ENVIRONMENTAL RESEARCH
LETTERS

LETTER • OPEN ACCESS

A new global hybrid map of annual herbaceous cropland at a 500 m resolution for the year 2019

To cite this article: Steffen Fritz *et al* 2024 *Environ. Res. Lett.* **19** 084025

ENVIRONMENTAL RESEARCH
LETTERS

CrossMark

OPEN ACCESS

RECEIVED
21 March 2024

REVISED
14 June 2024

ACCEPTED FOR PUBLICATION
2 August 2024

PUBLISHED
14 August 2024

LETTER

A new global hybrid map of annual herbaceous cropland at a 500 m resolution for the year 2019

Steffen Fritz[1,*], Myroslava Lesiv[1], Linda See[1], Dmitry Schepaschenko[1], Katya Pérez Guzmán[1], Juan Carlos Laso Bayas[1], Maria Schepaschenko[1], Ivelina Georgieva[1], Francesco Collivignarelli[2], Michele Meroni[3], Hervé Kerdiles[2], Felix Rembold[2] and Ian McCallum[1]

[1] International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria
[2] Joint Research Center of the European Commission (JRC), Ispra, Italy
[3] Seidor Consulting, Barcelona, Spain
[*] Author to whom any correspondence should be addressed.

E-mail: fritz@iiasa.ac.at

Keywords: global cropland mapping, food security, early warning, remote sensing

Supplementary material for this article is available online

Abstract

The global spatial extent of croplands is a crucial input to global and regional agricultural monitoring and modeling systems. Although many new remotely-sensed products are now appearing due to recent advances in the spatial and temporal resolution of satellite sensors, there are still issues with these products that are related to the definition of cropland used and the accuracies of these maps, particularly when examined spatially. To address the needs of the agricultural monitoring community, here we have created a hybrid map of global cropland extent at a 500 m resolution by fusing two of the latest high resolution remotely-sensed cropland products: the European Space Agency's WorldCereal and the cropland layer from the University of Maryland. We aggregated the two products to a common resolution of 500 m to produce percentage cropland and compared them spatially, calculating two kinds of disagreement: density disagreement, where the two maps differ by more than 80%, and absence-presence of cropland disagreement, where one map indicates the presence of cropland while the other does not. Based on these disagreements, we selected continuous areas of disagreement, referred to in the paper as hotspots of disagreement, for manual correction by experts using the Geo-Wiki land cover application. The hybrid map was then validated using a stratified random sample based on the disagreement layer, where the sample was visually interpreted by a different set of experts using Geo-Wiki. The results show that the hybrid product improves upon the overall accuracy statistics in the areas where the underlying cropland layer from the University of Maryland was improved with the WorldCereal product, but more importantly, it represents an improved spatially explicit cropland mask for early warning and food security assessment purposes.

1. Introduction

Various global maps of land cover as well as specific products focused on identifying croplands have been derived from satellite-based Earth observations over the last three decades. These products represent one of the most important sources of baseline terrestrial information and have been used in a wide variety of applications, e.g. as inputs to global models of land use and land use change (Foley *et al* 2011, Verburg *et al* 2011), to model changes in land surface patterns (Pielke 2005), for policy development and decision making (Justice *et al* 2015), to assess the land available for biofuels (Cai *et al* 2011), for monitoring crop health and yield prediction (Lobell *et al* 2015), as the basis for crop distribution modeling (You *et al* 2007), and for food security and early warning purposes (Liu *et al* 2008, Thenkabail *et al* 2009, Fritz *et al* 2019, Rembold *et al* 2019).

Due to an increase in the frequency and severity of agricultural droughts from climate change (Lee *et al* 2023), the accurate mapping of cropland extent is becoming more and more important. Crop failures and losses are the main direct impact of drought on agricultural sector productivity. Drought-induced production losses cause negative supply shocks and can therefore heavily impact the global supply of certain commodities, in particular, if they occur simultaneously, for example, as a multiple breadbasket failure (Gaupp *et al* 2020).

One way to mitigate the impacts of drought is through the provision of timely information from early warning and monitoring systems, which can be used to ensure an appropriate response. Early warning systems rely heavily on a cropland mask to define the areas where anomalies of the various indicators should be considered for early warning purposes (Fritz *et al* 2019). Different land cover products have become available recently, such as WorldCover (https://esa-worldcover.org/) and ESRI's land cover product (https://livingatlas.arcgis.com/landcover/), as well as cropland specific products such as WorldCereal (Van Tricht *et al* 2023) and the cropland layer produced by the University of Maryland (Potapov *et al* 2022). These products follow the trend in the increasing spatial resolution and accuracy observed by Herold *et al* (2016) over the last few decades.

However, these products have different cropland class definitions, and although overall accuracy may be high, they differ substantially when spatially compared with respect to the presence and absence of cropland. One approach for improving the spatial distribution of land cover maps in the past has been to produce hybrid maps, which merge more than one product into an integrated layer. These resulting hybrid maps also tend to have higher accuracies (when aggregated to a common resolution) than individual products and are produced using different approaches such as Geographically Weighted Regression, which uses independent reference data as inputs (Schepaschenko *et al* 2015, See *et al* 2015). Local accuracy metrics are computed, and the product is then merged based on the best locally performing accuracies. However, the downside of this approach (and any approach that employs a machine learning algorithm) is that it requires large amounts of input data, e.g. *in situ* field data or data collected through visual interpretation of high-resolution imagery.

In this paper, we use a novel approach to produce a hybrid cropland map that does not require large amounts of input data. Instead, it focuses on improving the areas where individual cropland maps disagree using visual inspection by experts, who decide which of the maps better captures the cropland extent and then correct the hybrid map accordingly. Such an approach was demonstrated previously by See and Fritz (2006), who combined the best performing classes from the GLC-2000 and MODIS land cover products into a single layer, resulting in a more spatially accurate land cover product. Here we focus on integrating two of the most recently available cropland products into a hybrid cropland map, namely, the WorldCereal cropland extent product (Van Tricht *et al* 2023) and the cropland map produced by Potapov *et al* (2022). A modified version of this hybrid map is now being used by the ASAP (Anomaly hotspots of Agricultural Production) system of the Joint Research Center of the European Commission (see https://agricultural-production-hotspots.ec.europa.eu) as a baseline cropland mask for early warning and food security monitoring.

2. Method

2.1. Selection of products for the hybrid cropland map

The rationale for the choice of products was based on a combination of fitness-for-purpose, high accuracy, the suitability of the cropland definition, and the time period for which the maps were produced. National maps were not considered here because some level of consistency across continents would be lost. Moreover, for some national and regional maps, accuracy statistics do not exist, or the minimum mapping unit is too large. For example, the European CORINE land cover product (Büttner 2014) has no truly independent accuracy statistic, and the minimum mapping unit of 25 ha is too large for our purposes because small cropland areas would not be captured. Similarly, when aggregating to a coarser resolution such as 500 m, the percentage of cropland might be overestimated.

For these reasons, we focus on the two latest remotely-sensed cropland extent products: the cropland extent layer from the WorldCereal product (Van Tricht *et al* 2023) at a 10 m resolution, hereafter referred to as WorldCereal, and the map produced by the Global Land Analysis and Discovery (GLAD) team at the University of Maryland (Potapov *et al* 2022) at a 30 m resolution, hereafter referred to as GLAD Cropland. These two products are superior to other existing land cover products since they both have substantially higher accuracies in detecting cropland compared to other very recent land cover maps such as WorldCover (https://esa-worldcover.org/), the ESRI land cover product (https://livingatlas.arcgis.com/landcover/) and the Dynamic World product (https://dynamicworld.app/). See the supplementary materials for a description of these products and their accuracy statistics.

2.2. Defining cropland for the hybrid product

The definitions of cropland in the WorldCereal and GLAD Cropland products are quite similar except for fallow land (see supplementary material). However,

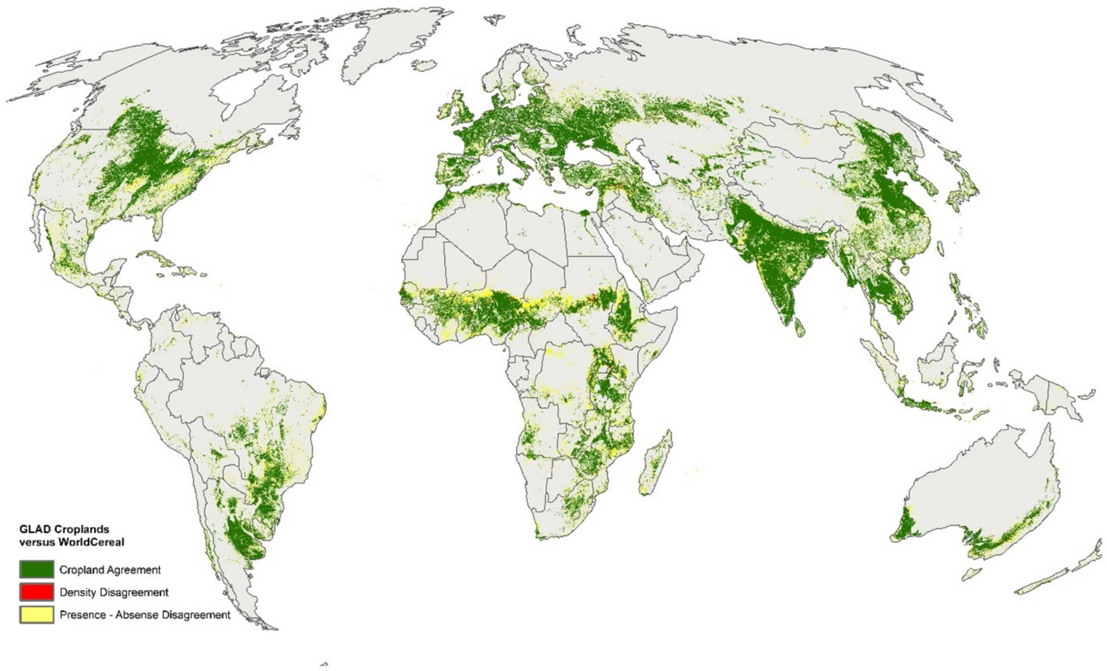

Figure 1. Comparison of the WorldCereal and GLAD Cropland products, showing high overall agreement in green, moderate levels of presence-absence disagreement in yellow and extremely minor levels of density disagreement in red.

for food security applications, we want to include fallow land since a crop mask is normally applied for several years and not updated annually. For our purposes, we adopt the GLAD definition for the hybrid herbaceous annual cropland map (Potapov *et al* 2022): Land used for annual and perennial herbaceous crops for human consumption, forage (including hay) and biofuel. Perennial woody crops, permanent pastures and shifting cultivation are excluded. The fallow length is limited to 4 years for inclusion as cropland.

2.3. Development of the hybrid cropland map

Three main steps were undertaken to develop the hybrid cropland map as described in the sections that follow.

2.3.1. Comparison of the WorldCereal and GLAD Cropland products

The first step involved comparing the WorldCereal and GLAD Cropland products in terms of (i) overall area and (ii) spatial agreement and disagreement. The area of cropland in WorldCereal was calculated to be 1137 355.692 kha for the year 2021 while the cropland area in the GLAD Cropland product for 2016–2019 was 1218 425.442 kha. One explanation for this difference is that the GLAD Cropland product includes forage and fallow land, which results in a higher estimate of cropland overall.

To meet the needs of different user communities, and the global agricultural monitoring community and early warning, in particular, aggregated maps of cropland at a 500 m resolution are sufficient. Therefore, the maps were aggregated to a resolution of 500 m and the cropland percentage was calculated. This resolution was also chosen since it allows for reconciliation of some degree of differences in the definitions of cropland in the two products, and it allows the two products to be compared spatially since they have different spatial resolutions. The frequency of cropland percentages in the two products was then calculated by 10% bins, which is shown in figure S1 in the supplementary material. When looking at this distribution of pixels, there are similarities across most bins except for the 0%–10% and 90%–100% categories. From this analysis, we can see that there are about 10 million more pixels in the 0%–10% class in the WorldCereal map that are not present in the GLAD Cropland map. There are about eight million more pixels in the 90%–100% class in the GLAD Cropland data than there are in the WorldCereal data.

We then undertook a spatial agreement/disagreement analysis between the two products to better understand how they compare spatially. We recorded two different types of disagreement between the maps: (i) density disagreement; and (ii) presence-absence disagreement.

Density disagreement (shown in red in figure 1) is recorded if one map has more than 80% disagreement compared to the other one. The 80% threshold of density disagreement was chosen to highlight those areas where the disagreement is substantial/very severe and cannot be attributed to differences in cropland definitions alone. Since we wanted to concentrate our initial efforts in identifying those places where we could improve the spatial representation

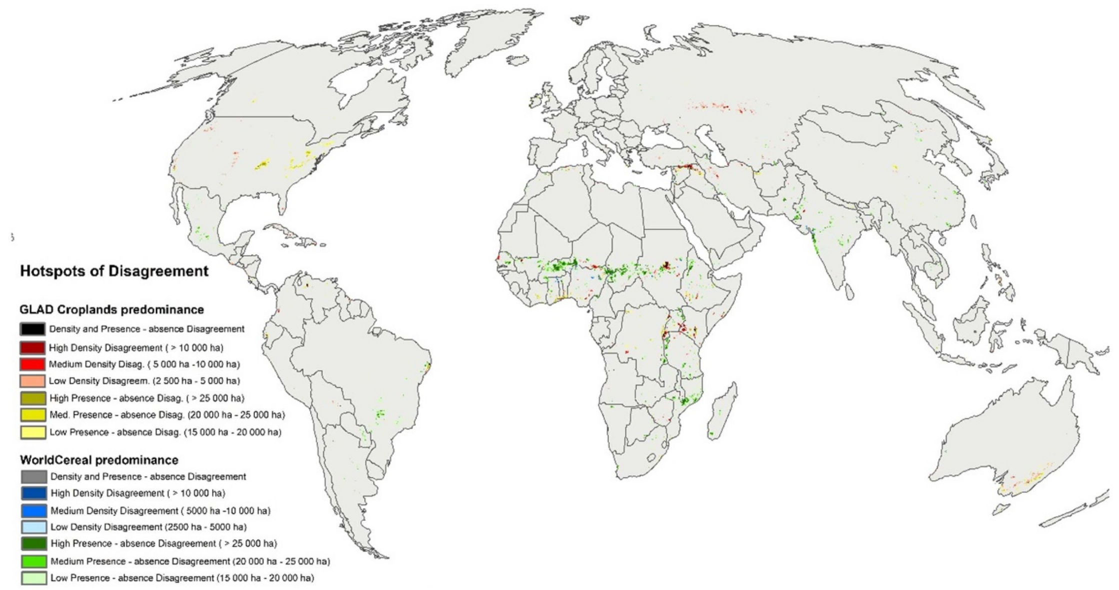

Figure 2. Hotspots of disagreement between the WorldCereal and GLAD Cropland products. See supplementary material for closer inspection of the hotspots of disagreement.

of cropland, this threshold was deemed appropriate, especially considering the amount of labor available to visually improve the map.

When the two maps are compared using density disagreement, 0.93% of pixels show 80% or more cropland in the GLAD Cropland map, and 0.27% of pixels show 80% or more cropland in WorldCereal.

This very high degree of density agreement points towards a very high convergence between the two independently derived global cropland products, demonstrating very high agreement regarding the location of major crop growing regions in the world.

Presence-absence disagreement occurs when one map shows any percentage of cropland in a 500 m pixel when the other map shows no cropland. When we examine the presence-absence disagreement (shown in yellow in figure 1), there is only 73% agreement (shown in green in figure 1). Hence, there is still quite some spatial disagreement between the two maps. In 10% of the pixels, the GLAD Cropland product records cropland with no corresponding cropland in WorldCereal, with 17% for the opposite case. Based on this comparison that shows relatively high presence and absence disagreement, a hybrid product will likely improve the representation of the spatial distribution of cropland extent.

2.3.2. Baseline selection

To fuse the two latest maps, an initial baseline product is needed that can be updated. We initially planned to use the continental accuracy numbers as a guide to choose either WorldCereal or GLAD Cropland depending on which product performed better at a continental level (see table S1, supplementary material). However, after some initial preliminary validation, we determined that using the baseline GLAD Cropland product as the underlying initial basemap would likely result in a more consistent product with the cropland definition being used (since WorldCereal does not include fallow or forage crops). Therefore, we decided to use the GLAD Cropland product as the baseline layer.

2.3.3. Hotspot-based area correction

We then undertook a disagreement analysis between the WorldCereal and GLAD Cropland products to identify hotspots of disagreement as shown in figure 2. We focused on areas in a 20 by 20 km grid where the density disagreement was larger than 80% for a contiguous area larger than 10 000 ha, 5000–10 000 ha and 2500–5000 ha, which we refer to as high, medium and low density disagreement (see figure 2). We also focused on areas within a 20 by 20 km grid where one map had any cropland and the other map had no cropland, which we refer to as presence/absence disagreement, highlighting areas larger than 25 000 ha, 20 000–25 000 ha and areas between 15 000–20 000 ha (see figure 2). These hotspot areas of disagreement were first gridded using a 2 by 2 km grid and then inspected by the two lead authors of this paper, who determined which map was more accurate in each 2 km grid cell within each hotspot, using freely available very high-resolution imagery in the Geo-Wiki application (Fritz *et al* 2012). We applied this correction to all 20 by 20 km grids showing a density disagreement larger than 10 000 ha and presence/-absence disagreement larger than 25 000 ha (corresponding to all the darker colors shown in figure 2).

Examples of 8 hotspots of disagreement are shown in the supplementary material. Note that we have three classes: WorldCover is more correct, the GLAD Cropland product is more correct, or the two

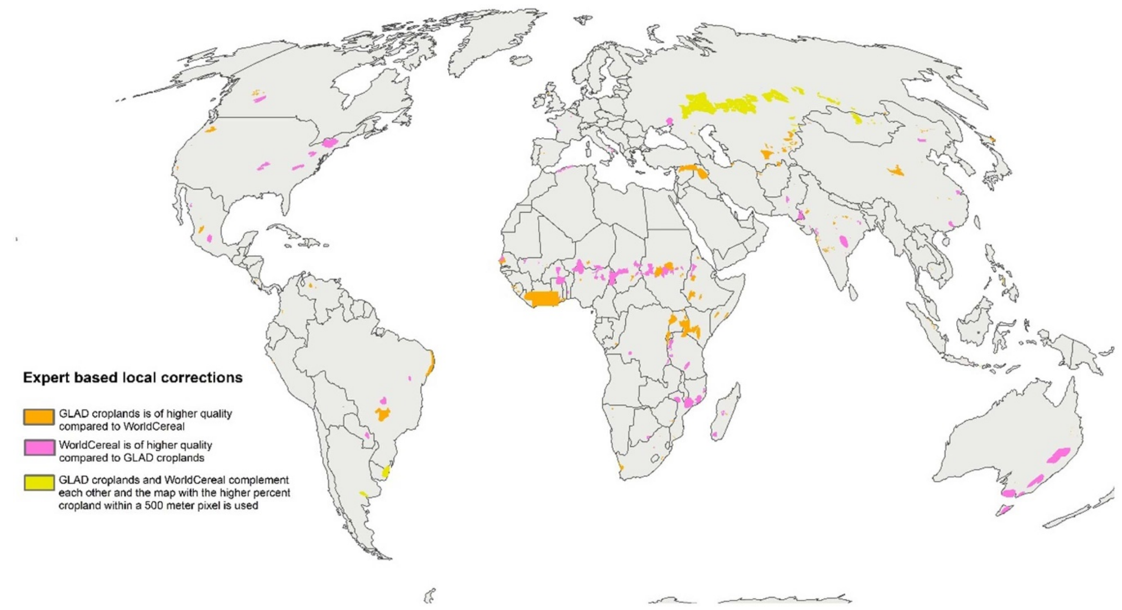

Figure 3. Location of area-based corrections to the GLAD Cropland product.

maps are complementary. We use the complementary class in landscapes where very large fields of greater than 100 ha dominate (e.g. in Kazakhstan), and where due to the crop rotation, some cropland fields have not been mapped in either the WorldCereal or GLAD Cropland products. Figure 3 shows all the places where corrections were made to the GLAD Cropland product and the type of corrections.

Using this approach and focusing on just major hotspots, we were able to make area-based corrections to 8.8% of the global cropland area (maximum extent), which captured 11.7% of the major disagreeing pixels globally (either density or presence/absence disagreement), 3.6% where the GLAD Cropland product is of higher quality (1.3% where GLAD Cropland indicates no cropland or little cropland and WorldCereal Cropland has a high percentage of cropland, 2.3% where GLAD says cropland or high percentage cropland and WorldCereal says no cropland or little cropland), 5.7% where WorldCereal is of higher quality (1.7% where GLAD says cropland or high cropland percentage and WorldCereal no cropland or little cropland) and 2.3% where the maps complement each other (cropland of 1.3% and 1% in GLAD Cropland and WorldCereal, respectively). Note that since we used the GLAD Cropland as a basemap, we could simply have delineated areas where WorldCereal is better. However, since we might want to change the basemap at a later stage, we pursued this more flexible approach. It should also be noted that the corrections we made affect large areas of cropland/non-cropland and are not due to changes in cropland between 2019 and 2021.

2.3.4. Validation of the hybrid cropland map

To understand to what degree the new hybrid map has improved, we validated the map using an independent probability-based sample following guidelines for validation data generation (Olofsson *et al* 2014). We applied a stratified random sample (496) in the corrected areas and calculated the overall accuracy for cropland/non-cropland. We also calculated the mean absolute error (MAE) and root mean squared error (RMSE) based on the % cropland per pixel between the validation samples from the hybrid and GLAD Cropland maps in the areas where improvements of GLAD Cropland were made using the WorldCereal product.

3. Results

3.1. The hybrid cropland map

The final hybrid herbaceous annual cropland map (including fallow) at a 500 m resolution is shown in figure 4 and is available for download from Zenodo (https://zenodo.org/records/10 818 824).

3.2. Results of the validation

We can see that the corrections made to the GLAD Cropland map based on the WorldCereal map result in an improvement in the overall accuracy from 71.8% (CI 3.9%) to 83.3% (CI 3.3%) (see the confusion matrix, table S2, in the supplementary material) in the areas where a manual delineation was undertaken and where WorldCereal is either better or complementary (see pink and yellow areas in figure 3). In particular, it can be noted that the omission error

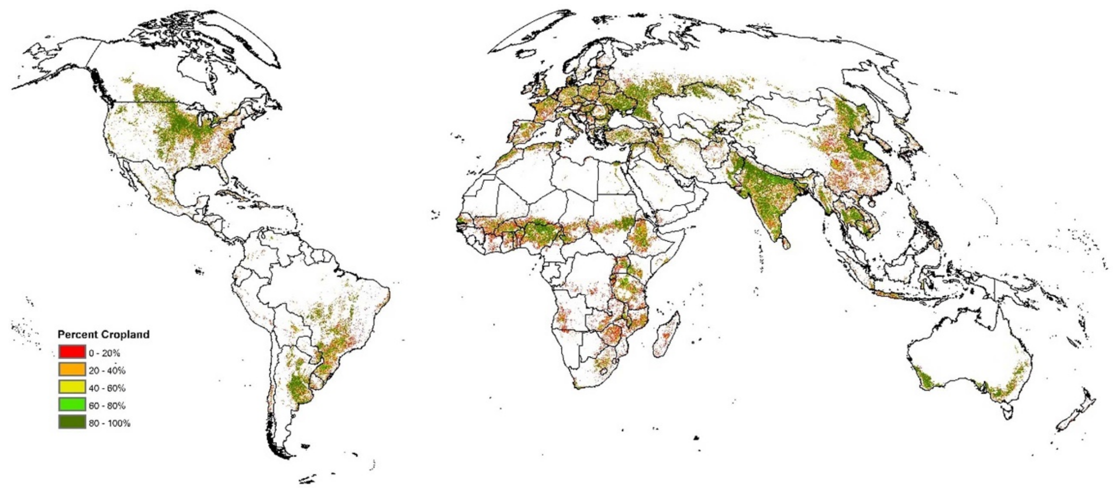

Figure 4. The final hybrid cropland map (as % cropland) at a 500 m resolution.

of cropland in the delineated areas is reduced (corresponding to an increase in the Producers Accuracy from 67.3% (CI 3.7%) to 92.2 (CI 2.8%) The MAE and RMSE also decreased from 20.8% and 32.3% in the GLAD Cropland map to 16.1% and 25.5% in the hybrid map, respectively, indicating a clear improvement in the hybrid map (table S3), not just in the presence-absence of cropland but also in the actual cropland percentages in the 500 m grids.

4. Discussion

In this paper we have implemented a new approach to hybrid map generation. In contrast, previous approaches such as kriging and GWR have used existing products and additional reference data to determine which map is better at a given location (Schepaschenko *et al* 2015, See *et al* 2015). The issue with these approaches is that they require a systematic sample that is independent of the original map. Obtaining an independent, sufficiently high-quality reference data set based on visual interpretation is challenging and generating it would likely require more time than the time spent here on the manual interpretation. By initially undertaking a disagreement analysis, we can more effectively identify those areas where one or the other map is more accurate. Visual interpretation of the areas of disagreement using very high-resolution images available in Microsoft Bing Maps, Google Earth, etc. in combination with Sentinel 2 images (which provides temporal information) and location specific NDVI profiles helps to visually identify cropland.

Even though the map is only available at a 500 m resolution, making the map available at a higher resolution such as 100 or 30 m (e.g. by aggregating WorldCereal to 30 m) will also result in issues related to the differences in the definitions (e.g. the inclusion of fallow in one product) and geolocation issues (Landsat and Sentinel are not aligned) although this could be additionally tried and tested as follow up work. The resolution of 500 m is, however, sufficient for many applications (e.g. food security). We also find that a higher resolution will not improve the depiction of cropland or non-cropland in many locations since very large areas still show issues in terms of cropland presence and absence.

It should be noted that some of the differences identified might be due to the slight differences in the herbaceous annual cropland definition. For example, herbaceous perennial crops as well as forage are generally included in the GLAD Cropland definition, but they are only partly included in the WorldCereal definition (e.g. sugar cane). However, the biggest difference between the two products is in fallow land.

The reason why GLAD includes higher frequencies in the 90%–100% bin (see figures S1 and S2 in the supplementary material) could partly be attributed to the inclusion of fallow and forage and possibly also some misclassification of permanent woody cropland (e.g. some larger vineyard areas in Europe). On the other hand, WorldCereal picks up 500 m grids with a low cropland intensity in the 0%–10% category that are not picked up by GLAD (see figures S1 and S3 in the supplementary material). Some of these areas are low intensity cropland areas in the dry parts of Africa, which are mixed systems that also include annual cropland although these are likely not harvested every year due to droughts.

Focusing on hotspot areas of disagreement can help to improve a map relatively quickly but this analysis has shown that most of the disagreement is not concentrated in the hotspot areas. However, targeting all areas of disagreement would be much harder and more labor intensive. A more refined and sensitive hotspot analysis that focuses on additional areas could follow, especially if more person hours could be dedicated to improving the map. Hence, this approach would need to be complemented by crowdsourcing to tackle smaller disagreeing areas than those identified in the hotspots. This could be achieved by opening up the Geo-Wiki tools to the broader cropland community, thereby improving the map further through a community-based effort. Other maps that have a cropland class (e.g. WorldCover, Google Dynamic World, ESRI land cover, etc.) could also be added. A similar approach could be implemented, and a composite disagreement map could be produced using recent global land cover products, highlighting which combinations of maps disagree at which location. Participants in this community-based map improvement process could delineate areas and indicate which map is better at which location.

The validation approach has demonstrated that the manual correction approach presented here can increase the accuracy of the original map (in this case GLAD Cropland) by more than 11% in the areas where another map (in this case WorldCereal) was used. This accuracy number in the improved areas is still lower than the original accuracy of the maps calculated for the entire globe. The reason is that in disagreeing areas, mapping is particularly challenging (e.g. in drier areas the remote sensing signal from cropland resembles that of grassland). However, we acknowledge that the approach as applied here does not correct areas where both maps are wrong. This issue could be corrected in the future by adding other maps to the analysis as outlined in the previous paragraph, which may help to identify some of these types of errors.

5. Conclusion

We have produced a new global annual cropland map for the year 2019 by integrating the two latest global cropland layers, namely WorldCereal for 2021 and GLAD Cropland for 2016–2019. This approach is new and only requires the manual visual interpretation of large areas where the products disagree. An independent accuracy of the GLAD Cropland and hybrid maps show that the hybrid map is of higher quality. To our knowledge, this is the best current global annual cropland map available for the user community. This map has demonstrated the advantage of the manual correction of hotspot areas of disagreement. Although only 11.7% of the disagreement has been corrected, further improvements using Geo-Wiki and a community-driven effort, possibly also integrating other additional maps, could be made to further improve the accuracy of the map, which will be the subject of future work. Such further improvements will also lead to a higher overall accuracy of the hybrid map if more disagreeing areas are corrected in the future.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://zenodo.org/records/10818824.

ORCID iDs

Steffen Fritz https://orcid.org/0000-0003-0420-8549
Linda See https://orcid.org/0000-0002-2665-7065
Dmitry Schepaschenko https://orcid.org/0000-0002-7814-4990
Juan Carlos Laso Bayas https://orcid.org/0000-0003-2844-3842
Francesco Collivignarelli https://orcid.org/0000-0001-7442-9051

References

Büttner G 2014 CORINE land cover and land cover change products *Remote Sensing and Digital Image Processing* vol 18 (Springer) pp 55–74 (available at: www.scopus.com/inward/record.uri?eid=2-s2.0-84979917650&doi=10.1007%2f978-94-007-7969-3_5&partnerID=40&md5=f11269711a298cf8833f603a9d1e700d)

Cai X, Zhang X and Wang D 2011 Land availability for biofuel production *Environ. Sci. Technol.* **45** 334–9

Foley J A *et al* 2011 Solutions for a cultivated planet *Nature* **478** 337–42

Fritz S *et al* 2019 A comparison of global agricultural monitoring systems and current gaps *Agric. Syst.* **168** 258–72

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. Model. Softw.* **31** 110–23

Gaupp F, Hall J, Hochrainer-Stigler S and Dadson S 2020 Changing risks of simultaneous global breadbasket failure *Nat. Clim. Change* **10** 54–57

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

Justice C, Gutman G and Vadrevu K P 2015 NASA land cover and land use change (LCLUC): an interdisciplinary research program *J. Environ. Manage.* **148** 4–9

Lee H, Calvin K, Dasgupta D, Krinner G, Mukherji A, Thorne P, Trisos C, Romero J, Aldunce P and Barrett K 2023 Climate change 2023: synthesis report. Contribution of working groups I, II and III to the sixth assessment report of the intergovernmental panel on climate change (The Australian National University)

Liu J, Fritz S, van Wesenbeeck C F A, Fuchs M, You L, Obersteiner M and Yang H 2008 A spatially explicit assessment of current and future hotspots of hunger in Sub-Saharan Africa in the context of global change *Glob. Planet. Change* **64** 222–35

Lobell D B, Thau D, Seifert C, Engle E and Little B 2015 A scalable satellite-based crop yield mapper *Remote Sens. Environ.* **164** 324–33

Olofsson P, Foody G M, Herold M, Stehman S V, Woodcock C E and Wulder M A 2014 Good practices for estimating area and assessing accuracy of land change *Remote Sens. Environ.* **148** 42–57

Pielke R A 2005 Land use and climate change *Science* **310** 1625–6

Potapov P, Turubanova S, Hansen M C, Tyukavina A, Zalles V, Khan A, Song X-P, Pickens A, Shen Q and Cortez J 2022 Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century *Nat. Food* **3** 19–28

Rembold F, Meroni M, Urbano F, Csak G, Kerdiles H, Perez-Hoyos A, Lemoine G, Leo O and Negre T 2019 ASAP: a new global early warning system to detect anomaly hot spots of agricultural production for food security analysis *Agric. Syst.* **168** 247–57

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 M and Fritz S 2006 A method to compare and improve land cover datasets: application to the GLC-2000 and MODIS land cover products *IEEE Trans. Geosci. Remote Sens.* **44** 1740–6

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

Thenkabail P S *et al* 2009 Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium *Int. J. Remote Sens.* **30** 3679–733

Van Tricht K *et al* 2023 WorldCereal: a dynamic open-source system for global-scale, seasonal, and reproducible crop and irrigation mapping *Earth Syst. Sci. Data* **15** 5491–515

Verburg P H, Neumann K and Nol L 2011 Challenges in using land use and land cover data for global change studies *Glob. Change Biol.* **17** 974–89

You L, Wood S and Wood-Sichra U 2007 Generating plausible crop distribution and performance maps for Sub-Saharan Africa using a spatially disaggregated data fusion and optimization approach (FPRI: International Food Policy Research Institute) (available at: www.ifpri.org/publication/generating-plausible-crop-distribution-and-performance-maps-sub-saharan-africa-using)