. The ESA Climate Change Initiative: Satellite data records for Essential Climate Variables. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 1541–1552. [[CrossRef](#)]
- Poulter, B.; MacBean, N.; Hartley, A.; Khlystova, I.; Arino, O.; Betts, R.; Bontemps, S.; Boettcher, M.; Brockmann, C.; Defourny, P.; et al. Plant functional type classification for earth system models: Results from the European Space Agency's Land Cover Climate Change Initiative. *Geosci. Model Dev.* **2015**, *8*, 2315–2328. [[CrossRef](#)]
- United Nations Statistical Commission (Ed.) *System of Environmental-Economic Accounting 2012: Central Framework*; United Nations Statistical Commission: New York, NY, USA, 2014; ISBN 978-92-79-35797-8.
- Becker-Reshef, I.; Justice, C.; Sullivan, M.; Vermote, E.; Tucker, C.; Anyamba, A.; Small, J.; Pak, E.; Masuoka, E.; Schmaltz, J.; et al. Monitoring global croplands with coarse resolution Earth Observations: The Global Agriculture Monitoring (GLAM) project. *Remote Sens.* **2010**, *2*, 1589–1609. [[CrossRef](#)]
- Smith, P.; Gregory, P.J.; van Vuuren, D.; Obersteiner, M.; Havlik, P.; Rounsevell, M.; Woods, J.; Stehfest, E.; Bellarby, J. Competition for land. *Philos. Trans. R. Soc. B Biol. Sci.* **2010**, *365*, 2941–2957. [[CrossRef](#)] [[PubMed](#)]
- Nel, J.L.; Le Maitre, D.C.; Nel, D.C.; Reyers, B.; Archibald, S.; van Wilgen, B.W.; Forsyth, G.G.; Theron, A.K.; O'Farrell, P.J.; Kahinda, J.-M.M.; et al. Natural hazards in a changing world: A case for ecosystem-based management. *PLoS ONE* **2014**, *9*, e95942. [[CrossRef](#)] [[PubMed](#)]
- Hackman, K.O. A method for assessing land-use impacts on biodiversity in a landscape. *Glob. Ecol. Conserv.* **2015**, *3*, 83–89. [[CrossRef](#)]
- De Fries, R.; Hansen, M.; Townshend, J.; Sohlberg, R. Global land cover classifications at 8 km spatial resolution: The use of training data derived from landsat imagery in decision tree classifiers. *Int. J. Remote Sens.* **1998**, *19*, 3141–3168. [[CrossRef](#)]

9. 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.; et al. Validation of the global land cover 2000 map. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1728–1737. [[CrossRef](#)]
10. Friedl, M.A.; Sulla-Menashe, D.; Tan, B.; Schneider, A.; Ramankutty, N.; Sibley, A.; Huang, X. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens. Environ.* **2010**, *114*, 168–182. [[CrossRef](#)]
11. Defourny, P.; Vancustem, C.; Bicheron, P.; Brockmann, C.; Nino, F.; Schouten, L.; Leroy, M. GLOBCOVER: A 300 m global land cover product for 2005 using ENVISAT MERIS time series. In Proceedings of the ISPRS Commission VII Mid-Term Symposium: Remote Sensing: From Pixels to Processes, Enschede, NL, USA, 8–11 May 2006.
12. Yu, L.; Wang, J.; Li, X.; Li, C.; Zhao, Y.; Gong, P. A multi-resolution global land cover dataset through multisource data aggregation. *Sci. China Earth Sci.* **2014**, *57*, 2317–2329. [[CrossRef](#)]
13. Chen, J.; Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Peng, S.; Lu, M.; et al. Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 7–27. [[CrossRef](#)]
14. Wulder, M.A.; Masek, J.G.; Cohen, W.B.; Loveland, T.R.; Woodcock, C.E. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sens. Environ.* **2012**, *122*, 2–10. [[CrossRef](#)]
15. Wulder, M.A.; Coops, N.C. Satellites: Make Earth observations open access. *Nature* **2014**, *513*, 30–31. [[CrossRef](#)] [[PubMed](#)]
16. Herold, M.; See, L.; Tsendbazar, N.-E.; Fritz, S. Towards an integrated global land cover monitoring and mapping system. *Remote Sens.* **2016**, *8*, 1036. [[CrossRef](#)]
17. Marconcini, M. Mapping urban areas globally by jointly exploiting optical and radar imagery—The GUF+ layer. In Proceedings of the World Cover 2017, ESA-ESRIN, Frascati, Italy, 14–16 March 2017.
18. Feddema, J.J.; Oleson, K.W.; Bonan, G.B.; Mearns, L.O.; Buja, L.E.; Meehl, G.A.; Washington, W.M. The Importance of Land-Cover Change in Simulating Future Climates. *Science* **2005**, *310*, 1674–1678. [[CrossRef](#)] [[PubMed](#)]
19. Intergovernmental Panel on Climate Change (IPCC). *Summary for Policymakers. Climate Change 2014, Mitigation of Climate Change; Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Edenhofer, O., Pichs-Madruga, R., Sokona, Y., Farahani, E., Kadner, S., Seyboth, K., Adler, A., Baum, I., Brunner, S., Eickemeier, P., et al., Eds.; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2014.
20. Group on Earth Observations (GEO). Earth Observations in Support of the 2030 Agenda for Sustainable Development. Japan Aerospace Exploration Agency: Tokyo, Japan, 2017; p. 19. Available online: http://earthobservations.org/documents/publications/201703_geo_eo_for_2030_agenda.pdf (accessed on 22 April 2017).
21. Defourny, P.; Brockmann, C.; Bontemps, S.; Achard, F.; Boettcher, M.; Maet, T.D.; Gamba, P.; Hagemann, S.; Hartley, A.; Hoffman, L.; et al. A consistent 300 m global land cover and land cover change time series from 1992 to 2015 derived from multi-mission reprocessed archives. In Proceedings of the GOFC-GOLD Land Cover Meeting, The Hague, The Netherlands, 4 November 2016.
22. Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.; Turubanova, S.A.; Tyukavina, A.; Thau, D.; Stehman, S.V.; Goetz, S.J.; Loveland, T.R.; et al. High-resolution global maps of 21st-century forest cover change. *Science* **2013**, *342*, 850–853. [[CrossRef](#)] [[PubMed](#)]
23. Pekel, J.-F.; Cottam, A.; Gorelick, N.; Belward, A.S. High-resolution mapping of global surface water and its long-term changes. *Nature* **2016**, *540*, 418–422. [[CrossRef](#)] [[PubMed](#)]
24. Pesaresi, M.; Huadong, G.; Blaes, X.; Ehrlich, D.; Ferri, S.; Gueguen, L.; Halkia, M.; Kauffmann, M.; Kemper, T.; Lu, L.; et al. A Global Human Settlement Layer from optical HR/VHR RS data: Concept and first results. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 2102–2131. [[CrossRef](#)]
25. Stehman, S.V.; Czaplewski, R.L. Design and analysis for thematic map accuracy assessment: Fundamental principles. *Remote Sens. Environ.* **1998**, *64*, 331–344. [[CrossRef](#)]
26. McRoberts, R.E.; Walters, B.F. Statistical inference for remote sensing-based estimates of net deforestation. *Remote Sens. Environ.* **2012**, *124*, 394–401. [[CrossRef](#)]

27. Strahler, A.H.; Boschetti, L.; Foody, G.M.; Friedl, M.A.; Hansen, M.C.; Herold, M.; Mayaux, P.; Morisette, J.T.; Stehman, S.V.; Woodcock, C.E. Global Land Cover Validation: Recommendations for Evaluation and Accuracy Assessment of Global Land Cover Maps. 2006. Available online: <http://landval.gsfc.nasa.gov/pdf/GlobalLandCoverValidation.pdf> (accessed on 22 April 2017).
28. Stehman, S.V. Estimating area from an accuracy assessment error matrix. *Remote Sens. Environ.* **2013**, *132*, 202–211. [CrossRef]
29. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57. [CrossRef]
30. Olofsson, P.; Foody, G.M.; Stehman, S.V.; Woodcock, C.E. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens. Environ.* **2013**, *129*, 122–131. [CrossRef]
31. Foody, G.M.; Boyd, D.S. Using volunteered data in land cover map validation: Mapping West African forests. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 1305–1312. [CrossRef]
32. Bastin, L.; Buchanan, G.; Beresford, A.; Pekel, J.-F.; Dubois, G. Open-source mapping and services for Web-based land-cover validation. *Ecol. Inform.* **2013**, *14*, 9–16. [CrossRef]
33. Brice, M.; Tsendbazar, N.-E.; Herold, M.; Arino, O. Global land cover mapping: Current status and future trends. In *Land Use and Land Cover Mapping in Europe*; Manakos, I., Braun, M., Eds.; Remote Sensing and Digital Image Processing; Springer: Dordrecht, The Netherlands, 2014; Volume 18, pp. 11–30. ISBN 978-94-007-7968-6.
34. Leutner, B.; Horning, T. RStoolbox: Tools for Remote Sensing Data Analysis in R. Available online: <https://bleutner.github.io/RStoolbox/> (accessed on 22 April 2017).
35. Kopeinig, R.; Equihua, J.; Gebhardt, S. QGIS Python Plugins Repository: Validation Tool. Available online: <https://plugins.qgis.org/plugins/vTool/> (accessed on 22 April 2017).
36. Pontius, R.G.; Millones, M. Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *Int. J. Remote Sens.* **2011**, *32*, 4407–4429. [CrossRef]
37. Conrad, O.; Bechtel, B.; Bock, M.; Dietrich, H.; Fischer, E.; Gerlitz, L.; Wehberg, J.; Wichmann, V.; Böhner, J. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geosci. Model Dev.* **2015**, *8*, 1991–2007. [CrossRef]
38. Geographic Resources Analysis Support System (GRASS). Development Team Geographic Resources Analysis Support System (GRASS) Software, Version 7.2. Open Source Geospatial Foundation, 2017. Available online: <http://grass.osgeo.org> (accessed on 31 October 2016).
39. International Institute for Geo-Information Science and Earth Observation (ITC). 52North ILWIS, the Free User-Friendly Raster and Vector GIS. 1984. Available online: <http://www.ilwis.org/> (accessed on 22 April 2017).
40. GISGeography. 13 Open Source Remote Sensing Software Packages. 2017. Available online: <http://gisgeography.com/open-source-remote-sensing-software-packages/> (accessed on 22 April 2017).
41. QA4EO Secretariat QA4EO. Available online: <http://qa4eo.org/> (accessed on 22 April 2017).
42. Bicheron, P.; Defourny, P.; Brockmann, C.; Schouten, L.; Vancutsem, C.; Huc, M.; Bontemps, S.; Leroy, M.; Achard, F.; Herold, M.; et al. *Globcover: Products Description and Validation Report*; MEDIAS-France: Saint Ouen, France, 2008; p. 47.
43. GOFCC-GOLD. GOFCC-GOLD Reference Data Portal. 2015. Available online: http://www.gofccgold.wur.nl/sites/gofccgold_refdataportal.php (accessed on 22 April 2017).
44. United States Geological Survey (USGS). Global Land Cover Validation Reference Dataset. Available online: <https://landcover.usgs.gov/glc/SitesDescriptionAndDownloads.php> (accessed on 22 April 2017).
45. Pengra, B.; Long, J.; Dahal, D.; Stehman, S.V.; Loveland, T.R. A global reference database from very high resolution commercial satellite data and methodology for application to Landsat derived 30 m continuous field tree cover data. *Remote Sens. Environ.* **2015**, *165*, 234–248. [CrossRef]
46. Olofsson, P.; Stehman, S.V.; Woodcock, C.E.; Sulla-Menashe, D.; Sibley, A.M.; Newell, J.D.; Friedl, M.A.; Herold, M. A global land-cover validation data set, part I: Fundamental design principles. *Int. J. Remote Sens.* **2012**, *33*, 5768–5788. [CrossRef]
47. Zhao, Y.; Gong, P.; Yu, L.; Hu, L.; Li, X.; Li, C.; Zhang, H.; Zheng, Y.; Wang, J.; Zhao, Y.; et al. Towards a common validation sample set for global land-cover mapping. *Int. J. Remote Sens.* **2014**, *35*, 4795–4814. [CrossRef]

48. Fritz, S.; See, L.; Perger, C.; McCallum, I.; Schill, C.; Schepaschenko, D.; Duerauer, M.; Karner, M.; Dresel, C.; Laso Bayas, J.C.; et al. A global dataset of crowdsourced land cover and land use reference data. *Sci. Data* **2017**, *4*, 170075. [[CrossRef](#)] [[PubMed](#)]
49. Tsendbazar, N.E.; de Bruin, S.; Herold, M. Assessing global land cover reference datasets for different user communities. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 93–114. [[CrossRef](#)]
50. Congalton, R.G.; Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*; Lewis Publishers: Boca Raton, FL, USA, 1999.
51. Finn, J.T. Use of the average mutual information index in evaluating classification error and consistency. *Int. J. Geogr. Inf. Syst.* **1993**, *58*, 349–366. [[CrossRef](#)]
52. Comber, A.; Fisher, P.; Brunsdon, C.; Khmag, A. Spatial analysis of remote sensing image classification accuracy. *Remote Sens. Environ.* **2012**, *127*, 237–246. [[CrossRef](#)]
53. Yu, L.; Liang, L.; Wang, J.; Zhao, Y.; Cheng, Q.; Hu, L.; Liu, S.; Yu, L.; Wang, X.; Zhu, P.; et al. Meta-discoveries from a synthesis of satellite-based land-cover mapping research. *Int. J. Remote Sens.* **2014**, *35*, 4573–4588. [[CrossRef](#)]
54. Büttner, G.; Kosztra, B.; Maucha, G.; Pataki, R. *Implementation and Achievements of CLC2006*; European Environment Agency: Copenhagen, Denmark, 2012.
55. Gong, P.; Wang, J.; Yu, L.; Zhao, Y.; Zhao, Y.; Liang, L.; Niu, Z.; Huang, X.; Fu, H.; Liu, S.; et al. Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data. *Int. J. Remote Sens.* **2013**, *34*, 2607–2654. [[CrossRef](#)]
56. Fonte, C.; Minghini, M.; Patriarca, J.; Antoniou, V.; See, L.; Skopeliti, A. Generating up-to-date and detailed land use and land cover maps using OpenStreetMap and GlobeLand30. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 125. [[CrossRef](#)]
57. Brovelli, M.A.; Molinari, M.E.; Hussein, E.; Chen, J.; Li, R. The first comprehensive accuracy assessment of GlobeLand30 at a national level: Methodology and results. *Remote Sens.* **2015**, *7*, 4191–4212. [[CrossRef](#)]
58. Arsanjani, J.J.; See, L.; Tayyebi, A. Assessing the suitability of GlobeLand30 for mapping land cover in Germany. *Int. J. Digit. Earth* **2016**, *9*, 873–891. [[CrossRef](#)]
59. Arsanjani, J.J.; Tayyebi, A.; Vaz, E. GlobeLand30 as an alternative fine-scale global land cover map: Challenges, possibilities, and implications for developing countries. *Habitat Int.* **2016**, *55*, 25–31. [[CrossRef](#)]
60. Ban, Y.; See, L.; Haas, J.; Jacob, A.; Fritz, S. Validation of the water layers of global land cover products using Geo-Wiki and national land cover maps. In Proceedings of the 35th EARSel Symposium, Stockholm, Sweden, 15–18 June 2015.
61. Chen, J.; Tong, X.; Li, S.; Perger, C.; Leinenkugle, P.; Wakhayanga, J.A.; Chen, L.; Liu, C.; Wang, Z.; Han, G.; et al. *Technical Specification for the International Validation of Global Land Cover Products*; NGCC: Beijing, China, 2016.
62. Maus, V. dtwSAT R Package, v.0.2.2. Time-Weighted Dynamic Time Warping for Satellite Image Time Series Analysis. 2017. Available online: <https://github.com/vwmaus/dtwSat> (accessed on 1 May 2017).
63. Laso Bayas, J.-C.; Fritz, S.; See, L.; Perger, C.; Dresel, C.; Hofer, M.; Weichselbaum, J.; McCallum, I.; Moorthy, I.; Kraxner, F.; et al. Extending the LACO-Wiki tool for land cover validation to crop area estimation. In Proceedings of the International Conference on Agricultural Statistics VII (ICAS), FAO, Rome, Italy, 26–28 October 2016.

