Global Food Security 4 (2015) 37–45

Contents lists available at ScienceDirect

Global Food Security

journal homepage: www.elsevier.com/locate/gfs

Improved global cropland data as an essential ingredient for food security

Linda See [a,*], Steffen Fritz [a], Liangzhi You [b,i], Navin Ramankutty [c], Mario Herrero [d], Chris Justice [e], Inbal Becker-Reshef [e], Philip Thornton [f], Karlheinz Erb [g], Peng Gong [h], Huajun Tang [i], Marijn van der Velde [a], Polly Ericksen [f], Ian McCallum [a], Florian Kraxner [a], Michael Obersteiner [a]

[a] International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria
[b] International Food Policy Research Institute (IFPRI), Washington DC, USA
[c] The University of British Columbia, Vancouver, Canada
[d] Commonwealth Scientific and Industrial Research Organisation (CSIRO), St. Lucia, QLD, Australia
[e] University of Maryland, College Park, MD, USA
[f] CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), International Livestock Research Institute (ILRI), Nairobi, Kenya
[g] Institute of Social Ecology, Vienna; Alpen-Adria-University Klagenfurt, Graz, Austria
[h] University of California, Berkeley, USA; Tsinghua University, Beijing, China
[i] Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing, China

ARTICLE INFO

Article history:
Received 8 June 2014
Received in revised form
30 September 2014
Accepted 3 October 2014

Keywords:
Cropland
Land cover
Food security
Crowdsourcing
Data sharing

ABSTRACT

Lack of accurate maps on the extent of global cropland, and particularly the spatial distribution of major crop types, hampers policy and strategic investment and could potentially impede efforts to improve food security in an environment characterized by continued market volatility and a changing climate. Here we discuss the pressing need for the provision of spatially explicit cropland datasets at a global scale and review the strengths and weaknesses of the various approaches used to develop such data.

1. Introduction

Ensuring food security from land that is increasingly under pressure is a key challenge of this century. By 2050 the global population will exceed 9 billion (Roberts, 2011), and with the growing wealth of populous low- and middle-income countries, a 60% to 70% increase in annual agricultural output is required (FAO, 2009; Alexandratos and Bruinsma, 2012), a rise unprecedented in human history.

Food security is monitored in near-real time by different organizations and initiatives at the international, regional and national scale, e.g. the Group on Earth Observations Global Agricultural Monitoring (GEOGLAM) initiative at the global level and by the agricultural departments of many countries. Along with other sources of information, such as road networks and market prices, forecasts of crop production are needed in order to anticipate production shortfalls. Production is estimated from yield and cropland area, which are often obtained through interviews with farmers or from agricultural surveys, where both methods have problems, e.g. area can simply be estimated as a difference from the previous year, leading to biases over time (Jayne and Rashid, 2010). At a national level, cropland area and yield are needed in order to make decisions about how much food is to be stored, distributed or exported and to make an assessment of food losses along the food supply chain. Hence, wrong trade decisions can lead to unwanted price fluctuations and food shortages. Jayne and Rashid (2010) provide a hypothetical example of how overestimating production by 13% and underestimating consumption by 8% can lead to a potentially disastrous shortfall in food of 21%, which could then lead to sharp rises in food prices if no food aid or trade was present in this situation. The authors then provide a real life example of how overestimates in maize

* Corresponding author.
E-mail address: see@iiasa.ac.at (L. See).

http://dx.doi.org/10.1016/j.gfs.2014.10.004
2211-9124/

surpluses reported in Malawi in 2007 led to maize prizes reaching record highs when that forecasted surplus failed to materialize.

At the global level, detailed crop yield and crop extent information can help to identify where investment from large donors might be most effective in terms of boosting agricultural output. At the regional level, this information can be used to help understand the impact of drought and other natural and manmade disasters on food production (Funk and Brown, 2009). At a national and sub-national level, accurate information on cropland can be used to measure trends in agricultural outputs and to evaluate if investments have led to the expected results.

More accurate cropland data are also required to address the multi-dimensional challenges related to global environmental change. To meet the growing demand for food in the future, agricultural land will either expand or production will be intensified in areas where there are current yield gaps or through new innovations. There are environmental impacts and tradeoffs associated with both of these pathways that need to be better understood if the effects are to be minimized (Tilman et al., 2011). Nevertheless, it is generally argued that large expansions of existing cropland are more disadvantageous than intensification, and it is essential to know where and how to increase crop yield on existing cropland areas (Foley et al., 2005; Tilman et al., 2002; Van Wart et al., 2013). Therefore, having accurate data on current cropland extent is critical for undertaking these types of analyses. However, the uncertainties in the available datasets on cropland are currently too high for use in many applications. A recent study showed that current global estimates of the amount of land under crop production vary by 300 Mha, i.e. around 1600 Mha from one global land cover product compared to around 1300 Mha from another (Fritz et al., 2011a). This variation introduces uncertainties in considering how other important drivers of change in agricultural systems, such as biofuel production, rising demand for livestock products, and expanding urban areas, might affect food production. The issue is aptly illustrated in a recent study by Smith et al. (2010), who compared the outputs from a number of integrated assessment models regarding the global change in cropland area as well as other land-use types in 2020 and 2050 under numerous drivers of change. Many of the predictions of cropland change from these models are within the 300 Mha range of uncertainty regarding total cropland area estimates from global land cover maps, some of which are used as inputs to these models. Other studies have shown that model outputs and analyses can vary substantially depending upon which land cover product has been used (Ge et al., 2007; Linard et al., 2010; Quaife et al., 2008). Examples such as these illustrate the need for an improved global spatially explicit cropland map, which is useful at multiple scales from global monitoring and assessment to planning at the national and sub-national levels.

This paper reviews current and emerging approaches for developing global cropland maps including an overview of their strengths and weaknesses. These include a range of options such as the use of satellite imagery, agricultural census and survey statistics, the incorporation of crowdsourcing approaches as well as hybrid methods. We further discuss the need to harmonize definitions of cropland, share data more openly and target new mapping efforts, which could yield substantial benefits for improving food security in the future.

2. Current and emerging approaches for developing global cropland maps

With the recognition that global land use and land cover change is a major driver of global environmental change (Foley et al., 2005), there have been numerous efforts to map land cover and its change globally. We have characterized these into five distinct approaches, where each produces different types of cropland extent information at varying spatial resolutions as outlined in Table 1. We also recognize that these approaches have strengths and limitations and have compared them based on whether they are consistent with FAO statistics, their relative costs, the accuracy of the products, the temporal frequency of production and updating, and any other issues related to these approaches; these are summarized in Table 2 and discussed in the sections that follow.

2.1. Global agricultural census- and survey-based statistics

A frequently used, standard, globally-complete, source of information on cropland extent is the Food and Agriculture Organization's (FAO) compilation of statistics reported by individual countries, which are based on censuses, agricultural samples and questionnaire-based surveys with major agricultural producers (FAO, 1996). These data are publicly available from the FAOSTAT database (http://faostat.fao.org/) from 1961 onward and are reported at the national level. They are frequently updated, often with revisions to the entire time-series. At their best, these data are based on comprehensive agricultural censuses conducted every 5 years and in some countries with conflicts or poor infrastructure, there is a lack of regular censuses. The database reports detailed land cover and land use statistics, and in addition to the extent of cropland, pasture and other land covers, provides land use information such as irrigation extent, fertilizer application rates, and mechanization. Some ongoing efforts aim to compile census and survey based information at the subnational level from individual country statistics/census reports (e.g., The Agro-MAPS project (http://kids.fao.org/agromaps/); Ramankutty et al., 2008; Monfreda et al., 2008). However cropland information at the sub-national scale is even scarcer, plagued by data gaps, and are tedious to compile. Moreover, independent evaluations (FAO, 2006; The World Bank, 2010) have recognized that there are both quality- and quantity-related problems in the agricultural information provided by different reporting countries, particularly those in Africa. Among several pitfalls of relying on national

Table 1
Cropland information that is produced by the five approaches.

Approach	Cropland information produced by the approach
Global agricultural census-and survey-based methods (Section 2.1)	Farm census or sample-survey estimates, reported at the sub-national level (by individual countries) or national level (reported by FAO)
Satellite-based global land cover classification (Section 2.2)	Global maps of presence/absence of cropland or the percentage cropland at resolutions of 30 m to 1 km. National and regional maps are also produced using satellite-based land cover classification.
Blending census and satellite data (Section 2.3)	Maps of presence/absence of cropland or percentage cropland at resolutions of 1 km to 10 km that are calibrated to FAO statistics and other agricultural census data
Use of crowdsourced data (Section 2.4)	Samples of varying resolutions (250 m to 1 km) of percentage cropland (or presence/absence of cropland) which can then be interpolated to create maps of cropland at a resolution that matches the sample size, e.g. 1 km
Synergy map that blends remote sensing, crowdsourcing and census (Section 2.5)	Similar approach to blending the census and satellite data except that existing maps (global, regional and national) are integrated to produce a hybrid product. The input maps are ranked based on correspondence with data collected from crowdsourcing of high resolution imagery.

Table 2
Comparison of various approaches to mapping agricultural land cover.

Approach	Consistent with FAO statistics	Relative costs	Accuracy	Temporal frequency	Other issues
Global agricultural census- and survey-based methods (Section 2.1)	Yes	High	Accuracy assessments not available. But quality is highly variable because of inconsistency in definitions across countries and time periods, data gaps, under/over reporting incentives, different types of census/survey methods used	Annual, with long historical records but lack of regular censuses in some countries with conflicts or lack of infrastructure	Provides additional thematic information on land use and agricultural inputs, e.g. fertilizers, mechanization, etc.
Satellite-based global land cover classification (Section 2.2)	No	High	Variable due to sensor misclassification errors (e.g. fallow land vs. bare soil). Reported Producer/User accuracies of 74/88% (rainfed croplands) and 55/81% (post-flooding or irrigated croplands) for GlobCover 2009 (Bontemps et al., 2011) and 83/93% (cropland) and 61/28% (cropland/natural vegetation mosaic) for MODIS Collection 5 (Friedl et al., 2010). However, there is large disagreement between different satellite products in some regions (Fritz et al., 2011a, 2011b).	Frequent and consistent measurements so could be updated often	Limited in thematic resolution; moderate-resolution global data are often inadequate in terms of spectral and spatial resolution; definitions of cropland vary between map producers; cloud cover
Blending census and satellite data (Section 2.3)	Yes and No, depends on approach used	Low	Accuracy assessments not available. More consistent with the census/survey data than individual global land cover maps used but new errors may be introduced due to inconsistencies between definition of cropland in satellite and census product.	Created on an ad hoc basis	
Use of crowdsourced data (Section 2.4)	No	Low	89% for Ethiopia example (See et al., 2013) but global maps need to be created	Created on an ad hoc basis	Data reliability and quality issues; images are often from different time periods; only a sample rather than comprehensive coverage; needs a volunteer community with incentives to participate; issues of sustainability beyond the running of finite crowdsourcing campaigns; participation can be open to all and therefore not exploit local expertise
Synergy map that blends remote sensing, crowdsourcing and census (Section 2.5)	Yes	Low	Around 83% (Fritz et al., submitted) but further improvements still needed	Created on an ad hoc basis	Input maps are made using different methodologies and for different time periods so this will introduce some errors into the map

reports, there are no attempts made to harmonize data sources and collection methods among different countries, and there are issues regarding the quality and accuracy of the data (The World Bank, 2010). For example, in 2008/09 in Malawi, cropland extent was estimated by combining household surveys with field measurements derived from a "pacing method" in which the size of crop fields is determined by the number of steps required to walk around them. However, this is a highly inaccurate and outdated method and farmer estimates of area were shown to be 30% higher than enumerator estimates (Dorward and Chirwa, 2010). Different agricultural survey methods used in Ethiopia result in production estimates of different crops that vary between 29% and 44% (Alemu et al., 2008). Another representative example of the very high uncertainties around crop area can be found in Uganda, where numbers from the Ministry of Agriculture (Uganda Bureau of Statistics (UBOS), 2007; FAO, 2012) and the officially reported FAO numbers from FAOSTAT in 2006 differ by 15% for maize and 75% for soya beans. Finally, the costs of this approach are relatively high in relation to the other approaches since this method requires surveys to be undertaken by individual countries and the further involvement of FAO in compiling, endorsing and reporting these statistics.

2.2. Satellite-based global land-cover classification maps

Cropland information is also available from remote sensing. While remote sensing data provide an objective, frequent and consistent measure of what the Earth's land cover looks like, they are limited by having to see through clouds, dust and other atmospheric constituents from more than 700 km above the surface of the Earth. Moreover, the operational global products (e.g., based on MODIS) have a spatial resolution of 500 m at best (250 m pan-chromatic), and have limited ability to capture land use patterns in complex landscapes (Ozdogan and Woodcock, 2006). Until recently, global land cover products were developed using moderate-resolution satellite imagery but with the opening up of the Landsat archive in 2008 (Wulder et al., 2012), higher resolution global land cover maps are now starting to appear, as described below. The relative costs of this approach are also high, i.e. the investment in satellite technology and the costs associated with image acquisition. However, Landsat is now free but higher resolution imagery, i.e. less than 10 m, is still costly.

2.2.1. Data from moderate-resolution satellites

The three most recent products that provide information on global cropland extent are: GLC-2000, which was produced by the Joint Research Centre (JRC) of the European Commission as a one-off product using SPOT Vegetation as baseline land cover for the year 2000; NASA's MODIS land cover obtained via the MODerate Resolution Imaging Spectroradiometer sensor (Friedl et al., 2010), which is produced on an annual basis; and the GlobCover maps produced by the European Space Agency's Medium Resolution Image Spectrometer (MERIS) (Bontemps et al., 2011) for 2005 and 2009. The problem with these products is that they are not accurate enough to provide a reliable estimate of croplands. For example, they are particularly poor at detecting croplands in areas of low agricultural intensification because the spectral signatures and temporal profiles are similar to grasslands; this would include most of Africa (Pittman et al., 2010). Other global cropland products are available (e.g. Biradar et al., 2009) but these are at an even coarser resolution of 10 km^2.

This lack of detail has led to the highly divergent estimates that one sees in today's cropland and land cover maps. Globally, cropland estimates derived from GlobCover are 20% higher than those derived from MODIS (Fritz et al., 2011a, 2011b). Regionally, the disagreement between estimates can be particularly large. For example, FAO estimates that there are 319 Mha of cropland in Africa (FAO, 2005) compared to the lower MODIS and GlobCover estimates of 277 Mha and 152 Mha, respectively. These discrepancies can be attributed to the use of different classification algorithms, different datasets used to train the algorithms, different satellite sensors, and different temporal windows used to develop these products, i.e., some use a single reference year while others use multiple years. The product that is used in subsequent monitoring and modeling exercises can therefore have a potentially large impact on the outcome.

2.2.2. Use of high-resolution satellite imagery

There are several satellite sources (e.g., Landsat and Sentinel, or the even more finely granulated satellite imagery such as IKONOS/Quickbird or GeoEye) that produce imagery of sufficient detail to capture croplands in regions with low agricultural intensification. A reasonable goal would be a cropland map that provides detail down to at least 30 m per pixel, if not lower, as the complexities of today's land-use policy considerations demand increasingly precise and accurate mapping. For example, there has been criticism that a 2011 forest survey in India underestimated tree cover loss even though it used satellite imagery at a resolution of 23.5 m per pixel (Gilbert, 2012). Critics say the survey should use images now available that go down to 5.8 m per pixel. There is even discussion of seeking satellite imagery where resolutions are measured in cm per pixel. Although remote sensing is advancing rapidly and cropland maps should improve accordingly, the quest for a more finely-detailed cropland map is a long-term solution and is not likely to provide improved products over the short-term. Although 30 m land cover maps are now starting to appear, their overall accuracy is still short of that required for food security applications (Gong et al., 2013; Yu et al., 2013a, 2013b). Moreover, high-resolution imagery may be needed to identify croplands but the final resolution of a useful global cropland extent map could still be medium resolution, such as 250 m. More accurate higher resolution regional products such as the EU's CORINE (Coordination of Information on the Environment) land cover product (Steemans, 2008) and AFRICOVER (FAO, 1998), an initiative which produced land cover products for a small number of African countries, are also available but they do not provide global coverage.

2.3. Blending census and satellite data

Another approach taken to map croplands globally blends remote sensing and ground-based census statistics (e.g. Ramankutty and Foley, 1998; Ramankutty et al., 2008) which are relatively low cost solutions since the main costs have already been absorbed through implementation of the approaches described in Sections 2.1 and 2.2. While the ground-based agricultural and sample survey statistics provide richer thematic detail, they are limited as the data are reported by administrative units, and because of reporting errors and important data gaps (see Section 2.1 for more details). Several approaches have been developed to statistically blend the satellite and census data (e.g., Hurtt et al., 2001; Cardille and Clayton, 2007; Ramankutty et al., 2008), with the aim of developing a single consistent product that combines the best aspects of both. The final products have coarse spatial resolution, typically 10 km, but the focus of these products is to increase accuracy in mapping cropland extent globally rather than fine-grained spatial precision. However, these approaches also have important limitations. To take a couple of examples, tree crops such as oil palm may be classified as tree cover by satellite data and croplands by census data; or fallow land may be classified as bare ground by satellite data but cropland by census data. Although the statistical methods attempt to correct for potential misclassifications by satellite data, important inconsistencies will nevertheless introduce major errors. In other words, approaches that blend the best of two different sources of information may also inadvertently introduce new errors due to inconsistencies. Moreover, these products tend to be produced for a baseline year so they are not updated on a regular basis.

2.4. Direct use of crowdsourced data

Crowdsourcing involves using citizens and interested experts to help collect and analyze data (Howe, 2006), which can then be used for scientific purposes. Crowdsourcing competitions are run by the International Institute for Applied Systems Analysis (IIASA) throughout the year whereby participants examine high resolution satellite imagery and indicate the location of cultivation and other types of land cover using Geo-Wiki crowdsourcing technology (Fritz et al., 2012). These participants are experts in remote sensing and geospatial sciences, scientists from a range of other disciplines, students and ordinary citizens, who have registered to participate in Geo-Wiki campaigns. Recruitment has been through mailing lists, conferences, academic papers, social networking, incentives offered for introducing friends, and media coverage, where the Geo-Wiki network now has around 5000 registered users. An example of the system is shown in Fig. 1 where the participant sees an area surrounded by a red rectangle and they are asked to choose the land cover type from a drop down list based on what is visible from Google Earth imagery. This process takes place online so does not require individuals to be located in those countries shown.

These crowdsourced data on land cover contribute to a growing, open source database of information that can be used to calibrate and validate land cover maps. For example, the US Agency for International Development (USAID) sponsored a crowdsourcing effort using Geo-Wiki where participants were asked to indicate the degree of cultivation and human settlements (Fig. 2). The campaign was carried out online without ground truthing so it did not require individuals to collect data on the

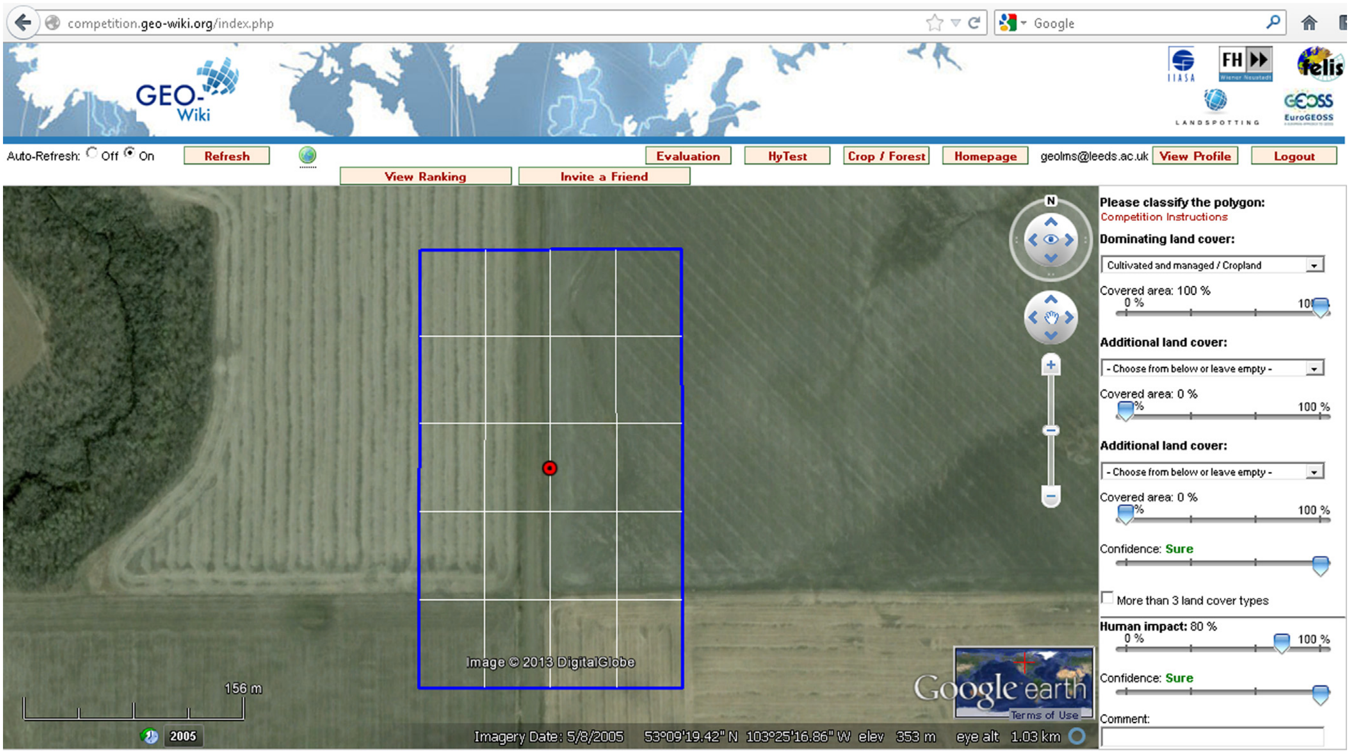

Fig. 1. Screenshot from the Geo-Wiki crowdsourcing tool for collecting information on land cover across the globe.

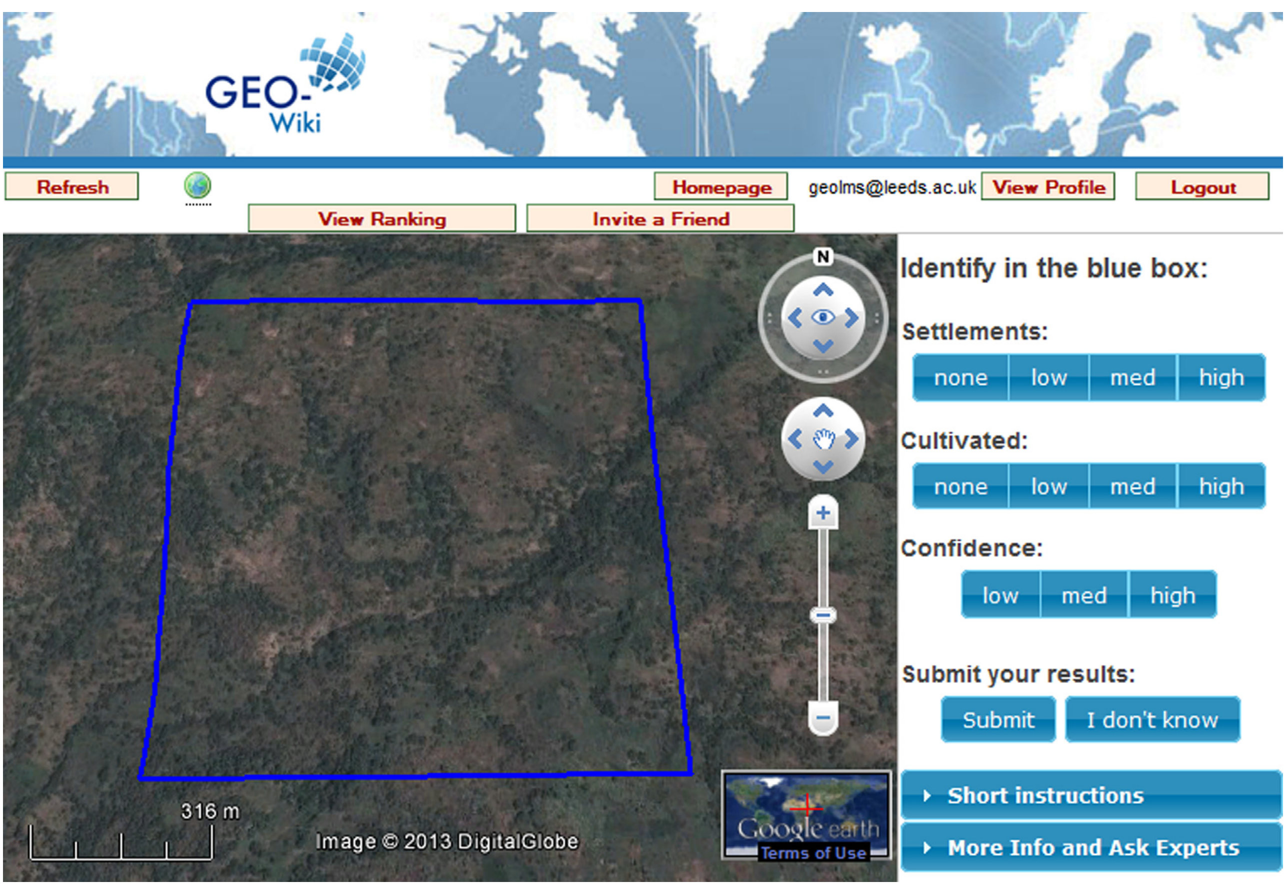

Fig. 2. Screenshot from the Geo-Wiki crowdsourcing tool for collecting information on degree of cultivation and degree of human settlement in Ethiopia.

ground in Ethiopia. Participants were provided with examples to help train them, a facebook site was set up to discuss images that were difficult to classify, and each area classified was provided to more than one individual so that majority agreement could be applied. A random sample of 1 km sized areas was created across Ethiopia. In just three weeks a sample equating to roughly five

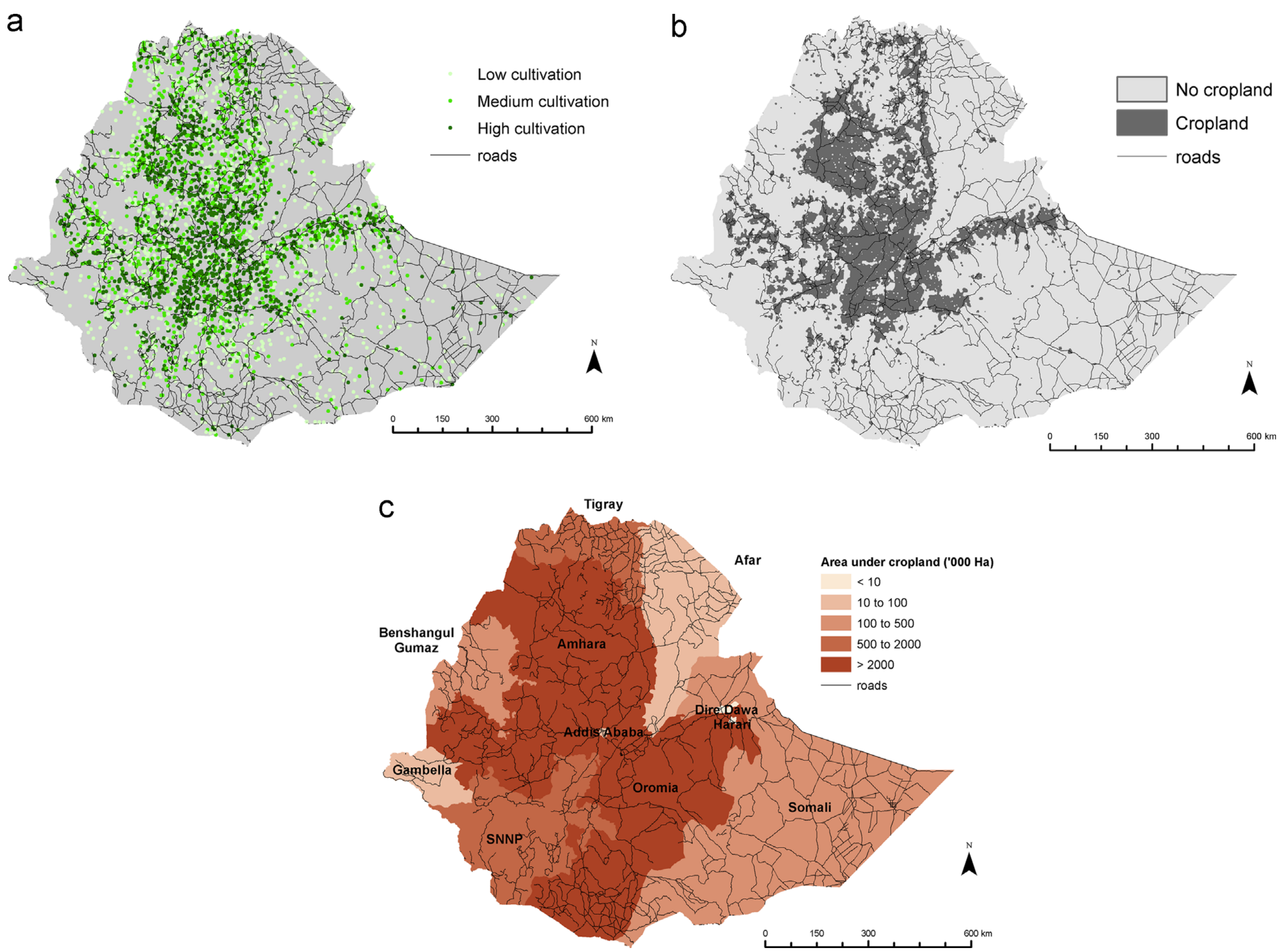

Fig. 3. (a) Data collected for Ethiopia on cropland through crowdsourcing (b) a cropland map for Ethiopia created by interpolation of crowdsourced data and (c) cropland area for Ethiopia by sub-national zone.

percent of the area of Ethiopia was collected (Fig. 3a). This sample was then turned into a gridded cropland map using a simple method of interpolation. Interpolation is a method that takes a discrete set of values (i.e. the collected sample) and determines the values at a regularly spaced grid, essentially filling in values for the whole country. The resulting cropland map is shown in Fig. 3b. This map was compared to other global land cover maps, which contain categories for cropland. Using an independent validation data set, the interpolated cropland map was shown to be more accurate than cropland extent from any of the individual global land cover maps (See et al., 2013). The point of the exercise was to show how a simple, low cost cropland map could be constructed using crowdsourcing, Google Earth and interpolation. Other than the staff time needed to run the campaign, the cost was roughly 25 Euros for a book prize awarded to the winning crowdsourcing team. In contrast, Fig. 3c shows the total amount of cropland area by sub-national zone, which is a much coarser representation of this information. Statistics at a lower level of administrative zone were not available for Ethiopia.

There are, however, limitations of a crowdsourcing approach for land cover data collection. For example, data collection tends to take place on an ad hoc basis, i.e. campaigns run for finite periods of time, so data collection is not continual in space or time or related to any update cycle. There are justifiable concerns regarding data quality and there is a growing body of literature addressing methods for improving or ensuring data quality (e.g. Alabri and Hunter, 2010; Allahbakhsh et al., 2013). The satellite imagery used by Geo-Wiki is constantly changing over time. Thus data collected by the crowd are from different time periods, which means that some errors could be introduced as a result of changes in land cover. Moreover, it is also only realistically possible to sample a relatively small portion of the Earth's surface, as spatially comprehensive data collection would require participation on the order of that found in the gaming industry or from social media. Of the 5000 registered Geo-Wiki users, only a small number actually provide the majority of the data, a phenomenon that is very common in crowdsourcing (Bryant et al., 2005). Motivating and incentivizing participation are therefore ongoing challenges, as is sustainability over the long term. Finally, crowdsourcing is often open to anyone who registers so may not benefit from local knowledge of an area. Each of these limitations has potential solutions but they are clearly quite different challenges to those of the other mapping approaches.

2.5. Developing synergy maps that blend remote sensing, crowdsourcing and census

This approach is similar to that described in Section 2.3. However, there are a number of differences. In this approach, we seek out all available cropland or land cover maps (global, regional and national), resolve disparities in their estimates, and integrate them into a single product. A workshop in 2010, led by IIASA, focused on taking a number of different national and regional land cover products and merging them with global products to produce a new global cropland map. A second difference is in the way the maps are combined, which uses crowdsourced data from Geo-Wiki to rank the accuracy of the individual map products at a country level. This product is also consistent with FAO statistics and is more accurate than current global products. Similar to the approach described in Section 2.3, the costs are relatively low and a single product was created for the year 2005. This product is openly available in beta version on http://beta-hybrid.geo-wiki.org and is currently being used by the International Food Policy Research Institute (IFPRI) to map the distribution of crop types using their SPAM model (http://www.mapSPAM.info). The hybrid map has also been recognized as an important input to GEO's Global Agricultural Monitoring (GEOGLAM) initiative (G8 Research Group, 2013). There are plans to create a second product for 2010. However, we recognize that the input maps have all been made using different methodologies and for different time periods so this will introduce some errors into the final product. For this reason we have provided an expert crowdsourcing tool within Geo-Wiki, which allows the hybrid map to be visualized on top of Google Earth, where it can be annotated using drawing tools to mark up areas that are incorrectly mapped. This ongoing feedback will be used to update the map in the future.

3. Charting the path forward

There is no single answer to solving the problem of improving global cropland data but there are simple ways in which we can move forward. From a technological standpoint, there are no significant barriers to providing more detailed insights into cropland attributes of countries vulnerable to food insecurity. Instead the issues are related to the need to harmonize the definition of cropland, share existing data more openly and target future mapping efforts in areas where this information is currently lacking.

3.1. Harmonization of cropland definitions

FAO defines agricultural land as land that is suitable for crop and livestock production, which can be further broken down into land under annual crops or land that is temporarily fallow (i.e. arable land), land under permanent crops (e.g. orchards) and permanent pastures (FAO, 2014a). However, the term cropland, which is often used to describe one or more classes in land cover maps, is not clearly defined and there is no agreement as to what this constitutes. As a starting point, we propose the definition used by IFPRI in the application of their SPAM model, which is arable land and land under permanent crops; this would exclude hay and fodder crops or managed grasslands. Agreeing upon a definition is necessary if maps shared from different sources are to be harmonized and therefore integrated. National map legends must be detailed enough to allow for exclusion of classes such as hay and fodder crops. If legend definitions are incompatible with a globally harmonized legend of cropland, then they cannot be integrated into a hybrid product. Fortunately many developing countries do use the FAO proposed LCCS (Land Cover Classification System) (e.g. AFRICOVER and ASIACOVER), which allows for a clear understanding of the class definitions employed (Di Gregorio and Jansen, 2000).

3.2. More open data sharing

The composite cropland map (section 2.5) was only possible through data sharing by different countries yet it only contains data from 25 countries worldwide (in addition to all European countries through CORINE) where we ensured that cropland was compatible with the definition provided in Section 3.1. There are clearly many countries around the world where detailed information on land cover, including croplands, has been collected but has not been made available. For example, China has shared a wealth of information, while Russia has been reluctant to provide what is believed to be extensive data sets on croplands. In Africa, there is high-quality information available for Ethiopia from the AFRICOVER initiative yet this has only recently become available in an aggregated and integrated form as part of the FAO's GLC-SHARE product (FAO, 2014b). The GLC-SHARE product is a new global land cover product that has been created by integrating existing national and regional land cover products (shared by FAO member states and from other open sources) into a single global map.

At the very least, data now being collected through multilateral organizations such as the UN and the World Bank should be widely shared. The UN, for example, is working with the European Space Agency to provide a variety of crop and land use maps for Botswana, Niger, the Gambia, and Vietnam. Aside from sharing this information with multilateral development banks, there is nothing in these projects that guarantees open access to these mapping products. However, through the concerted efforts of the United

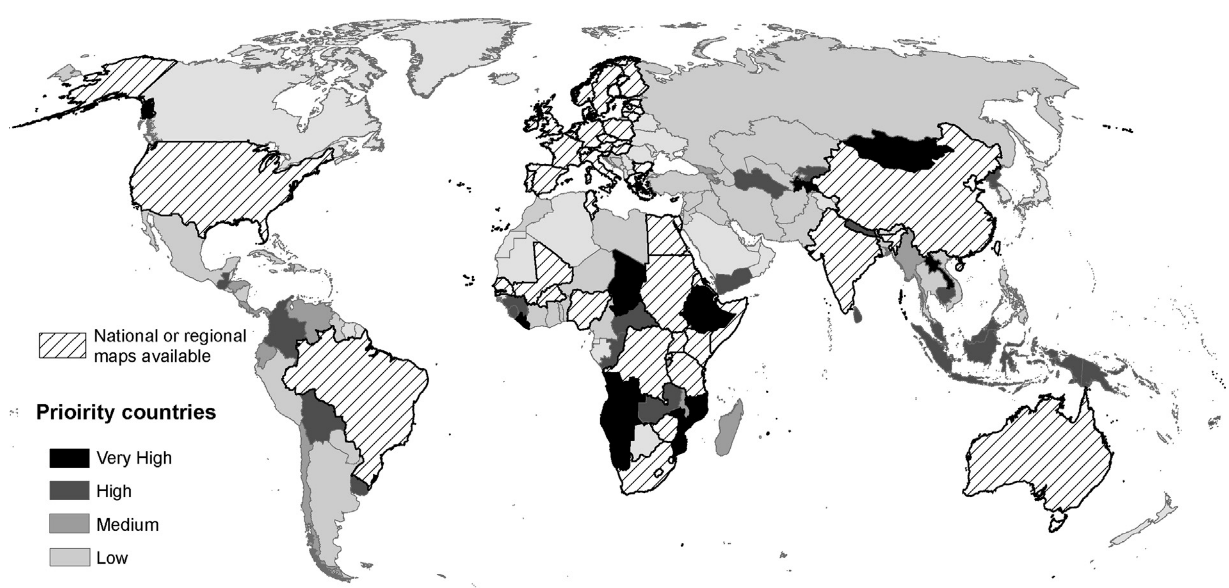

Fig. 4. Areas of priority for improving global cropland maps. Areas with diagonal hatching indicate maps that were shared at the national level or where regional products are available, e.g. CORINE and AFRICOVER.

Nations, the Group on Earth Observations, which promotes the use of Earth Observation and the sharing of data, and national governments, e.g. Data.gov (the US Government's open data website), the recent executive order in the USA and the trend towards open data policies in a number of EU countries and research-for-development organizations such as CGIAR (CGIAR, 2013), the situation is clearly improving. The sharing and integration of data is an inexpensive yet effective solution. Moreover, opening up data can lead to increased innovation and entrepreneurship along with substantial financial gains (Huijboom and Van den Broek, 2011).

3.3. Targeting efforts towards mapping the most vulnerable

As resources are finite, we need to concentrate our efforts on mapping areas that are highly vulnerable from a food security perspective and where the highest uncertainties lie. These are generally not the areas where good maps are already available. Here we present an analysis of those countries where more accurate information on cropland is needed as the only available information at present is through global cropland maps. Two criteria were used to map priority countries: (i) the level of spatial disagreement in the cropland domain between different global land cover maps, normalized by the amount of cropland according to FAO statistics (FAO, 2010); and (ii) the Global Hunger Index (GHI), which combines undernourishment, the proportion of children underweight, and child mortality into a single index (Von Grebmer et al., 2011). The criteria have been ranked by country and equally weighted to produce a mapping priority index, shown in Fig. 4, where only those countries with a minimum cropland of 500,000 ha have been chosen in order to place more emphasis on countries with larger agricultural areas.

The top 10 very high priority countries highlight those areas where mapping efforts should now be focused as we are still relying on global land cover maps for this information. Chad is at the top of the list, followed by Liberia, Ethiopia, Angola, Guinea-Bissau, Mozambique, Lao PDR, Mongolia, Namibia and Tajikistan. Areas of high priority indicate the next 20 countries that require attention, concentrated mostly in the lower latitudes. There are clearly other countries where food security is an issue but information beyond global land cover maps are available so they may not appear in this list.

4. The bottom line

Effort needs to be directed towards mapping those countries identified in Section 3.3, building sustainable crowdsourcing tools underpinned by a motivated community of volunteers and citizen scientists, and in building relationships at the national level through GEO in order to share data more freely. However, moving towards a high-quality cropland product that is updated on a regular basis (perhaps annually) will still require input from remote sensing. Given the availability of high and very high resolution satellite imagery, any long-term solution should look to capitalize on these data-rich information streams. This pressing need has been recognized recently in the context of monitoring deforestation (Lynch et al., 2013). A remote sensing solution will be more expensive than other solutions, but this can easily be put into perspective. The amount of money that has been invested in wall-to-wall mapping exercises, i.e. complete mapping of the land surface in the past, has been relatively small compared with the costs that are incurred as a result of droughts or mitigation of climate change in the agricultural sector. The cost of creating a global land cover map is on the order of US$ 1 million based on discussions with producers. These amounts are small when viewed in light of the costs incurred by droughts, e.g. in Kenya, the droughts in 2008–2011 were estimated to cost around US$ 12 billion Similarly, these costs are dwarfed when compared with the UNFCCC projected cost of adapting to climate change in developing countries of anywhere between US$ 28–67 billion per year by 2030, where an additional US$ 7 billion will be required per year just to keep up with climate change in the agricultural sector alone (Parry et al., 2009). Thus, strategic investment to improve global cropland information is urgently called for.

With higher quality and more accurate maps on the spatial and temporal patterns of agricultural land use, decision-makers at the global, regional and national level will be much better placed to reliably evaluate land-use policies, explore long-term sustainability alternatives, and assess the impact of climate change as well as the trade-offs between food production and the provision of other ecosystem services. Consolidating and providing accurate cropland data at the global level, and targeting more accurate and precise map production for those countries where this information is currently lacking, e.g. in sub-Saharan Africa, will permit informed public and private investment decisions which in turn will contribute to enhanced food security for the world's rural and urban poor. Such value goes beyond any measurement of monetary returns.

Acknowledgments

This research was supported by the International Food Policy Research Institute (IFPRI), the International Institute on Livestock Research (ILRI), the CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), and the EU FP7 funded SIGMA (No. 603719) and ERC CrowdLand (No. 612755) projects.

References

Alabri, A.A., Hunter, J., 2010. Using trust metrics to enhance the quality and reliability of citizen science data. In: Proceedings of the 4th eResearch Australasia Conference, Gold Coast, Australia.

Alemu, D., Ayele, G., Behute, B., Beyone, Y., Dewana, R., Fedaku, B., Vargas Hill, R., Minot, N., Rashid, S., Taffesse, A., Tefera, N., 2008. Develoment of a spatial equilibrium modelling approach to study the impact of policy intervention on cereals availability (JRC Scientific and Technical Reports), Cereals Availability Study in Ethiopia, 2008. JRC, Ispra, Italy.

Alexandratos, N., Bruinsma, J., 2012. World agriculture towards 2030/2050. In: Proceedings of the 2012 Revision (ESA Working Paper No. 12-03). Food and Agriculture Organization of the United Nations, Rome, Italy.

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

Biradar, C.M., Thenkabail, P.S., Noojipady, P., Li, Y., Dheeravath, V., Turral, H., Velpuri, M., Gumma, M.K., Gangalakunta, O.R.P., Cai, X.L., Xiao, X., Schull, M. A., Alankara, R.D., Gunasinghe, S., Mohideen, S., 2009. A global map of rainfed cropland areas (GMRCA) at the end of last millennium using remote sensing. Int. J. Appl. Earth Obs. Geoinf. 11, 114–129. http://dx.doi.org/10.1016/j.jag.2008.11.002.

Bontemps, S., Defourny, P., van Bogaert, E., Arino, O., Kalogirou, V., Perez, J.R., 2011. GLOBCOVER 2009: Products description and validation report.

Bryant, S.L., Forte, A., Bruckman, A., 2005. Becoming Wikipedian: Transformation of Participation in a Collaborative Online, Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work, New York NY, pp. 1–10. http://dx.doi.org/10.1145/1099203.1099205.

Cardille, J.A., Clayton, M.K., 2007. A regression tree-based method for integrating land-cover and land-use data collected at multiple scales. Environ. Ecol. Stat. 14, 161–179. http://dx.doi.org/10.1007/s10651-007-0012-5.

CGIAR, 2013. Making progress towards Open Access for CGIAR [WWW Document]. URL 〈http://www.cgiar.org/consortium-news/progress-towards-open-access-for-cgiar/〉 (accessed 08.08.13.).

Di Gregorio, A., Jansen, L., 2000. Land Cover Classification System: Classification Concepts and User Manual. Food and Agriculture Organization of the United Nations, Rome, Italy.

Dorward, A., Chirwa, E., 2010. A review of methods for estimating yield and production impacts.

FAO, 1996. Conducting Agricultural Censuses and Surveys (FAO Statistical Development Series No. 6). FAO UN, Rome, Italy.

FAO, 1998. Land Cover and Land Use: The FAO AFRICOVER Programme.

FAO, 2005. FAOSTAT agricultural data.

FAO, 2006. Independent External Evaluation of the FAO. Food and Agriculture Organization of the United Nations, Rome, Italy (Inception Report Submitted to

the Council Committee for the Independent External Evaluation (CC-IEE) by the Independent External Evaluation Core Team).
FAO, 2009. How to Feed the World in 2050. Report from the High-Level Expert Forum.
FAO, 2010. FAOSTAT agricultural data.
FAO, 2012. FAOSTAT.
FAO, 2014a. Glossary List [WWW Document]. FAOSTAT Methods & Standards. URL ⟨http://faostat3.fao.org/faostat-gateway/go/to/mes/glossary/*/E⟩.
FAO, 2014b. Global Land Cover-SHARE of year 2014 – Beta-Release 1.0 [WWW Document]. FAO Global Land Cover Network (GLCN). URL ⟨http://www.glcn.org/databases/lc_glcshare_en.jsp⟩.
Foley, J.A., DeFries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E.A., Kucharik, C.J., Monfreda, C., Patz, J.A., Prentice, I.C., Ramankutty, N., Snyder, P.K., 2005. Global consequences of land use. Science 309, 570–574. http://dx.doi.org/10.1126/science.1111772.
Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets. Remote Sens. Environ. 114, 168–182. http://dx.doi.org/10.1016/j.rse.2009.08.016.
Fritz, S., See, L., McCallum, I., You, L., Bun, A., Albrecht, F., Schill, C., Perger, C., Duerauer, M., Havlik, P., Mosnier, A., Thornton, P., Wood-Sichra, U., Herrero, M., Becker-Reshef, I., Justice, C., Hansen, M., Gong, P., Abdel Aziz, S., Cipriani, A., Cumani, R., Cecchi, G., Conchedda, G., Ferreira, S., Gomez, A., Haffani, M., Kayitakire, F., Malanding, J., Mueller, R., Newby, T., Nonguierma, A., Olusegun, A., Ortner, S., Ram, R., Rocha, J., Schepaschenko, D., Schepaschenko, M., Terekhov, A., Tiangwa, A., Vancutsem, C., Vintrou, E., Wenbin, W., van der Velde, M., Dunwoody, A., Kraxner, F., Obersteiner, M., submitted. Mapping global cropland and field size. Glob. Change Biol.
Fritz, S., McCallum, I., Schill, C., Perger, C., See, L., Schepaschenko, D., van der Velde, M., Kraxner, F., Obersteiner, M., 2012. Geo-Wiki: an online platform for improving global land cover. Environ. Model. Softw. 31, 110–123. http://dx.doi.org/10.1016/j.envsoft.2011.11.015.
Fritz, S., See, L., McCallum, I., Schill, C., Obersteiner, M., Velde, M., van der, Boettcher, H., Havlík, P., Achard, F., 2011a. Highlighting continued uncertainty in global land cover maps for the user community. Environ. Res. Lett. 6, 044005. http://dx.doi.org/10.1088/1748-9326/6/4/044005.
Fritz, S., You, L., Bun, A., See, L., McCallum, I., Schill, C., Perger, C., Liu, J., Hansen, M., Obersteiner, M., 2011b. Cropland for sub-Saharan Africa: a synergistic approach using five land cover data sets. Geophys. Res. Lett. 38, L04404. http://dx.doi.org/10.1029/2010GL046213.
Funk, C.C., Brown, M.E., 2009. Declining global per capita agricultural production and warming oceans threaten food security. Food Secur. 1, 271–289. http://dx.doi.org/10.1007/s12571-009-0026-y.
G8 Research Group, 2013. G8 – The UK Summit: Lough Erne. Newsdesk Media, London UK.
Ge, J., Qi, J., Lofgren, B.M., Moore, N., Torbick, N., Olson, J.M., 2007. Impacts of land use/cover classification accuracy on regional climate simulations. J. Geophys. Res., 112. http://dx.doi.org/10.1029/2006JD007404.
Gilbert, N., 2012. India's forest area in doubt. Nature 489, 14–15. http://dx.doi.org/10.1038/489014a.
Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., Li, C., Li, X., Fu, W., Liu, C., Xu, Y., Wang, X., Cheng, Q., Hu, L., Yao, W., Zhang, H., Zhu, P., Zhao, Z., Zhang, H., Zheng, Y., Ji, L., Zhang, Y., Chen, H., Yan, A., Guo, J., Yu, L., Wang, L., Liu, X., Shi, T., Zhu, M., Chen, Y., Yang, G., Tang, P., Xu, B., Giri, C., Clinton, N., Zhu, Z., Chen, J., Chen, J., 2013. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. Int. J. Remote Sens. 34, 2607–2654. http://dx.doi.org/10.1080/01431161.2012.748992.
Howe, J., 2006. The rise of crowdsourcing. Wired Magazine. 14.06, ⟨http://archive.wired.com/wired/archive/14.06/crowds.html⟩.
Huijboom, N., Van den Broek, T., 2011. Open data: an international comparison of strategies. Eur. J. ePract., 12.
Hurtt, G.C., Rosentrater, L., Frolking, S., Moore, B., 2001. Linking remote-sensing estimates of land cover and census statistics on land use to produce maps of land use of the conterminous United States. Glob. Biogeochem. Cycles 15, 673–685. http://dx.doi.org/10.1029/2000GB001299.
Jayne, T.S., Rashid, S., 2010. The value of accurate crop production forecasts., in: Agricultural Risks Management in Africa: Taking Stock of What Has and Hasn't Worked. Presented at the Fourth African Agricultural Markets Program (AAMP) Policy Symposium, Lilongwe, Malawi.
Linard, C., Gilbert, M., Tatem, A.J., 2010. Assessing the use of global land cover data for guiding large area population distribution modelling. GeoJournal 76, 525–538. http://dx.doi.org/10.1007/s10708-010-9364-8.
Lynch, J., Maslin, M., Balzter, H., Sweeting, M., 2013. Sustainability: choose satellites to monitor deforestation. Nature 496, 293–294. http://dx.doi.org/10.1038/496293a.
Monfreda, C., Ramankutty, N., Foley, J.A., 2008. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Glob. Biogeochem. Cycles, 22(1), GB1022 http://dx.doi.org/10.1029/2007GB002947.
Ozdogan, M., Woodcock, C.E., 2006. Resolution dependent errors in remote sensing of cultivated areas. Remote Sens. Environ. 103, 203–217. http://dx.doi.org/10.1016/j.rse.2006.04.004.
Parry, M.L., Arnell, N., Berry, P., Dodman, D., Fankhauser, S., Hope, C., Kovats, S., Nicholls, R., Satterthwaite, D., Tiffin, R., Wheeler, T., 2009. Assessing the Costs of Adaptation to Climate Change: A Review of the UNFCCC Estimates. International Institute for Environment and Development: Grantham Institute for Climate Change, Imperial College, London.
Pittman, K., Hansen, M.C., Becker-Reshef, I., Potapov, P.V., Justice, C.O., 2010. Estimating global cropland extent with multi-year MODIS data. Remote Sens. 2, 1844–1863. http://dx.doi.org/10.3390/rs2071844.
Quaife, T., Quegan, S., Disney, M., Lewis, P., Lomas, M., Woodward, F.I., 2008. Impact of land cover uncertainties on estimates of biospheric carbon fluxes. Glob. Biogeochem. Cycles 22, GB4016. http://dx.doi.org/10.1029/2007GB003097.
Ramankutty, N., Evan, A.T., Monfreda, C., Foley, J.A., 2008. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. Glob. Biogeochem. Cycles, 22. http://dx.doi.org/10.1029/2007GB002952.
Ramankutty, N., Foley, J.A., 1998. Characterizing patterns of global land use: An analysis of global croplands data. Glob. Biogeochem. Cycles 12, 667–685. http://dx.doi.org/10.1029/98GB02512.
Roberts, L., 2011. 9 Billion? Science 333, 540–543. http://dx.doi.org/10.1126/science.333.6042.540.
See, L., McCallum, I., Fritz, S., Perger, C., Kraxner, F., Obersteiner, M., Deka Baruah, U., Mili, N., Ram Kalita, N., 2013. Mapping cropland in Ethiopia using crowdsourcing. Int. J. Geosci. 4, 6–13. http://dx.doi.org/10.4236/ijg.2013.46A1002.
Smith, P., Gregory, P.J., van Vuuren, D., Obersteiner, M., Havlik, P., Rounsevell, M., Woods, J., Stehfest, E., Bellarby, J., 2010. Competition for land. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 365, 2941–2957. http://dx.doi.org/10.1098/rstb.2010.0127.
Steemans, C., 2008. Coordination of information on the environment (CORINE). In: Kemp, K. (Ed.), Encyclopedia of Geographic Information Science. Sage Publications Inc, Thousand Oaks, CA, pp. 49–50.
The World Bank, 2010. Global strategy to improve agricultural and rural statistics (Economic and Sector Work No. 56719-GLB).
Tilman, D., Balzer, C., Hill, J., Befort, B.L., 2011. Global food demand and the sustainable intensification of agriculture. PNAS 108, 20260–20264. http://dx.doi.org/10.1073/pnas.1116437108.
Tilman, D., Cassman, K.G., Matson, P.A., Naylor, R., Polasky, S., 2002. Agricultural sustainability and intensive production practices. Nature 418, 671–677.
Uganda Bureau of Statistics (UBOS), 2007. Uganda National Household Survey 2005/2006: Report on the Agricultural Module. UBOS, Kampala, Uganda.
Van Wart, J., Kersebaum, K.C., Peng, S., Milner, M., Cassman, K.G., 2013. Estimating crop yield potential at regional to national scales. Field Crops Res. 143, 34–43. http://dx.doi.org/10.1016/j.fcr.2012.11.018.
Von Grebmer, K., Torero, M., Olofinbiyi, T., Fritschel, H., Wiesmann, D., Yohannes, Y., Schofield, L., von Oppeln, C., 2011. 2011 Global Hunger Index. The Challenge of Hunger: Taming Price Spikes and Excessive Food Price Volatility (IFPRI Issue Brief No. 69). IFPRI, Washington DC.
Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R., Woodcock, C.E., 2012. Opening the archive: how free data has enabled the science and monitoring promise of Landsat. Remote Sens. Environ. 122, 2–10. http://dx.doi.org/10.1016/j.rse.2012.01.010.
Yu, L., Wang, J., Clinton, N., Xin, Q., Zhong, L., Chen, Y., Gong, P., 2013a. FROM-GC: 30 m global cropland extent derived through multisource data integration. Int. J. Digit. Earth 6, 521–533. http://dx.doi.org/10.1080/17538947.2013.822574.
Yu, L., Wang, J., Gong, P., 2013b. Improving 30 m global land-cover map FROM-GLC with time series MODIS and auxiliary data sets: a segmentation-based approach. Int. J. Remote Sens. 34, 5851–5867. http://dx.doi.org/10.1080/01431161.2013.798055.