ENVIRONMENTAL RESEARCH
LETTERS

LETTER • OPEN ACCESS

Lessons learned in developing reference data sets with the contribution of citizens: the Geo-Wiki experience

To cite this article: Linda See *et al* 2022 *Environ. Res. Lett.* **17** 065003

ENVIRONMENTAL RESEARCH
LETTERS

CrossMark

OPEN ACCESS

RECEIVED
10 January 2022

REVISED
11 April 2022

ACCEPTED FOR PUBLICATION
27 April 2022

PUBLISHED
13 May 2022

LETTER

Lessons learned in developing reference data sets with the contribution of citizens: the Geo-Wiki experience

Linda See*, Juan Carlos Laso Bayas, Myroslava Lesiv, Dmitry Schepaschenko, Olga Danylo, Ian McCallum, Martina Dürauer, Ivelina Georgieva, Dahlia Domian, Dilek Fraisl, Gerid Hager, Santosh Karanam, Inian Moorthy, Tobias Sturn, Anto Subash and Steffen Fritz

Novel Data Ecosystems for Sustainability (NoDES) Research Group, Advancing Systems Analysis Program, International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria

* Author to whom any correspondence should be addressed.

E-mail: see@iiasa.ac.at

Keywords: remote sensing, reference data, citizen science, crowdsourcing, land cover, environmental monitoring

Supplementary material for this article is available online

Abstract

The development of remotely sensed products such as land cover requires large amounts of high-quality reference data, needed to train remote sensing classification algorithms and for validation. However, due to the lack of sharing and the high costs associated with data collection, particularly ground-based information, the amount of reference data available has not kept up with the vast increase in the availability of satellite imagery, e.g. from Landsat, Sentinel and Planet satellites. To fill this gap, the Geo-Wiki platform for the crowdsourcing of reference data was developed, involving visual interpretation of satellite and aerial imagery. Here we provide an overview of the crowdsourcing campaigns that have been run using Geo-Wiki over the last decade, including the amount of data collected, the research questions driving the campaigns and the outputs produced such as new data layers (e.g. a global map of forest management), new global estimates of areas or percentages of land cover/land use (e.g. the amount of extra land available for biofuels) and reference data sets, all openly shared. We demonstrate that the amount of data collected and the scientific advances in the field of land cover and land use would not have been possible without the participation of citizens. A relatively conservative estimate reveals that citizens have contributed more than 5.3 years of the data collection efforts of one person over short, intensive campaigns run over the last decade. We also provide key observations and lessons learned from these campaigns including the need for quality assurance mechanisms linked to incentives to participate, good communication, training and feedback, and appreciating the ingenuity of the participants.

1. Introduction

Data from satellite remote sensing can provide comprehensive spatial and temporal coverage of the Earth. In the past, these data have been used to map many different environmental variables of interest, including land cover and various types of land use (Szantoi *et al* 2020), vegetation extent as well as biomass, tree species, crop types (Atzberger 2013, Grabska *et al* 2019, Santoro *et al* 2021), and human settlements (Corbane *et al* 2017), among many others. Although Landsat was first launched in 1972, it was the opening up of the Landsat archive in 2008, along with increasing computer power and storage, that has resulted in the proliferation of many new applications (Zhu *et al* 2019). The data from Landsat are now complemented by the Sentinel satellites from the European Space Agency, satellites from other national space agencies, the new generation of nanosatellites as well as very high-resolution imagery openly available for viewing through Google Earth and Microsoft Bing Maps.

This exponential growth in big data from remote sensing has, however, not seen similar increases in the availability of reference data; these are needed to support the training of remote sensing algorithms and the validation of remotely sensed products, which are still largely lacking in many areas (Szantoi *et al* 2020). This paucity of reference data also impacts the overall accuracy of remotely sensed products, where large amounts of high-quality reference data have already been shown to be more important than the choice of the classification algorithm for improving the accuracy of the final product (Maxwell *et al* 2018). For example, the number of global land cover products now available has increased over the last two decades due to the growth in satellite imagery available, yet there has been little improvement in the overall accuracy of these products (Herold *et al* 2016). Moreover, both training and validation data have traditionally been collected on the ground, but this is expensive and not always possible, particularly in the case of validation data where the locations in the statistical sample may not be physically accessible and hence could lead to underrepresentation in the reference data set. One of the few openly available reference data sets available for use in remote sensing is the Land Cover/Use Area Frame Survey (LUCAS), which takes place every three years and involves data collection in the field (Eurostat 2019). It consists of a systematic sample containing around 300K locations across countries in the European Union, but it is a costly exercise (Laso Bayas *et al* 2020). The data are also collected at single coordinates, so they do not provide the area-based data needed for remote sensing, although a module was recently added to the 2018 LUCAS to address this shortcoming (D'Andrimont *et al* 2021). However, not all LUCAS data are surveyed on the ground, but some locations are visually interpreted from satellite and aerial imagery. In fact, there has been an increasing trend to build training and validation data sets from visual interpretation of satellite and aerial imagery, which can provide data of a comparable quality to field-based ground truth data (Copass *et al* 2018). However, much of the reference data used in remote sensing is not openly shared.

For this reason, the Geo-Wiki platform was developed, which involves citizens in the collection of reference data through visual interpretation of very high-resolution satellite and aerial imagery (Fritz *et al* 2012). This application has been used in numerous crowdsourcing campaigns since 2011, where each campaign has been framed around data collection to answer a specific research question. Much of this research has been published as individual papers, e.g. Fritz *et al* (2013b) and Lesiv *et al* (2019), but there has been no attempt at reflecting on the collective contributions of these campaigns and the lessons learned through a decade of experience in crowdsourced reference data collection. Hence, the aim of this paper is to demonstrate how crowdsourcing using the Geo-Wiki platform has contributed to improving land cover and land use reference data sets and products, which would not have been possible without the participation of citizens. We also present the lessons that have been learned throughout this process of running campaigns with Geo-Wiki and how these can guide future crowdsourcing and participatory efforts in land cover and land use science.

2. The Geo-Wiki platform

The Geo-Wiki platform was developed in 2009 as a way of visualizing the three main global land cover maps available at the time, i.e. GLC-2000 (Fritz *et al* 2003), MODIS v.4 (Friedl *et al* 2002) and GlobCover 2005 (Defourny *et al* 2006), in a single place, displayed on top of very high-resolution satellite imagery from Google Earth. The interface allowed users to zoom into any location, view the area covered by individual pixels from these three products, and then indicate if the land cover class matched what the users could see from the imagery (see figure S1 in the supplementary material available online at stacks.iop.org/ERL/17/065003/mmedia). This was also part of ongoing research highlighting the large disagreements between the three global land cover products (Fritz *et al* 2011), and layers of disagreement were provided in Geo-Wiki to guide the contributions.

However, there was little incentive for citizens to provide us with this information, and hence, very little reference data were collected. To increase participation, it became clear that crowdsourcing campaigns were needed, with clear scientific research questions and appropriate incentives. Hence, new capabilities were added to Geo-Wiki to collect data in 'campaign' mode, where campaigns were run for finite periods of time and prizes were offered to incentivize participation. The first campaign, entitled 'Human impact', was focused on validating maps of land availability for biofuels, produced by Cai *et al* (2011). They used a top-down approach to combine different coarse resolution layers to derive estimates of the additional land available for bioenergy production. To validate these maps and revise the estimates, Geo-Wiki was used to collect data on land cover and the degree of human impact using a sample of pixels drawn from the land availability map (see figure S2 in the supplementary material). The results from the campaign demonstrated that the original estimates of land availability for biofuel production were far higher than they should be, which led to a considerable downward adjustment (Fritz *et al* 2013b). We also included the top ten data collectors as co-authors on the publication as an incentive, which is also being advocated by others in the field of citizen science (Ward-Fear *et al* 2020).

3. The Geo-Wiki crowdsourcing campaigns

Since the 'Human Impact' campaign, we have run an additional nine campaigns with Geo-Wiki, which are summarized in table 1. Each campaign had a clear research question, which was intended to act

Table 1. Summary of the Geo-Wiki campaigns that ran from 2011 to 2020.

Campaign name & number of contributors in brackets	Research questions	Incentives	Dates of the campaign (mm/year)	Resolution and grid	# of unique locations where data were collected & [total collected]
Human impact (65 people)	Can we validate a map of land availability for biomass and adjust the global estimates?	Co-authorship to top ten contributors & electronic devices	08/2011–11/2011	1 km MODIS	33 817 [53 278]
Disagreement (61 people)	Can we build a hybrid map of land cover based on crowdsourced data sampled from disagreeing areas from global land cover maps?	Co-authorship to top ten contributors	01/2012–03/2012	300 m GlobCover	8582 [30 359]
Reference data campaign (26 people)	Can we collect additional reference data to support development of hybrid land cover maps?	Micro-payments funded externally	04/2012–07/2012	1 km MODIS	33 473 [35 444]
Wilderness (65 people)	Can we develop a global wilderness map using a bottom-up approach?	Co-authorship to top ten contributors	07/2012–11/2012	1 km MODIS	28 146 [32 861]
Land grabbing (36 people)	Can we gather data on land use in Ethiopia to determine if there are overlaps with proposed land grabbing projects?	Co-authorship to top three contributors and books on aerial imagery	09/2012–12/2012	1 km MODIS	37 350 [82 397]
Cropland (80 people)	Can we gather a reference data set for validating any global cropland product?	Co-authorship to the top contributors (number varied between campaigns) and/or Amazon vouchers of varying amounts and certificates (only implemented occasionally)	09/2016 (13 d)	300 m PROBA-V 60 m sub-cells	36 000 [3 times each]
Field size (130 people)	Can we characterize field sizes globally?		06/2017	80 m and 1 km	130 000 [3 times each]
Forest management (130 people)	Can we create a global, harmonized map of forest management?		04/2018–06/2018	100 m Sentinel 2	49 982 [110 000]
Built-up areas (61 people)	Can we gather a reference data set for validating any global built-up product?		09/2020 (7 d)	80 m 10 m sub-cells	50 000 [5 times each]
Drivers of tropical forest loss (58 people)	Can we develop a map of drivers of tropical forest loss?		12/2020 (15 d)	100 m and 1 km	115 000 [3 times each]

as an intrinsic motivation to participate, i.e. the altruistic idea of helping to make advances in scientific research. We maintained co-authorship as an incentive as well as adding small prizes in the form of Amazon vouchers. Table 1 also shows that the amount of data collected per campaign has increased over time while the duration of the campaigns has decreased. Moreover, the number of unique observations relative to the total has also increased with each campaign; this was a result of gaining a better understanding of how much data could realistically be collected and the desire to collect a minimum number of observations at each location from the Cropland campaign (table 1) onwards for quality control. Finally, we modified some campaigns to gather sub-pixel information, starting with the Cropland campaign. Instead of identifying whether cropland was present in a single pixel of 300 m, we divided it into a grid of 5 × 5 cells (of 60 m each) and asked volunteers to shade those cells where cropland was present (see figure S3 in the supplementary material). This increased the complexity of the task but also the information richness, allowing us to create percentage values for each grid cell rather than only dominant presence of cropland.

Based on table 1, the total number of observations collected through these ten campaigns has been around 1.68 million at around 533K unique locations. Although each campaign involved different visual interpretation tasks, each of which took a different amount of time to complete, if we assume 10 s on average as a very conservative estimate, with an 8 hour day and a 220 day working year, this equates to more than 2.5 years of effort. If this is raised to 20 s on average per observation (e.g. to account for shading of sub-pixels), this value increases to more than 5.3 years, which is still relatively conservative. Hence, the collective contribution of the citizens from Geo-Wiki during the past decade has been considerable.

Table 2 is a summary of the outputs from the campaigns, written up as scientific papers, as well as the open-source publication of the reference data in various data repositories. Not only have the Geo-Wiki volunteers helped us to increase the amount of reference data available for remote sensing more generally, they have also indirectly contributed to the development of new layers (e.g. the forest management map), hybrid maps (e.g. global land cover and forest extent), and new global estimates (e.g. agricultural field sizes). This collective body of research in land cover and land use, driven by crowdsourcing, would not have been possible without the participation of citizens. As the data and products have been openly shared, there are also examples of usage in further studies, e.g. the field size layer from Fritz *et al* (2015) was used as an input to a study on nutrients from farming (Herrero *et al* 2017), the hybrid cropland layer derived from crowdsourcing (Fritz *et al* 2015) has been used to allocate crop types globally in the SPAM database (Anderson *et al* 2014) while the drivers of tropical forest loss data set was used to examine the drivers within protected areas (Fritz *et al* 2022). We expect to see other uses of the data for scientific research in the future, whether as data to train classification algorithms or as new input layers to models.

4. Key observations and lessons learned

As we gained experience through subsequent campaigns, we modified the campaign workflow based on key observations and lessons learned. These are summarized below:

4.1. Campaigns should be short and intense, and provide incentives and recognition for participation

One of the early lessons we learned is that some type of incentive linked to a finite campaign duration can result in a critical mass of contributors willing to participate. From the very beginning, we used co-authorship and small prizes as incentives, which have remained throughout all ten campaigns. However, we have increased the number of contributors that could receive co-authorship and/or small prizes from 10 to around 30. We observed that contributions declined once the ranking of contributors was established, which discouraged participation from those no longer in contention for a prize. Hence, we wanted to encourage as much participation as possible, even if the incentives were quite small near the bottom of the ranking. Without these incentives in place, we are convinced that participation would not have been as high. This is quite different from citizen science projects that are continuously collecting data, e.g. eBird, because they simultaneously fulfil the ongoing needs of an existing bird watching community (Sullivan *et al* 2014). We also moved from campaigns that took place over a three-month period (e.g. Human impact) and condensed these to periods of less than one month (table 1), which was based on feedback from the participants. However, it is also a function of the campaign incentives, which drive participants to compete against one another for co-authorship and prizes. Our approach also has similarities to a mapathon (Quill 2018) or a BioBlitz (Lundmark 2003), both of which concentrate efforts over short periods of time. Finally, we need to find better ways to recognize our contributors, especially those who have participated in multiple campaigns. Ideas include certificates that acknowledge participation across all campaigns and the use of titles, e.g. expert in visual interpretation of cropland or trainer when used in classroom settings, using techniques from other citizen science projects as surveyed by Reeves*et al* (2017).

Table 2. Outputs, publications and reference data sets produced from the Geo-Wiki campaigns. Below IIASA refers to the International Institute for Applied Systems Analysis.

Campaign name	Outputs and publications	Reference data sets
Human impact	Revised estimates of land availability for biofuels (Fritz *et al* 2013b) Raised awareness of the limitations of top-down approaches to land availability for biofuel estimates (Fritz *et al* 2013a) A global field size map (Fritz *et al* 2015)	
Disagreement	A global hybrid land cover map (See *et al* 2015)	Published in PANGAEA: https://doi.pangaea.de/10.1594/PANGAEA.869682 and described in Fritz *et al* (2017)
Reference data campaign	Contributed towards the global hybrid land cover map (See *et al* 2015) and a global forest hybrid map (Schepaschenko *et al* 2015)	
Wilderness	A global map of the degree of human impact/wilderness (See *et al* 2014)	
Land grabbing	A cropland map of Ethiopia (See *et al* 2013)	Data are available to download from Geo-Wiki.org
Cropland	A reference data set to validate remotely sensed cropland layers and a quality assessment of crowdsourced data (Laso Bayas *et al* 2017, Waldner *et al* 2019)	Published in PANGAEA: https://doi.pangaea.de/10.1594/PANGAEA.873912 and described in Laso Bayas *et al* (2017)
Field size	Bottom-up estimates of the global distribution of field sizes (Lesiv *et al* 2019)	Published in IIASA's repository: http://pure.iiasa.acat/id/eprint/15526/
Forest management	A global, harmonized map of forest management (Lesiv *et al* 2022)	Will be published in Zenodo
Built-up areas	A reference data set to validate remotely sensed built-up area layers (See *et al* 2022)	Published in IIASA's repository: http://pure.iiasa.acat/id/eprint/17534/
Drivers of tropical forest loss	A reference data set of potential primary and secondary drivers of tropical deforestation (Laso Bayas *et al* 2022) used to examine the effectiveness of protected areas (Fritz *et al* 2022)	Published in IIASA's repository: http://pure.iiasa.acat/id/eprint/17539/

4.2. Good methods are needed to ensure data quality

The quality assurance methods used during the campaigns have evolved over time. However, they have all used expert control points, i.e. locations where experts visually interpret the imagery and agree on the answer as a way of comparing answers from participants with a gold standard. Moreover, we have always combined the amount of contribution with the quality to ensure that participants take the need to provide high quality data seriously. Here is a summary of this evolution:

- During the first five campaigns, we used a small number of expert control points (i.e. between 60 and 300). Each participant had to visually interpret these control points at the start of the competition. We then weighted the number of contributions by the quality based on the agreement with the expert control points, adjusting the leaderboard at the end.
- In these early campaigns we did not systematically gather more than one observation at each location but relied on the expert control points for determining quality so there was no option to do some type of data aggregation post-campaign, e.g. using majority voting (Kestler *et al* 2011).
- In the latter campaigns, we adopted a more dynamic approach. First, we gathered a much larger number of control points so that these could be provided to participants during the entire campaign and could be linked to the scores. We adjusted the formulas for scoring across these different campaigns based upon experience and user behavior. For example, the formulae for the scoring of the field size campaign is provided in Lesiv *et al* (2019), but we soon realized two major issues: (i) the participants could see a pattern to the allocation of control points, which was not random enough; and (ii) the participants were not penalized enough for certain behaviors such as choosing to say there is 0% cropland when one of the sub-pixels contained cropland in order to classify the images more quickly. We then tweaked these formulae and the scoring with each subsequent campaign; these different approaches to the random delivery of control points and the scoring are documented in different research papers describing the various campaigns (table 2) but see the supplementary information in See *et al* (2022) for an example of the latest approach taken.
- Finally, in these latter campaigns, we required each location in the sample to be visually interpreted a minimum number of times (from 3 to 5) so that we could apply different approaches to aggregating the crowdsourced data into a consolidated reference data set. See, e.g. Laso Bayas *et al* (2022) in which they used the agreement between participants and the overall quality to filter the data.

From these experiences, we have learned that control points, a dynamic process of quality control during the campaign and post-processing of the data using multiple observations from each location are three essential elements for producing a high-quality reference data set. How these have been implemented has varied across the campaigns but we will continue to refine this combination of approaches for future campaigns.

4.3. Communication, training and feedback are key components to any successful campaign

This observation may seem like commonsense, but communication, training and feedback all require considerable resources that must be planned for. We learned that participants need sufficient notice of the start of a campaign (communicated through email, our website, our newsletter, Facebook and Twitter) but that good and constant communication during the campaign has been absolutely critical. For the first time in the last campaign on 'Drivers of tropical forest loss', we used the Discord messaging tool so that participants could have a dedicated chat channel for interaction within the group and with the experts, which was well received (Laso Bayas *et al* 2022). In addition to a training video and a gallery with examples as guidance on visual interpretation, we added an 'Ask the Expert' button to the Geo-Wiki interface. When pressed by a user, it would send an email to the Geo-Wiki team with a link to the image for interpretation. The email was then answered by someone from the team as soon as possible; some answers were also posted on Facebook so that other participants could learn from these queries. In the case of the Field size campaign, the task was quite complex, so the Geo-Wiki team made short videos during the campaign to illustrate answers to incoming queries (Lesiv *et al* 2019). For future campaigns, we will continue to invest heavily in supporting the participants using all of these different communication channels and resources.

4.4. The experts are not always right

In later campaigns, we substantially increased the number of control points used in the competition because of the link to the quality control mechanism and the scoring of points (see section 4.2 above). Building up a control data set took a considerable amount of time, and it was not always possible to collect data from multiple experts at more than one location. Visual interpretation is also subjective, and studies have shown that experts do not always agree (Van Coillie *et al* 2014). Hence, we asked the volunteers to use the 'Ask the Expert' button for feedback but also to challenge the expert in their interpretation. If the user was found to be correct, then they either received bonus points, or as a minimum, they had their scores revised (e.g. Laso Bayas *et al* 2022, See *et al* 2022). This also helped us to improve the expert control data set as the competition ran.

But in the majority of cases, the users were incorrect, and this allowed us to have a dialogue with users and help them to learn from their mistakes. The main lesson here is that much more time should be invested in building a very good control data set for each campaign.

4.5. Do not waste the time of your contributors

In latter campaigns, we chose the size of the sample and the number of times each location could be interpreted based on what could be collected over a short period of time while still obtaining multiple observations at each location for quality control purposes. We did, however, increase the amount of data collected in latter campaigns considerably once we gained experience with what was achievable (e.g. Lesiv *et al* 2019). In the future we would like to move towards an automated system whereby the number of times a location is provided to the crowd is based on the statistical probability that it is correct with a certain confidence, thereby removing locations that are easy to interpret and targeting the efforts towards locations that are more difficult to visually interpret based on, e.g. Bayesian approaches (Salk *et al* 2022). In this way we could maximize the time and effort of the contributors and not use the simple approach of a minimum number of observations at each location. Ultimately, we would like to use the data in machine learning approaches, which require large amounts of input data. Hence, using more intelligent methods for optimizing contributor time is a priority for the future.

4.6. Do not underestimate human ingenuity and the desire to win

In each campaign, we discovered loopholes that were exploited by some participants or behaviors that maximized their ranking on the leaderboard (and hence their contention for a prize). In some cases, participants alerted us to these loopholes as they could see this behavior occurring during the campaign. For example, in the built-up campaign, equal points were awarded for built-up and non-built-up images, but non-built-up images were much easier to interpret. Hence, a small set of users refreshed their browser to avoid built-up areas. We stopped the campaign once this was discovered, removed the non-built-up areas from the control set and resumed the campaign (See *et al* 2022). In the next campaign, we fixed this issue technologically. In another campaign, one user programmed a bot to increase the number of contributions and hence their ranking on the leaderboard. Once discovered, this user and their data were removed from the campaign (Laso Bayas *et al* 2022). We were surprised at these types of behaviors, but we acknowledge that it is a function of the incentive scheme that we have developed over time. Based on questionnaires administered after each campaign from the Cropland Campaign onwards, the majority of participants selected 'helping science' as one incentive for participation (Laso Bayas *et al* 2017, 2022, See *et al* 2022). Co-authorship has clearly been a strong motivator given the academic backgrounds of many of the participants. However, we are also increasingly aware that the nature of the competition in the campaigns has generated behaviors that are not always altruistic. In contrast, we have also experienced participants donating the monetary value of the Amazon vouchers to the student program at the International Institute of Applied Systems Analysis (IIASA). Overall, we feel that the incentives used have been successful, and we will continue to refine this type of model in the future.

5. Conclusions

By involving citizens in the visual interpretation of very high-resolution satellite imagery via the Geo-Wiki platform, we have been able to collect much more data than would be possible using only research scientists. By running several different campaigns focused on a range of research questions related to improving the reference data available for remote sensing applications as well as using the data to create new layers or new global estimates, e.g. of field size distributions, we have been able to make scientific advances in the field of land cover and land use.

In addition to Geo-Wiki, crowdsourcing is also being used in other remote sensing applications. For example, reference data were collected via crowdsourcing in collaboration with Google for validating a built-up surface layer (Marconcini *et al* 2020) although the reference data are not openly shared. Moreover, a recent review by Saralioglu and Gungor (2020) reveals the considerable extent to which crowdsourcing is being used for visual interpretation. Hence, it is evident that citizen science and crowdsourcing are being used for improving reference data sets more generally but that the potential is still largely untapped. This is reflected in the increasing appearance of citizen science and crowdsourcing in proposal call texts as well as legislation that acknowledges citizen generated data as a valid source of information for decision making (Fritz *et al* 2019). Hence, the next decade should see many more examples of involving citizens in reference data generation.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://doi.pangaea.de/10.1594/PANGAEA.869682.

ORCID iDs

Linda See https://orcid.org/0000-0002-2665-7065
Juan Carlos Laso Bayas https://orcid.org/0000-0003-2844-3842
Dmitry Schepaschenko https://orcid.org/0000-0002-7814-4990
Dilek Fraisl https://orcid.org/0000-0001-7523-7967

References

Anderson W, You L, Wood S, Wood-Sichra U and Wu W 2014 *A Comparative Analysis of Global Cropping Systems Models and Maps* (Rochester, NY: Social Science Research Network)
Atzberger C 2013 Advances in remote sensing of agriculture: context description, existing operational monitoring systems and major information needs *Remote Sens.* **5** 949–81
Cai X, Zhang X and Wang D 2011 Land availability for biofuel production *Environ. Sci. Technol.* **45** 334–9
Copass C, Antonova N and Kennedy R 2018 Comparison of office and field techniques for validating landscape change classification in Pacific Northwest national parks *Remote Sens.* **11** 3
Corbane C *et al* 2017 Big earth data analytics on Sentinel-1 and Landsat imagery in support to global human settlements mapping *Big Earth Data* **1** 118–44
d'Andrimont R, Verhegghen A, Meroni M, Lemoine G, Strobl P, Eiselt B, Yordanov M, Martinez-Sanchez L and van der Velde M 2021 LUCAS Copernicus 2018: earth-observation-relevant *in situ* data on land cover and use throughout the European Union *Earth Syst. Sci. Data* **13** 1119–33
Defourny P, Vancustem C, Bicheron P, Brockmann C, Nino F, Schouten L and Leroy M 2006 GLOBCOVER: a 300 m global land cover product for 2005 using ENVISAT MERIS time series *Proc. ISPRS Commission VII Mid-Term Symp.: Remote Sensing: From Pixels to Processes* (Enscede NL)
Eurostat 2019 LUCAS—land use and land cover survey (available at: https://ec.europa.eu/eurostat/statistics-explained/index.php/LUCAS_-_Land_use_and_land_cover_survey#The_LUCAS_survey)
Friedl M A *et al* 2002 Global land cover mapping from MODIS: algorithms and early results *Remote Sens. Environ.* **83** 287–302
Fritz S *et al* 2013b Downgrading recent estimates of land available for biofuel production *Environ. Sci. Technol.* **47** 1688–94
Fritz S *et al* 2015 Mapping global cropland and field size *Glob. Change Biol.* **21** 1980–92
Fritz S *et al* 2017 A global dataset of crowdsourced land cover and land use reference data *Sci. Data* **4** 170075
Fritz S *et al* 2019 Citizen science and the United Nations sustainable development goals *Nat. Sustain.* **2** 922–30
Fritz S *et al* 2022 A continental assessment of the drivers of tropical deforestation with a focus on protected areas *Front. Conserv. Sci.* **3** 830248
Fritz S, Bartholomé E, Belward A, Hartley A, Stibig H-J, Eva H and Mayaux P 2003 *Harmonisation, Mosaicing and Production of the Global Land Cover 2000 Database (Beta Version)* (Luxembourg: Office for Official Publications of the European Communities)
Fritz S, McCallum I, Schill C, Perger C, See L, Schepaschenko D, van der Velde M, Kraxner F and Obersteiner M 2012 Geo-Wiki: an online platform for improving global land cover *Environ. Modelling Softw.* **31** 110–23
Fritz S, See L, McCallum I, Schill C, Obersteiner M, van der Velde M, Boettcher H, Havlík P and Achard F 2011 Highlighting continued uncertainty in global land cover maps for the user community *Environ. Res. Lett.* **6** 044005
Fritz S, See L and Valin H 2013a Current issues and uncertainties in estimating global land availability for biofuel production *Biofuels* **4** 343–5
Grabska E, Hostert P, Pflugmacher D and Ostapowicz K 2019 Forest stand species mapping using the Sentinel-2 time series *Remote Sens.* **11** 1197
Herold M, See L, Tsendbazar N-E and Fritz S 2016 Towards an integrated global land cover monitoring and mapping system *Remote Sens.* **8** 1036
Herrero M *et al* 2017 Farming and the geography of nutrient production for human use: a transdisciplinary analysis *Lancet Planet. Heath* **1** e33–e42
Kestler H A, Lausser L, Lindner W and Palm G 2011 On the fusion of threshold classifiers for categorization and dimensionality reduction *Comput. Stat.* **26** 321–40
Laso Bayas J C *et al* 2017 A global reference database of crowdsourced cropland data collected using the Geo-Wiki platform *Sci. Data* **4** 170136
Laso Bayas J C *et al* 2022 Drivers of tropical forest loss between 2008 and 2019 *Sci. Data* **9** 146
Laso Bayas J C, See L, Bartl H, Sturn T, Karner M, Fraisl D, Moorthy I, Busch M, van der Velde M and Fritz S 2020 Crowdsourcing LUCAS: citizens generating reference land cover and land use data with a mobile app *Land* **9** 446
Lesiv M *et al* 2019 Estimating the global distribution of field size using crowdsourcing *Glob. Change Biol.* **25** 174–86
Lesiv M *et al* 2022 Global forest management data for 2015 at a 100 m resolution *Sci. Data*
Lundmark C 2003 BioBlitz: getting into backyard biodiversity *BioScience* **53** 329
Marconcini M *et al* 2020 Outlining where humans live, the World Settlement Footprint 2015 *Sci. Data* **7** 242
Maxwell A E, Warner T A and Fang F 2018 Implementation of machine-learning classification in remote sensing: an applied review *Int. J. Remote Sens.* **39** 2784–817
Quill T M 2018 Humanitarian mapping as library outreach: a case for community-oriented mapathons *J. Web Librariansh.* **12** 160–8
Reeves N, Tinati R, Zerr S, Van Kleek M G and Simperl E 2017 From crowd to community: a survey of online community features in citizen science projects *Proc. 2017 ACM Conf. on Computer Supported Cooperative Work and Social Computing CSCW '17: Computer Supported Cooperative Work and Social Computing* (Portland, OR: ACM) pp 2137–52
Salk C, Moltchanova E, See L, Sturn T, McCallum I and Fritz S 2022 How many people need to classify the same image? A method for optimizing volunteer contributions in binary geographical classifications *PLoS One*
Santoro M *et al* 2021 The global forest above-ground biomass pool for 2010 estimated from high-resolution satellite observations *Earth Syst. Sci. Data* **13** 3927–50
Saralioglu E and Gungor O 2020 Crowdsourcing in remote sensing: a review of applications and future directions *IEEE Geosci. Remote Sens. Mag.* **8** 89–110
Schepaschenko D *et al* 2015 Development of a global hybrid forest mask through the synergy of remote sensing, crowdsourcing and FAO statistics *Remote Sens. Environ.* **162** 208–20
See L *et al* 2015 Building a hybrid land cover map with crowdsourcing and geographically weighted regression *ISPRS J. Photogramm. Remote Sens.* **103** 48–56
See L *et al* 2022 A crowdsourced global data set for validating built-up surface layers *Sci. Data* **9** 13
See L, McCallum I, Fritz S, Perger C, Kraxner F, Obersteiner M, Baruah U D, Mili N and Kalita N R 2013 Mapping cropland in Ethiopia using crowdsourcing *Indian J. Gastroenterol.* **04** 6–13
See L *et al* 2014 Mapping human impact using crowdsourcing *Mapping Wilderness: Concepts, Techniques and Applications of GIS* ed S Carver (Dordrecht: Springer) pp 89–102

Sullivan B L *et al* 2014 The eBird enterprise: an integrated approach to development and application of citizen science *Biol. Conserv.* **169** 31–40

Szantoi Z, Geller G N, Tsendbazar N-E, See L, Griffiths P, Fritz S, Gong P, Herold M, Mora B and Obregón A 2020 Addressing the need for improved land cover map products for policy support *Environ. Sci. Policy* **112** 28–35

Van Coillie F M B, Gardin S, Anseel F, Duyck W, Verbeke L P C and Wulf R R D 2014 Variability of operator performance in remote-sensing image interpretation: the importance of human and external factors *Int. J. Remote Sens.* **35** 754–78

Waldner F *et al* 2019 Conflation of expert and crowd reference data to validate global binary thematic maps *Remote Sens. Environ.* **221** 235–46

Ward-Fear G, Pauly G B, Vendetti J E and Shine R 2020 Authorship protocols must change to credit citizen scientists *Trends Ecol. Evol.* **35** 187–90

Zhu Z *et al* 2019 Benefits of the free and open Landsat data policy *Remote Sens. Environ.* **224** 382–5