

Data Descriptor

A Unified Cropland Layer at 250 m for Global Agriculture Monitoring

François Waldner ^{1,*}, Steffen Fritz ², Antonio Di Gregorio ³, Dmitry Plotnikov ⁴, Sergey Bartalev ⁴, Nataliia Kussul ⁵, Peng Gong ⁶, Prasad Thenkabail ⁷, Gerard Hazeu ⁸, Igor Klein ⁹, Fabian Löw ¹⁰, Jukka Miettinen ¹¹, Vinay Kumar Dadhwal ¹², Céline Lamarche ¹, Sophie Bontemps ¹ and Pierre Defourny ¹

¹ Earth and Life Institute, Université catholique de Louvain, Croix du Sud, 1348 Louvain-la-Neuve, Belgium; celine.lamarche@uclouvain.be (C.L.); sophie.bontemps@uclouvain.be (S.B.); pierre.defourny@uclouvain.be (P.D.)

² International Institute for Applied Systems Analysis (IIASA), 2361 Laxenburg, Austria; fritz@iiasa.ac.at

³ Food and Agriculture Organisation (FAO), 00153 Rome, Italy; antoniodigregorio12@gmail.com

⁴ Space Research Institute (IKI), Russian Academy of Sciences, Profsoyuznaya str. 117997 Moscow, Russia; dmitplot@d902.iki.rssi.ru (D.P.); bartalev@d902.iki.rssi.ru (S.B.)

⁵ Space Research Institute NAS & SSA (SRI), Glushkov Prospekt 40, 03680 Kyiv, Ukraine; nataliia.kussul@gmail.com

⁶ Tsinghua University, 30 Shuang Qing Lu, Tsinghua University, 30 Shuang Qing Lu Beijing, China; penggong@tsinghua.edu.cn

⁷ United States Geological Survey, Flagstaff, Arizona 86001, USA; pthenkabail@usgs.gov

⁸ Alterra, Droevendaalsesteeg, 6708 PB Wageningen, The Netherlands; gerard.hazeu@wur.nl

⁹ Deutschen Zentrums für Luft- und Raumfahrt (DLR), Linder Höhe, 51147 Cologne, Germany; igor.klein@dlr.de

¹⁰ Bonn University, 53113 Bonn, Germany; fabian.loew@gmx.de

¹¹ Centre for Remote Imaging, Sensing and Processing (CRISP) at the National University of Singapore (NUS), Singapore, Singapore; crsjukk@nus.edu.sg

¹² National Remote Sensing Centre, ISRO, Balanagar, Hyderabad 500037, India; dadhwalvk@hotmail.com

* Correspondence: francois.waldner@uclouvain.be; Tel.: +32(0)104-736-80

Academic Editors: Juan C. Jiménez-Muñoz and Jamal Jokar Arsanjani

Received: 22 January 2016; Accepted: 9 March 2016; Published: 19 March 2016

Abstract: Accurate and timely information on the global cropland extent is critical for food security monitoring, water management and earth system modeling. Principally, it allows for analyzing satellite image time-series to assess the crop conditions and permits isolation of the agricultural component to focus on food security and impacts of various climatic scenarios. However, despite its critical importance, accurate information on the spatial extent, cropland mapping with remote sensing imagery remains a major challenge. Following an exhaustive identification and collection of existing land cover maps, a multi-criteria analysis was designed at the country level to evaluate the fitness of a cropland map with regards to four dimensions: its timeliness, its legend, its resolution adequacy and its confidence level. As a result, a Unified Cropland Layer that combines the fittest products into a 250 m global cropland map was assembled. With an evaluated accuracy ranging from 82% to 95%, the Unified Cropland Layer successfully improved the accuracy compared to single global products.

Data Set: <https://dx.doi.org/10.6084/m9.figshare.2066742> or <http://maps.elie.ucl.ac.be>

Data Set License: CC-BY

Keywords: global; agriculture; cropland map; multi-criteria analysis; accuracy assessment

1. Summary

Mapping the global cropland extent is of paramount importance for agricultural production assessment. Timely and accurate cropland information directly feeds global crop monitoring systems [1] and early warning systems such as the Global Information and Early Warning System (GIEWS), the Early Warning Crop Monitor and the Famine Early Warning Systems Network (FEWSNET) [2,3]. It also serves environmental climate change studies [4]. In both agriculture monitoring and climate modeling, cropland maps mask terrestrial areas dedicated to agriculture in order to (1) assess and compare crop growth and conditions and (2) investigate how agricultural land use could respond to different climatic change scenarios.

Satellite remote sensing provides opportunities for global cropland monitoring in a spatially explicit, economical, efficient, and objective fashion [5]. In the last forty years, numerous initiatives aimed at deriving cropland either as a single class in land cover products or as a collection of agricultural land use classes (crops) via classification of satellite images. A large diversity of mapping strategies ranging from the local to the global scale and associated with various degrees of accuracy are documented in the literature [6–10]. The existing satellite derived information on the spatial extent of croplands might be categorized in four types of products. First, global land cover maps (GLC2000 [11], GlobCover 2005/2009 [12,13], MODerate resolution Imaging Spectroradiometer (MODIS) Land Cover [14] and the European Space Agency Climate Change Initiative Land Cover products [15]) often describe the cropland according to a Land Cover Classification System (LCCS) typology and focus mainly on vegetation types. As some cropland areas are also often included in mosaic or mixed classes, their integration in agricultural applications is not trivial. In addition, when analyzing their consistency, Fritz *et al.* [16] highlighted that they underestimate cropland compared to the official statistics and that they also disagree with one another. Earlier, Ramankutty *et al.* [17] quantified this uncertainty at the global scale: they estimated that the global cropland extent varies between 1.22 and 1.71 billion hectares, *i.e.*, more than 40% of the terrestrial surface. With the first 30 m resolution global land cover map, an increase of accuracy was expected but not observed in the case of the cropland class [18]. Second, high resolution national or regional land cover mapping efforts such as Africover and Corine Land Cover have also provided detailed cropland information, but the update frequency is too low for operational crop monitoring. Third, cropland maps were produced at the global, continental or national scale [5–7]. Some of them were strictly devoted to cropland mapping with an emphasis on water management: the global map of rainfed cropland areas (GMRCA) [19] and the global irrigated area map (GIAM) [20]. The relatively coarse spatial resolution of these products (10 km) does not meet the needs for operational applications and suffers from large uncertainties [3]. Fourth, some countries (e.g., the USA, Canada and India) have established dedicated annual national crop type mapping based on satellite remote sensing data. Those initiatives remain limited in number as they require advanced operational remote sensing programs and processing capabilities as well as intensive field data collection. Finally, numerous countries maintain a land parcel information system for managing the farmers' declarations, e.g., Land Parcel Identification System in Europe.

Given the known limited accuracy of existing global data sets and the still unresolved challenges for mapping specific areas, an evaluation and merger of existing data sets appear as worthwhile alternatives to creating new data layers from scratch. Hence, this paper presents the first 250 m global cropland map generated by combination of existing maps, hereafter referred to as "Unified Cropland Layer". After an exhaustive identification and collection of existing maps, the selection of the maps to use was done through a multi-criteria analysis. The Unified Cropland Layer aims at providing a map of the annually cultivated areas fitting the JECAM (Joint Experiment of Crop Assessment and Monitoring) cropland definition. The JECAM network has adopted a shared cropland definition that defines the annual cropland as a piece of land of minimum 0.25 ha (minimum width of 30 m) that is sowed/planted and harvestable at least once within the 12 months after the sowing/planting date. The annual cropland produces an herbaceous cover and is sometimes combined with some tree or woody vegetation. There are three known exceptions to this definition. The first one concerns

the sugarcane plantation and cassava crop which are included in the cropland class although they have a longer vegetation cycle and are not yearly planted. Second, taken individually, small plots such as legumes do not meet the minimum size criteria of the cropland definition. However, when considered as a continuous heterogeneous field, they should be included in the cropland. The third case is the greenhouse crops that cannot be monitored by remote sensing and are thus excluded from the definition. This shared definition facilitates across-site comparisons in benchmarking activities as well as map integration.

2. Data Description

2.1. Data and Metadata

In order to produce the best global cropland map, the fittest existing country-level cropland maps were identified by means of a multi-criteria analysis [21]. Four criteria were selected to evaluate the fitness of a cropland map: the adequacy of the legend, the spatial resolution, the timeliness and the confidence level (see Section 4). These fittest maps were resampled at 250 m in a reference grid with an average resampling method and then joined spatially into the Unified Cropland Layer (Figure 1). This implies that, wherever possible, the Unified Cropland Layer provides cropland proportion rather than a binary cropland/non-cropland information. Average aggregation was found to retain a greater amount of information than other statistical aggregation techniques [22] and provides more flexibility to the users. For products provided on an annual basis, e.g., the US Cropland Data Layer, 2014 was chosen as the reference year.

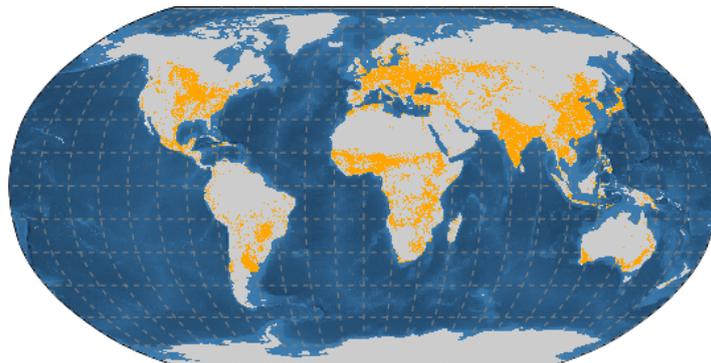


Figure 1. The Unified Cropland Layer at 250 m for 2014.

Table 1 provides a summary of the main characteristics of the data set. It is provided as a single GeoTiff in a latitude/longitude grid of about 355 MB. The valid data range spans from 0 to 100 where 0 represents non cropped pixels and 100 corresponds to fully cropped pixels.

Table 1. Data set characteristics of the Unified Cropland Layer 2014.

Characteristic	Description
Data format	GeoTiff
Epoch	2014
Coordinate system	latitude/longitude WGS84 (EPSG:4326)
Image dimensions	172,800 × 83,294 (rows × columns)
Size	355 MB
Data type	Byte using LZW compression
No Data value	255
Number of layers	1
Value	Cropland proportion in percent
Data range	0–100

2.2. Accuracy Assessment

The accuracy of the Unified Cropland Layer was assessed in two different ways: (1) with the receiver operating characteristic (ROC) curve and (2) with traditional confusion matrices. In both cases, three different and independent reference data sets were used: the consolidated GlobCover 2005 data set, the VIIRS (Visible Infrared Imaging Radiometer Suite) data set and the geoWiki data set (see Section 3.2).

As the Unified Cropland Layer provides continuous cropland proportion values, an ROC curve analysis was implemented. The ROC curve analysis for two classes (cropland/non-cropland) plots the true positive rate on the y -axis and their equivalent false-positive rate on the x -axis for every possible cut-off cropland proportion [23]. Accuracy was measured by the area under the ROC curve (AUC), which provides a single measure of the probability that the classifier will rank a randomly-chosen positive instance higher than a randomly-chosen negative instance [24]. An area of 1 represents a perfect accuracy; an area of 0.5 represents the performance of a random classification. Validation with the three sets show that the AUCs range from 0.8 to 0.93 (Figure 2).

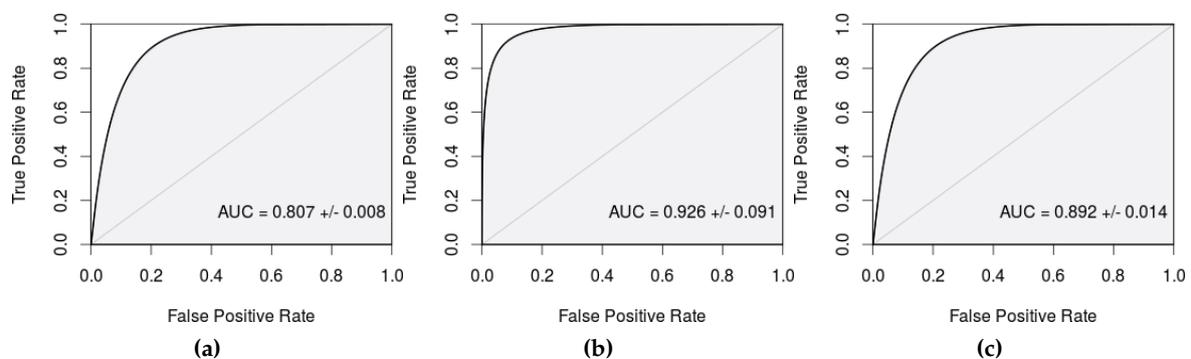


Figure 2. Receiver Operator Characteristics curves with the (a) geoWiki; (b) GlobCover 2005 and (c) VIIRS validation data sets.

The accuracy was further evaluated by means of confusion matrices (Table 2a–c) and their derived indices [25,26]: (i) the overall accuracy (OA) which expresses the probability of a randomly-selected pixel to be classified accurately; (ii) the producers' accuracy which is defined for a given class as the conditional probability that a pixel classified as category c by the reference data is classified as category c by the map; and (iii) the users' accuracy which is the conditional probability that an area classified as category c by the map is classified as category c by the reference data. Validation with confusion matrices requires discrete inputs, *i.e.*, cropland/non-cropland. A conservative cut-off was chosen: pixel with at least one percent of cropland were considered as cropland. According to the three validation data sets, the overall accuracy figures span from 84% to 94%. This is close to the similar—but at 1 km resolution—International Institute for Applied Systems Analysis-International Food Policy Research Institute (IIASA-IFPRI) map [27], and its overall accuracy is improved with respect to individual cropland products. While the producers' and users' accuracies range between 71–99% for the non-cropland class, large differences appear for the cropland class. This might be partially explained by the thematic mismatch between the JECAM and the validation set definitions, which penalizes the Unified Cropland Layer that is more restrictive semantically. The highest accuracy figures for the cropland class are obtained with the VIIRS data set: 94% of producers' accuracy and 80% of users' accuracy.

Table 2. Accuracy assessments of the Unified Cropland Layer of 2014 with the three different validation data sets.

(a) Confusion matrix obtained with the GlobCover2005 data set			
	Non-Cropland	Cropland	Users' Accuracy [%]
Non-Cropland	158	9	94.6
Cropland	2	16	88.9
Producers' Accuracy [%]	98.8	64.0	Overall Accuracy [%]: 94.1
(b) Confusion matrix obtained with the VIIRS data set			
	Non-Cropland	Cropland	Users' Acc. (%)
Non-Cropland	631	63	90.9
Cropland	251	985	79.7
Producers' Accuracy [%]	71.5	94.0	Overall Accuracy [%]: 83.7
(c) Confusion matrix obtained with the geoWiki data set			
	Non-Cropland	Cropland	Users' Acc. (%)
Non-Cropland	8490	1698	83.3
Cropland	384	2055	84.3
Producers' Accuracy [%]	95.7	54.8	Overall Accuracy [%]: 83.5

Overall, the validation exercise revealed an improved accuracy over individual cropland products. Figures from the GlobCover accuracy assessment might be too optimistic as this consolidated validation data mainly focuses on areas with obvious interpretation (see Section 3.2 for more details). These figures as well as the differences between the three accuracy assessments still highlight the need for both improved spatial cropland information (especially in southeast Asia as well as western and southern Africa) and validation data. It should be noted that none of the validation data sets entirely fit the cropland definition of the Unified Cropland Layer, and the accuracy figures presented might be revised when using one that matches. However, no such global validation data set yet exists.

2.3. Perspectives of Evolution

In the moderate and long term, the Unified Cropland Layer will evolve with future map releases (e.g., with the complete Corine Land Cover 2012 coverage), changes in data policy and with new contributions from different national and international institutions. Thus, the accuracy of the Unified Cropland Layer is expected to increase over time. However, without legend harmonization, multi-product maps will remain inconsistent by definition. The FAO Land Cover Meta Language (LCML) provides a robust theoretical framework for legend definition and would certainly play a key role in class definition harmonization.

To facilitate the identification and use of land cover products by the community, this study recommends systematically registering Earth Observation resources into the GEOSS Portal which is a main entry point to Earth Observation data from all over the world with linking world-wide community of practice in nine societal benefit areas among which agriculture. With the new and upcoming high resolution satellites such as Landsat-8 and the Sentinel-1 and -2, the number of high resolution land cover products is expected to increase and their accuracies to improve.

3. Material

3.1. Land Cover and Cropland Maps

Global, regional and national data sets were identified by means of systematic review during working sessions with key resource people and expert networks, literature review and web-based search. About 80 global, regional and national maps were identified (Table 3). It rapidly became

necessary to distinguish the existing data sets from the ones available; the former having a distribution policy that prevents its use or having issues to access the geo-referenced data source. Therefore, the general rule was to consider only the data sets granting rights of reuse and technically available.

Table 3. Input maps for the analysis.

Extent	Product Name-Reference	Epoch
Global	FROM-GLC [18]	2013
	Global Cropland extent [7]	2000–2008
	GlobCover 2009 [13]	2009
	ESA LandCover CCI [15]	2008–2012
	MOD12Q1 NASA	2005
	FAO GLC-Share [28]	1990–2012
	IIASA-IFPRI Cropland [27]	1990–2012
	GLC2000 [11]	1999–2000
	IGBP [29]	1992–1993
GLCNMO [30]	2007–2009	
Regional	Corine Land Cover EEA	2006, 2012
	SADC land cover database-CSIR	2002
	JRC Cropland Mask [3]	2012
	North American Environmental Atlas CEC	2005
	SERENA LAC [31]	2008
	Congo Basin Map [32]	2000–2007
	SEA CRISP [33]	2010
	Land Cover of Central Asia [34]	2009
Northern Eurasia Land Cover (NELC) [35]	2005	
Uzbekistan, Tajikistan, Kyrgyzstan, Turkmenistan	Wurzburg University	2003
Congo, Burundi, Egypt, Eritrea, Kenya, Rwanda, Somalia, Sudan, Tanzania, Uganda	Africover FAO	1999–2001
Senegal, Bhutan, Nepal	Global Land Cover Network	2005–2007
France, Belgium, The Netherlands	Land Parcel Identification System	2012–2014
Barbados, Rep Dom, Dominica, Grenada, Puerto Rico, Saint Kit and Nevis US and B Virgin Islands	USGS	2000–2001
Fiji, Solomon Islands, Timor Leste, Niue, Naurau, Palau, Tonga, Tuvalu, Vanuatu, Kiribati, Marshall Islands, Micronesia, Cook Islands	SOPAC	1999–2010
Botswana, Namibia, Rwanda, Zambia, Tanzania, Malawi	Land Cover Scheme II SERVIR	2010
China	GlobeLand30 NGCC [36]	2009–2011
Japan	JAXA HR LU-LCMap [37]	2006–2011
Tajikistan	ACCA [6]	2010
Burkina Faso	Corine Database of Burkina Faso	2000
Canada	Annual Crop Inventory-AAFC	2013
USA	Cropland Data Layer USDA	2013
China	National Land Cover Map of China [38]	1995–1996
Australia	Digital Land Cover Database GA-Australia [39]	2011
Cambodia	JICA Land Cover of Cambodia	2002
New Zealand	Land Cover DataBase v4 Ministry for the Environment	2004
South Africa	National Land Cover CSIR	2000–2001
South Africa	National Land Cover SANBI	2009
Canada	National Resources of Canada	2005
Uruguay	Land Cover Uruguay UNA-ONU	2010
Mexico	Land Cover of Mexico CONABIO	1999
Argentina	Cobertura y uso del suelo-INTA	2006
Ecuador	Uso del Suelo departamento de Inf. Ambiental	2001
Thailand	Royal Forest Department of Thailand	2000
Chile	Chile Corporacion Nacional Forestal	1999
India	Land Use Land Cover of India NRSC [40]	2011–2012
Gambia	[41]	2013
Ukraine	Land Cover Ukraine [42]	2010
Russia	TerraNorte Arable Lands of Russia [43]	2014

3.2. Global Validation Data Sets and Ancillary Data

Four global validation data sets were collected for two purposes: (1) one for the assessment of the individual maps and (2) three for the validation of the unified map (Table 4). For the criteria assessment purpose, an independent set of well-distributed validation samples produced by Zhao *et al.* [44] was utilized. Two reasons motivated this choice: (1) it was the most populated data set, and (2) its legend was closer to the one used in this study than the other available validation data sets. It consists of a global punctual data set based on interpreting Landsat Thematic Mapper and Enhanced TM+ images for a total of 38,664 sample units pre-determined with an equal-area stratified sampling scheme. This was supplemented by MODIS enhanced vegetation index time series data and other high-resolution imagery on Google Earth.

Table 4. Validation data sets collected, their geometries and percentage of cropland samples.

Validation Set	Geometry	Sample Size	Cropland [%]	Use
Zhao <i>et al.</i>	Point	38 664	7	Multi-criteria analysis
GlobCover 2005	Polygon (225 ha)	186	9	Validation
VIIRS	Polygon (5 × 5 km)	3664	27	Validation
geoWiki	Polygon (1 × 1 km)	12 833	29	Validation

Regarding the validation purpose, the GlobCover 2005, the VIIRS and the geoWiki data sets were selected (Figure 3). The GlobCover 2005 validation data set was built relying on a network of experts familiar with image interpretation and land cover over large areas [12]. In total, 16 experts committed to interpret validation samples through a dedicated working environment relying on Google Earth and also providing 10-day NDVI profiles to illustrate the seasonal dynamics. For a given sample, the expert saw not only the sample point but also a 225 ha box that coincided with a 5×5 MEdium Resolution Imaging Spectrometer (MERIS) pixel (MERIS being the sensor used to generate the GlobCover product). For each sample, the experts could describe up to three land cover types and had to provide their level of confidence. This dataset counts 4258 samples, from which 500 samples were randomly selected and re-interpreted by a second independent expert to generate a consolidated dataset. Only the samples with high confidence (186) were kept in this study, which might lead to an optimistic bias in the estimation of the map accuracy. The VIIRS Surface Type validation database is based on a stratified random sample of 500 5 × 5 km blocks [45,46]. The samples were extracted from strata derived from an intersection of a Koppen climate classification modified with a human population density layer. Samples were interpreted and labeled with very high resolution imagery. The allocation of samples within each stratum was targeted towards heterogeneous and complex land cover types that are more difficult to map. Recently, Fritz *et al.* [47] proposed a tool known as the geoWiki to collect volunteered geographic information on land cover from crowd-sourcing. The geoWiki Project capitalizes on a global network of volunteers who wish to help to improve the quality of global land cover maps. The volunteers were asked to review hot-spot areas where existing global land cover maps disagree. They had to determine, based on what they actually see in Google Earth and their local knowledge, if the land cover maps are correct or incorrect [48].

A global field size data set was also made available as a result of a geoWiki crowd-sourcing campaign [48] in which volunteers labelled the field size in four categories (very small, small, medium and large) based on very high resolution images and on a reference 1 × 1 km square box [27]. Crowd-sourced observations were then interpolated within the cropland using an inverse distance approach.

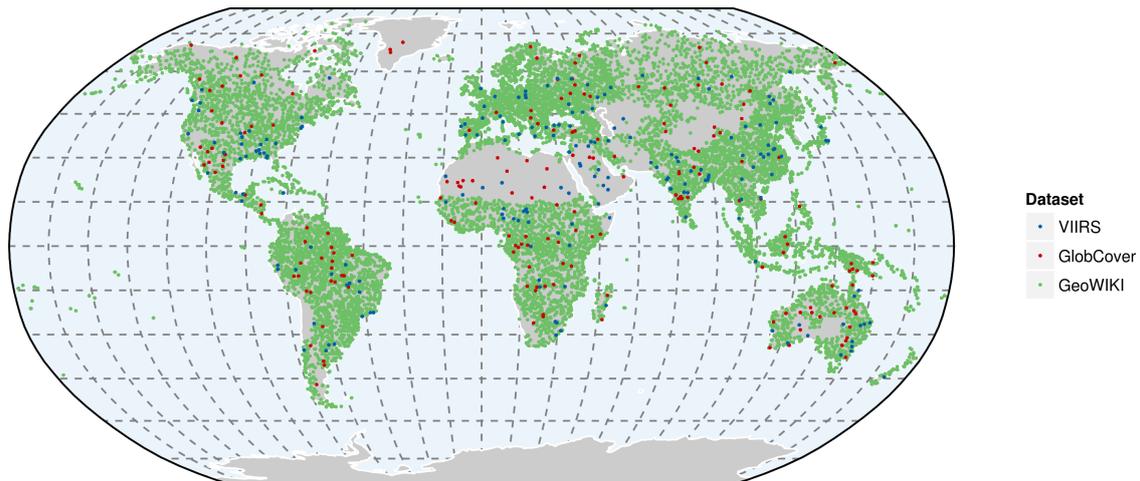


Figure 3. Distribution of the reference samples used for the validation of the 250 m Unified Cropland Layer.

4. Methodology

A cropland map must combine different criteria such as its adequacy with the legend and the spatial resolution, its timeliness and its confidence level. Therefore, the assessment of the cropland products must consider these four different criteria. To handle these four dimensions, a multi-criteria analysis was designed. A multi-criteria analysis is a particularly powerful tool when it comes down to combining the conflicting objectives described by different data sources into a single index form [21] in order to support decision making and priority analysis [49]. The general outline of the methodology included the following steps:

- Constructing a spatial information data base;
- Translating the criteria into scores;
- Defining the weight of each criterion;
- Aggregating the criteria in the output index and selecting the product that maximizes this index

Four criteria have been selected to evaluate the fitness of the cropland maps: the adequacy of the current legend (ThC), the adequacy of the spatial resolution (RC), the timeliness (TiC) and the confidence level (CC). The scores were attributed following a default rule (Table 5) and were then reinterpreted by experts to ensure consistency with their experience and/or visual assessment. The four criteria were finally aggregated into a single fitness indicator (FI) before being again reinterpreted by the experts. The fitness indicator FI of cropland map j for a country i is computed as:

$$FI_{ij} = ThC_{ij} + RC_{ij} + TiC_{ij} + CC_{ij}. \quad (1)$$

Low scores indicate low fitness with regard to the four prerequisite of a good cropland mask. The Unified Cropland Layer is the spatial combination of the fittest national maps—the products which maximize the fitness indicator. For more details about the methodology, the reader is referred to Waldner *et al.* [50].

Table 5. Default rules of the four components considered in the multi-criteria analysis.

(a) Rules for the thematic criterion		
Thematic Criterion	Code	Score
3	Good Thematic Agreement	4
2	Moderate Thematic agreement	3
1	Low Thematic agreement	2
0	No thematic agreement	1
(b) Rules for the timeliness criterion		
Timeliness Criterion	Code	Score
1–2	Up to date	4
2–5	Recent	3
5–10	Old	2
10>	Out of date	1
(c) Rules for the resolution adequacy criterion		
Resolution Adequacy Criterion	Code	Score
>0	Perfectly adequate	4
1	Adequate	3
2	Inadequate	2
3	Totally Inadequate	1
(d) Confidence Level Scoring		
Confidence Level Criterion	Code	Score
80%–100%	High Confidence Level	4
70%–80%	Good Confidence Level	3
60%–70%	Low Confidence	2
0%–60%	Very Low Confidence Level	1

4.1. Thematic Consistency Criterion

This criterion evaluates the consistency between the cropland definition of the individual product and the Unified Cropland Layer's. Indeed, as there is no broad agreement on cropland definition, a variety of definitions can be found in the individual products, that matches the Unified Cropland Layer's in various degrees. Even if the JECAM definition appears as the most appropriate, a more pragmatic definition was adopted to cope with the current limitations of global land cover legends. The definition was thus modified as follows: "the cropland is a specific area occupied by an herbaceous crop under permanent or fallow cultivation period (including active shifting cultivation fields)". To evaluate the thematic distance to the proposed definition, we compare it to other definitions on the basis of a set of binary criteria, *i.e.*, presence/absence of a given component. The components were the following:

1. Absence of woody crops (WC);
2. Presence of fallows and bare fields (FB);
3. Absence of managed pasture and meadows (MPM).

If, for a given product, a criterion is met, this product scores 1 and 0 conversely. Individual WC, FB, MPM scores are then summed to obtain the final score of the Thematic distance Criterion (ThC):

$$\text{ThC} = \text{WC} + \text{FB} + \text{MPM}. \quad (2)$$

In order to allow compatibility with the three other criteria TiC, RC and CC, this score originally ranging from 0 to 3, is reclassified between 1 and 4 (Table 5a).

4.2. Timeliness Criterion

The timeliness criterion (TiC) characterizes the number of years elapsed since the epoch of a map. The Timeliness Criterion is computed as the difference between the reference year of the unified cropland layer (2014) and the epoch of a given map:

$$\text{TiC} = 2014 - t_p \quad (3)$$

where t_p is the reference epoch of the map. The differences are then reclassified in four groups for which a score between 1 and 4 is assigned (Table 5b).

4.3. Resolution Adequacy Criterion

The spatial resolution required to accurately map the cropland of a given area is a function of the field size, the landscape fragmentation and some extent of the crop diversity. Areas with small adjacent parcels but low crop diversity tend to behave similarly to large fields, whereas higher spatial resolution would be required in areas with similar field size but higher crop diversity. Here, only the field size was taken as proxy to derive the spatial resolution requirements. Country-level field size histograms were derived based on the global field size data set. It was assumed that the adequate field size was the one allowing to map 75% of the largest fields. In practice, according to the distribution, this resolution would tend to fit more than 75% of the cropland extent. The extracted resolution was then related to GEOGLAM spatial resolution requirements (Table 6) that define the required resolution for cropland mapping for different field sizes [51]. The number of categories separating the field size class and the class that meets the requirements provided the final resolution criterion value (Table 5c).

Table 6. Linking the observed field size by crowd-sourcing to GEOGLAM’s spatial resolution requirements.

GeoWiki Field Size	GEOGLAM Field Size (ha)	GEOGLAM Resolution Requirements (m)
Large	>15	100–500
Medium	>1.5	20–100
Small	>0.15	5–20
Very Small	<0.15	<5

4.4. Confidence Level Criterion

Prior to the confidence level assessment, it was necessary to harmonize the legend of both the maps and the validation data sets. The legend of each data set was thus translated into binary legends that fit most of the legends chosen for this research. Then, the confidence level of the maps was assessed by means of confusion matrices obtained from the Zhao *et al.* [44] reference data set. The overall accuracy was then derived and reclassified into four confidence level categories (Table 5d).

5. User Notes

The cropland definition used for the Unified Cropland Layer was a pragmatic one, still constrained by the current diverse definitions. Principally, the pragmatic approach was to use the level of thematic details that is generally achieved with land covers at medium to coarse resolution, *i.e.*, multi-annual cropland that includes the fallow sections. However, users shall bear in mind that the products discard the fallow sections and the permanent crops whenever possible.

The crop proportions provided by the Unified Cropland Layer might be affected by some geometric uncertainty that would affect the average aggregator when going to 250 m pixels. Successive resamplings might also result in artifacts, especially in areas where the spatial resolution of the It is expected that the accuracy of the fractional estimates vary depending on the original resolution of the data product.

The accuracy assessment based on confusion matrices is based on a conservative threshold. However, users might use a different threshold according to their specific needs. They could also fine-tune the cut-off value to maximize the local/regional accuracy of the product, e.g., with a Youden index approach.

Acknowledgments: The authors would like to thank all the experts and data providers for sharing their expertise and/or data sets. This work would not have been possible if it was not for them. The research was conducted in the framework of the SIGMA (Stimulating Innovation for Global Monitoring of Agriculture and Its Impact on the Environment in Support of GEOGLAM) project funded by the European Commission in the Seventh Programme for research, technological development and demonstration under grant agreement No. 603719. The cropland map for Ukraine was also created within the SIGMA project. The arable land mapping over the Russian Federation was supported by the Russian Science Foundation under Grant number 14-17-00389.

Author Contributions: François Waldner performed the analysis under the supervision of Pierre Defourny. All authors participated in the writing and in the revision the manuscript as well as in the collection of data sets.

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

References

1. Wu, B.; Gommes, R.; Zhang, M.; Zeng, H.; Yan, N.; Zou, W.; Zheng, Y.; Zhang, N.; Chang, S.; Xing, Q.; *et al.* Global Crop Monitoring: A Satellite-Based Hierarchical Approach. *Remote Sens.* **2015**, *7*, 3907–3933.
2. Hannerz, F.; Lotsch, A. Assessment of remotely sensed and statistical inventories of African agricultural fields. *Int. J. Remote Sens.* **2008**, *29*, 3787–3804.
3. Vancutsem, C.; Marinho, E.; Kayitakire, F.; See, L.; Fritz, S. Harmonizing and combining existing land cover/land use datasets for cropland area monitoring at the African continental scale. *Remote Sens.* **2012**, *5*, 19–41.
4. Lobell, D.; Bala, G.; Duffy, P. Biogeophysical impacts of cropland management changes on climate. *Geophys. Res. Lett.* **2006**, *33*, doi:10.1029/2005GL025492.
5. Yu, L.; Wang, J.; Clinton, N.; Xin, Q.; Zhong, L.; Chen, Y.; Gong, P. FROM-GC: 30 m global cropland extent derived through multisource data integration. *Int. J. Digit. Earth* **2013**, *6*, 521–533.
6. Thenkabail, P.S.; Wu, Z. An automated cropland classification algorithm (ACCA) for Tajikistan by combining Landsat, MODIS, and secondary data *Remote Sens.* **2012**, *4*, 2890–2918.
7. Pittman, K.; Hansen, M.C.; Becker-Reshef, I.; Potapov, P.V.; Justice, C.O. Estimating global cropland extent with multi-year MODIS data. *Remote Sens.* **2010**, *2*, 1844–1863.
8. Xiao, X.; Boles, S.; Froking, S.; Li, C.; Babu, J.Y.; Salas, W.; Moore, B. Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. *Remote Sens. Environ.* **2006**, *100*, 95–113.
9. Waldner, F.; Canto, G.S.; Defourny, P. Automated annual cropland mapping using knowledge-based temporal features. *ISPRS J. Photogramm. Remote Sens.* **2015**, *110*, 1–13.
10. Matton, N.; Canto, G.S.; Waldner, F.; Valero, S.; Morin, D.; Inglada, J.; Arias, M.; Bontemps, S.; Koetz, B.; Defourny, P. An Automated Method for Annual Cropland Mapping along the Season for Various Globally-Distributed Agrosystems Using High Spatial and Temporal Resolution Time Series. *Remote Sens.* **2015**, *7*, 13208–13232.
11. Bartholomé, E.; Belward, A. GLC2000: A new approach to global land cover mapping from Earth observation data. *Int. J. Remote Sens.* **2005**, *26*, 1959–1977.
12. Defourny, P.; Schouten, L.; Bartalev, S.; Bontemps, S.; Cacetta, P.; Gérard, B.; Giri, C.; Gond, V.; Hazeu, G.; Heinimann, A.; *et al.* Accuracy assessment of a 300 m global land cover map: The GlobCover experience. In Proceedings of the 33rd International Symposium on Remote Sensing of Environment, Stresa, Italy, May 2009.
13. Arino, O.; Bicheron, P.; Achard, F.; Latham, J.; Witt, R.; Weber, J.L. The most detailed portrait of Earth. *Eur. Space Agency* **2008**, *136*, 25–31.
14. Friedl, M.A.; McIver, D.K.; Hodges, J.C.; Zhang, X.; Muchoney, D.; Strahler, A.H.; Woodcock, C.E.; Gopal, S.; Schneider, A.; Cooper, A.; *et al.* Global land cover mapping from MODIS: Algorithms and early results. *Remote Sens. Environ.* **2002**, *83*, 287–302.
15. Defourny, P.; Kirches, G.; Brockmann, C.; Boettcher, M.; Peters, M.; Bontemps, S.; Lamarche, C.; Schlerf, M.; Santoro, M. Land Cover CCI. In *PRODUCT USER GUIDE VERSION 2*; 2012.

16. Fritz, S.; See, L.; McCallum, I.; Schill, C.; Obersteiner, M.; Van der Velde, M.; Boettcher, H.; Havlík, P.; Achard, F. Highlighting continued uncertainty in global land cover maps for the user community. *Environ. Res. Lett.* **2011**, *6*, 044005.
17. Ramankutty, N.; Evan, A.T.; Monfreda, C.; Foley, J.A. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Glob. Biogeochem. Cycles* **2008**, *22*, doi:10.1029/2007GB002952.
18. Gong, P.; Wang, J.; Yu, L.; 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.
19. Biradar, C.M.; Thenkabail, P.S.; Noojipady, P.; Li, Y.; Dheeravath, V.; Turrall, H.; Velpuri, M.; Gumma, M.K.; Gangalakunta, O.R.P.; Cai, X.L.; *et al.* A global map of rainfed cropland areas (GMRCA) at the end of last millennium using remote sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2009**, *11*, 114–129.
20. Thenkabail, P.S.; Biradar, C.M.; Noojipady, P.; Dheeravath, V.; Li, Y.; Velpuri, M.; Gumma, M.; Gangalakunta, O.R.P.; Turrall, H.; Cai, X.; *et al.* Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium. *Int. J. Remote Sens.* **2009**, *30*, 3679–3733.
21. Setegn, S.G.; Srinivasan, R.; Dargahi, B.; Melesse, A.M. Spatial delineation of soil erosion vulnerability in the Lake Tana Basin, Ethiopia. *Hydrol. Process.* **2009**, *23*, 3738–3750.
22. Bian, L.; Butler, R. Comparing effects of aggregation methods on statistical and spatial properties of simulated spatial data. *Photogramm. Eng. Remote Sens.* **1999**, *65*, 73–84.
23. Zweig, M.H.; Campbell, G. Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clin. Chem.* **1993**, *39*, 561–577.
24. Fawcett, T. Introduction to Receiver Operator Curves. *Pattern Recognit. Lett.* **2006**, *27*, 861–874.
25. Foody, G.M. Status of land cover classification accuracy assessment. *Remote Sens. Environ.* **2002**, *80*, 185–201.
26. Stehman, S.V. Selecting and interpreting measures of thematic classification accuracy. *Remote Sens. Environ.* **1997**, *62*, 77–89.
27. Fritz, S.; See, L.; McCallum, I.; You, L.; Bun, A.; Moltchanova, E.; Duerauer, M.; Albrecht, F.; Schill, C.; Perger, C.; *et al.* Mapping global cropland and field size. *Glob. Chang. Biol.* **2015**, doi:10.1111/gcb.12838.
28. Latham, J.; Cumani, R.; Rosati, I.; Bloise, M. *Global Land Cover SHARE (GLC-SHARE) Database Beta-Release Version 1.0*; Global Land Cover Network: Rome, Italy, 2014.
29. Eidenshink, J.C.; Faundeen, J.L. The 1 km AVHRR global land data set: First stages in implementation. *Int. J. Remote Sens.* **1994**, *15*, 3443–3462.
30. Tateishi, R.; Hoan, N.T.; Kobayashi, T.; Alsaaidh, B.; Tana, G.; Phong, D.X. Production of Global Land Cover Data–GLCNMO2008. *J. Geogr. Geol.* **2014**, *6*, doi:10.5539/jgg.v6n3p99.
31. Blanco, P.D.; Colditz, R.R.; Saldaña, G.L.; Hardtke, L.A.; Llamas, R.M.; Mari, N.A.; Fischer, A.; Caride, C.; Aceñolaza, P.G.; Del Valle, H.F.; *et al.* A land cover map of Latin America and the Caribbean in the framework of the SERENA project. *Remote Sens. Environ.* **2013**, *132*, 13–31.
32. Verhegghen, A.; Mayaux, P.; De Wasseige, C.; Defourny, P. Mapping Congo Basin vegetation types from 300 m and 1 km multi-sensor time series for carbon stocks and forest areas estimation. *Biogeosciences* **2012**, *9*, 5061–5079.
33. Miettinen, J.; Shi, C.; Tan, W.J.; Liew, S.C. 2010 land cover map of insular Southeast Asia in 250-m spatial resolution. *Remote Sens. Lett.* **2012**, *3*, 11–20.
34. Klein, I.; Gessner, U.; Kuenzer, C. Regional land cover mapping and change detection in Central Asia using MODIS time-series. *Appl. Geogr.* **2012**, *35*, 219–234.
35. Sulla-Menashe, D.; Friedl, M.A.; Krankina, O.N.; Baccini, A.; Woodcock, C.E.; Sibley, A.; Sun, G.; Kharuk, V.; Elsakov, V. Hierarchical mapping of Northern Eurasian land cover using MODIS data. *Remote Sens. Environ.* **2011**, *115*, 392–403.
36. 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 30m resolution: A POK-based operational approach. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 7–27.
37. Takahashi, M.; Nasahara, K.N.; Tadono, T.; Watanabe, T.; Dotsu, M.; Sugimura, T.; Tomiyama, N. *JAXA High Resolution Land-Use and Land-Cover Map*; Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International: Melbourne, Australia, July 2013, 2384–2387.

38. Liu, J.; Liu, M.; Tian, H.; Zhuang, D.; Zhang, Z.; Zhang, W.; Tang, X.; Deng, X. Spatial and temporal patterns of China's cropland during 1990–2000: An analysis based on Landsat TM data. *Remote Sens. Environ.* **2005**, *98*, 442–456.
39. Lymburner, L.; Australia, G. *The National Dynamic Land Cover Dataset*; Geoscience Australia: Canberra, Australia, 2011.
40. Sreenivas, K.; Sekhar, N.S.; Saxena, M.; Paliwal, R.; Pathak, S.; Porwal, M.; Fyzee, M.; Rao, S.K.; Wadodkar, M.; Anasuya, T.; et al. Estimating inter-annual diversity of seasonal agricultural area using multi-temporal resourcesat data. *J. Environ. Manag.* **2015**, *161*, 433–442.
41. Holecz, F.; Collivignarelli, F.; Barbieri, M.; Gatti, L.; Boschetti, M.; Manfron, G.; Brivio, P.A.; Abukari, M.; Bondo, T. *Establishing National Baseline Land Cover Map Including Annual and Seasonal Variations for the Understanding of Current Agricultural Practices in the Gambia*; National Agricultural Land and Water Management Development Project (Nema); 2013; unpublished.
42. Lavreniuk, M.; Kussul, N.; Skakun, S.; Shelestov, A.; Yailymov, B. Regional retrospective high resolution land cover for Ukraine: methodology and results (Submitted). In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2015), Milan, Italy, 26–31 July 2015.
43. Bartalev, S.A.; Plotnikov, D.E.; Loupian, E.A. Mapping of arable land in Russia using multiyear time series of MODIS data and the LAGMA classification technique. *Rem. Sens. Let.* **2016**, *7*, 269–278.
44. 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.
45. 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.
46. Stehman, S.V.; Olofsson, P.; Woodcock, C.E.; Herold, M.; Friedl, M.A. A global land-cover validation data set, II: Augmenting a stratified sampling design to estimate accuracy by region and land-cover class. *Int. J. Remote Sens.* **2012**, *33*, 6975–6993.
47. Fritz, S.; McCallum, I.; Schill, C.; Perger, C.; Grillmayer, R.; Achard, F.; Kraxner, F.; Obersteiner, M. Geo-Wiki.org: The use of crowdsourcing to improve global land cover. *Remote Sens.* **2009**, *1*, 345–354.
48. Fritz, S.; McCallum, I.; Schill, C.; Perger, C.; See, L.; Schepaschenko, D.; van der Velde, M.; Kraxner, F.; Obersteiner, M. Geo-Wiki: An online platform for improving global land cover. *Environ. Model. Softw.* **2012**, *31*, 110–123.
49. Mourão, K.R.M.; Sousa Filho, P.W.M.; de Oliveira Alves, P.J.; Frédou, F.L. Priority areas for the conservation of the fish fauna of the Amazon Estuary in Brazil: A multicriteria approach. *Ocean Coast. Manag.* **2014**, *100*, 116–127.
50. Waldner, F.; Fritz, S.; Di Gregorio, A.; Defourny, P. Mapping priorities to focus cropland mapping activities: Fitness assessment of existing global, regional and national cropland maps. *Remote Sens.* **2015**, *7*, 7959–7986.
51. Committee on Earth Observation Satellites. *CEOS Acquisition Strategy for GEOGLAM Phase 1—Version 1.0*; 2013.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).