

You, Liangzhi; Sun, Zhanli

Article — Published Version

Mapping global cropping system: Challenges, opportunities, and future perspectives

Crop and Environment

Provided in Cooperation with:

Leibniz Institute of Agricultural Development in Transition Economies (IAMO), Halle (Saale)

Suggested Citation: You, Liangzhi; Sun, Zhanli (2022) : Mapping global cropping system: Challenges, opportunities, and future perspectives, Crop and Environment, ISSN 2773-126X, Elsevier, Amsterdam, Vol. 1, Iss. 1, pp. 68-73,
<https://doi.org/10.1016/j.crope.2022.03.006>

This Version is available at:

<https://hdl.handle.net/10419/251854>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by-nc-nd/4.0/>



Review

Mapping global cropping system: Challenges, opportunities, and future perspectives

Liangzhi You^{a,b,*}, Zhanli Sun^c^a International Food Policy Research Institute, 1201 Eye Street, NW, Washington, DC, 20005, USA^b Macro Agriculture Research Institute, College of Economics and Management, Huazhong Agricultural University, Wuhan, Hubei, 430070, China^c Leibniz Institute of Agricultural Development in Transition Economies (IAMO), Theodor-Lieser-Str. 2, 06120, Halle, Saale, Germany

ARTICLE INFO

Keywords:

Cropping system
Food security
Global crop mapping
Remote sensing
SPAM

ABSTRACT

Spatially explicit global cropping system data products, which provide critical information on harvested areas, crop yields, and other management variables, are imperative to tackle current grand challenges such as global food security and climate change. These cropping system datasets are also very useful for researchers as they can support various scientific analyses in research projects. Yet, effectively searching, navigating, and fully understanding various global datasets can be a daunting task for researchers and policy analysts. In this review, we first compare a few selected global data products, which use crop census and statistical data as the main data source, and identify key problems and challenges of the global crop mapping such as data accuracy and consistency. We then pointed out the future perspectives and directions in further improving the global cropping data products. Collective mechanisms and efforts with the support of open-access data hosting platforms, standard protocols, and consistent financial support are necessary to produce high-quality datasets for researchers, practitioners, and policymakers. Moreover, machine learning and data fusion approaches can also be further explored in future mapping exercises.

1. Introduction

Agriculture has provided livelihood and fed the world for thousands of years, and agricultural activities have dramatically transformed the biosphere in terms of land-use change, freshwater withdrawal, nitrogen cycle (e.g. fertilizer use in agriculture), and biodiversity loss. Agriculture is thus at the heart of a few intertwined global challenges such as food security, climate change, poverty reduction, environment and ecosystem protection, and biodiversity preservation. Providing enough food for the seven billion people on the planet and at the same time reducing the harmful effects of agricultural production are some of the biggest challenges in the 21st century (Foley et al., 2011). Currently, 9% of the world population i.e. around 700 million people are suffering severe food insecurity, despite more than enough food having been produced for everyone worldwide (Anderson et al., 2015; Roser and Ritchie, 2019). Agriculture plays a direct role in achieving the second Sustainable Development Goal (SDG 2), Zero Hunger, which seeks to simultaneously address food security challenges and global environmental sustainability (Blesh et al., 2019).

In the future, agriculture needs to produce more to meet the ever-increasing food demand, which is expected to be doubled by 2050 due to the rising population and, more importantly, the shifts in consumption patterns from a diet higher in starchy foods to one that is richer in protein, including meats and dairy products—which demand more land to produce—as the affluence increases. Up to 1.3 billion people may be at risk of food insecurity including undernutrition and malnutrition in 2050 in present low-income economies, if they fail to increase agricultural productivity, secure enough croplands for food production instead of renewable energy, and/or imports from other countries (Baldos and Hertel, 2014; Fader et al., 2013).

On the other hand, agriculture and associated land use and land cover change, e.g. deforestation, contributed over 20% of the global carbon emission (Tubiello et al., 2015). Agriculture is now a major driver behind many environmental threats, and it is a major force driving the environment beyond the “planetary boundaries” (Foley et al., 2011; Rockström et al., 2009). This means that the increase of food production also needs to be sustainable and not at the cost of the environment, i.e. mitigating carbon emission and protecting soil and water resources. At

* Corresponding author. International Food Policy Research Institute, 1201 Eye Street, NW, Washington, DC, 20005, USA.

E-mail address: l.you@cgiar.org (L. You).

<https://doi.org/10.1016/j.crope.2022.03.006>

Received 11 January 2022; Received in revised form 22 February 2022; Accepted 24 February 2022

2773-126X/© 2022 The Author(s). Published by Elsevier Ltd on behalf of Huazhong Agricultural University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

the same time, agricultural production, which is sensitive to weather conditions, namely precipitation, temperature, and natural extreme weather events (e.g. droughts and floods) has been witnessed the adverse effects of the changing climate (Fischer et al., 2005).

Facing all of these challenges, we need good data and information to provide more responsive and evidence-based analysis on interrelated aspects of improving the use and allocation of resources across food, land, water, energy, and environment. Data, in particular spatially explicit data on the scale and intensity of agricultural and pastoral systems are essentially the first order of business in the identification of priority land use information for policy-relevant research. This information is essential as a starting point for beginning to understand land use patterns in a way that is meaningful from a quantitative, analytical perspective. The design of policy mechanisms for reducing poverty and vulnerability is dependent on targeting rural economies and populations, and particularly in stimulating agricultural growth. The efficacy of any plans developed along these lines is dependent upon knowing where and how agriculture is practiced. Cropland, which refers to the land area that is arable and under permanent crops according to the Food and Agriculture Organization (FAO) of the United Nations, takes about 1.6 billion hectares of land, accounting for 12% of the earth's terrestrial surface (FAOSTAT, 2021). Croplands are the basis for food production and supply over 90% of all food calories (Kastner et al., 2012). While we have relatively good data on the cropland distribution, the data on the crop types, yields, and management at fine resolution and global scale are still lacking. Data on agricultural production (i.e. harvested area, production quantity, and yield) are usually collected and reported at national and sub-national geo-political boundaries (e.g. provinces and districts), but this level of statistics does not give a sense of the diversity and spatial patterns in agricultural production and is not spatially explicit which is critical for many environmental and biophysical models. These data gaps undermine our efforts in addressing the grand challenges, e.g. food security (Annan, 2018). Thus, there is pressing demand in creating well-suited global gridded cropping systems data products with adequate resolutions, both spatial and temporal resolutions, and with areas, yields, and, even better, management variables of multiple major crops.

Despite the importance and necessity of geospatial data on cropping systems, information on the distribution and performance of specific crops is often available only through national or sub-national statistics. There are some attempts to produce global crop maps. SPAM2000 produces the first global crop distribution for 20 major crops for the year circa 2000 (You et al., 2014). Besides International Food Policy Research Institute (IFPRI), there are three other independent efforts to incorporate the detailed information available from statistical surveys with supplemental spatial information to produce a spatially explicit global dataset specific to individual crops for the year 2000. These global studies have been reported by Fischer et al. (2013), Monfreda et al. (2008), and Portmann et al. (2010). Anderson et al. (2015) compared all the four global cropping system models (including SPAM) and found that there are substantial differences among the model results. The recent rapid increase in public availability of fine to moderate resolution satellite imagery, namely from the Landsat OLI and Copernicus Sentinel-1 and Sentinel-2 sensors, provides a unique opportunity to improve cropping system datasets. And yet the applying at a global scale at individual crop types remains a challenge.

Kim et al. (2021) recently conducted a systematic and comprehensive review on the global gridded cropping system data products, including global and regional data products; multiple-crops and single crop (e.g. only rice) products; datasets including multiple variables such as area, yields; and intensity and datasets with a single variable (e.g. only yield); and census-based datasets and remote sensing and other source-based datasets. The review gave a sweeping overview of the literature and datasets in the mapping of global cropping systems. While the review covered a broad range of data products, the analysis and comparison are understandably lack in-depth assessment on the particular type of data products given the wide disparity of various data products.

Here we build upon this review and focus on the global gridded cropping system data products that are global scale, mainly used census and statistical data as inputs, and covering multiple crops. The global scale multi-crops datasets, reflecting the spatial distribution and heterogeneity of the various crop productions, are particularly imperative to study the cross-regional interactions such as the tele-coupling of the world food systems (Levers and Müller, 2019), the spatial decoupling of the food production and consumption (Fader et al., 2013), and long-term interactions between crop production and climate change (e.g. Sacks et al., 2010). Compared to remote sensing-based global products, the statistical data-based products can provide much richer information on the crop types and more accurate yields. In addition, the farm management data (e.g. fertilizer inputs) and socioeconomic data (e.g. population and income) coming with the statistical data are relatively easy to be incorporated and combined for further in-depth analysis. Multiple crop type classification over a large geographic extent remains a challenge for remote sensing-based approaches, particularly for smallholder farming where land is quite fragmented and cropping is complex (Anderson et al., 2015; Song et al., 2017).

This review thus looks into the evolution of the census-based global cropping system mapping by comparing a few key datasets, identifies existing problems and challenges, and points out the future directions of this line of research. We only focus on crops, while livestock, forestry, grassland, etc. are not included. In the following sections, we first review the existing efforts of global crop mapping, then we describe the current state of the art in global cropping maps. In Section 3, we look into the existing weakness and shortcomings of the crop mapping products, briefly discuss the challenges. In Section 4, we point out the opportunities and future perspectives in global cropping system mapping. Finally, we end the paper with a conclusion section.

2. Current status of global census-based cropping system mapping

Recent studies have generated increasingly sophisticated results portraying the down-scaling of crop production statistics or agricultural census data across regional and global extents at a moderately high spatial resolution. Here we focus on three major global datasets: M3 (Monfreda et al., 2008), MIRCA (Portmann et al., 2010), and spatial production allocation model (SPAM) (You et al., 2014). We briefly describe the input data and methods underpinning these studies and attempt to summarize and compare their key attributes as well as their likely strengths and weaknesses with regard to the range of applications to which their results may be applied.

LRF (Leff et al., 2004) synthesized satellite-derived land cover data and agricultural census data worldwide. Then LRF further allocated the 18 major crops to a global 5 Arc minute grid using simply the share of the harvested area of each of the 18 crops in the statistics reporting unit. After masking non-cropland areas and applying a smoothing algorithm to correct for abrupt and arbitrary changes across administrative boundaries, LRF intersected the resulting proportions of individual crops with a global cropland distribution data set (Ramankutty and Foley, 1998) to obtain an assessment of the per-pixel proportion of each of the crops within the cropland extent. Ramankutty et al. (2008) further improved on this work and developed a Year 2000 global land cover data set for croplands and pasture on a 5 arc minute spatial resolution by combining agricultural land use inventory data and land cover imagery and, in a companion paper (Monfreda et al., 2008), the development of the M3 dataset was described using pixel-level cropland area shares as weights to distribute administrative crop statistics into pixels for the year 2000 (harvested) of 175 distinct crops of the world (Monfreda et al., 2008). By combining work from Ramankutty et al. (2008), LRF and the global map of irrigation areas (GMIA),¹ MIRCA produced a global monthly growing areas distributions of 26 irrigated crops on the same 5 arc minute grid.

¹ http://www.fao.org/ag/agl/aglw/aquastat/irrigation_map/index.stm.

While the primary focus of past downscaling efforts has been to assess the spatial distribution of crops at the mesoscale or landscape scale (all used a spatial resolution of 5 arc minutes), uniquely SPAM goes further in a number of ways. First, SPAM concerns not just with the distribution of crop production, but also with how crops are grown, the specific farming or production systems. Crop yields, the intensity of input use, and the potential for unintended environmental consequences (externalities) of production are quite different between rainfed and irrigated production systems, and between commercially oriented and subsistence producers within the rainfed systems. Second, SPAM relies on an additional collection of sub-national statistical data with a special focus on developing countries. Third, the SPAM model relies on using further ancillary information including crop prices, population density (CIESIN, IFPRI and WRI, 2004) and crop-specific biophysical suitability (Fischer et al., 2010) to distribute sub-national statistics within the cropland extent (Ramankutty et al., 2008) based on a method known as cross entropy (Golan et al., 1996).

All three datasets draw on the Sustainability and the Global Environment (SAGE) data source (Ramankutty et al., 2008) to define the distribution and area intensity of cropland, and MIRCA and SPAM datasets that distinguish between rainfed and irrigated cultivation draw in the same GMIA (Siebert et al., 2005) to assess the irrigated portion of the crop production. All three models used the FAOSTAT data to provide national totals for cropland area, and the harvested area and average yields of individual crops, while also making great efforts to collect sub-national crop statistics to the lowest administrative level as possible (e.g. up to the third sub-national administrative unit). Since MIRCA used M3 to provide its input data on the spatial allocation of the total area and average yield, it relies initially on the same sources of subnational crop statistics. SPAM relies on a separate collection of sub-national statistical data sources, focusing on increased coverage in African and Asian countries, although both draw from the FAO Agro-MAPS data source (<http://kids.fao.org/agromaps/>). MIRCA and SPAM both use FAO's AQUASTAT and global map of irrigated area (GMIA) for identifying crop-specific irrigation extents and practices, but MIRCA relies on a more complete collection of national/sub-national crop calendars and cropping intensities to produce monthly and not annual crop distribution maps. Because the SPAM approach attempts further disaggregation of its rainfed production into commercial and subsistence categories (beyond

the irrigated production), it needs additional sets of input data. These include crop-specific area shares and yield differences amongst irrigated production and large-scale/commercial and smallholder rainfed production, the spatial differences in the biophysical suitability of individual crops for irrigated and rainfed (commercial and subsistence) production, and the spatial patterns of population density as well as crop prices.

The scope, approach, and results of the M3, MIRCA, and SPAM initiatives are summarized in Fig. 1 and Table 1. In broad terms, the M3 approach covers the most complete coverage of crops (a total of 175 crops including tree and forage crops and managed grasslands) including both harvested area and yield. MIRCA downscales 26 crops each divided into rainfed and irrigated production areas, as well as two aggregate categories of “other annual” and “other perennial” crops. Furthermore, MIRCA, unique amongst the three approaches, applied a temporal downscaling to estimate rainfed and irrigated area by month. SPAM 2000 has the most limited crop coverage, just 20 crops (and an “other crops” category), but SPAM downscales the area and yield of each crop into three different production systems: irrigated, commercial rainfed, and subsistence rainfed.

Table 1 summarizes key features of both the granularity of input data and methodological details of the three approaches. The major determinants of the accuracy and reliability of downscaling efforts are (1) the quality of the cropland extent indicating the physical extent and area intensity of cropland, and (2) the spatial resolution and reliability of the sub-national crop statistics. Since all approaches rely on the same cropland dataset (Ramankutty et al., 2008), a primary difference between M3 and MIRCA on the one hand and SPAM on the other lies in the collection of sub-national production statistics used as input data. M3 has a total of 22106 statistical reporting units globally, of which 56 were national, 2299 were first level sub-national disaggregation (e.g. US state level), and 19751 were second level (e.g. US county level) reporting units. SPAM has 24507 statistical units, of which 251 were national level, 2758 were the first sub-national level, and 21498 were the second sub-national level. SPAM focused its data collection efforts particularly in developing countries. For example, in Africa SPAM used 4150 second-level statistical reporting units, while M3 has only 300 second-level statistical reporting units. MIRCA took M3 downscaled crop data results as its starting point for allocating the total harvested area for each crop into rainfed and irrigated areas. MIRCA team collected detailed information on crop

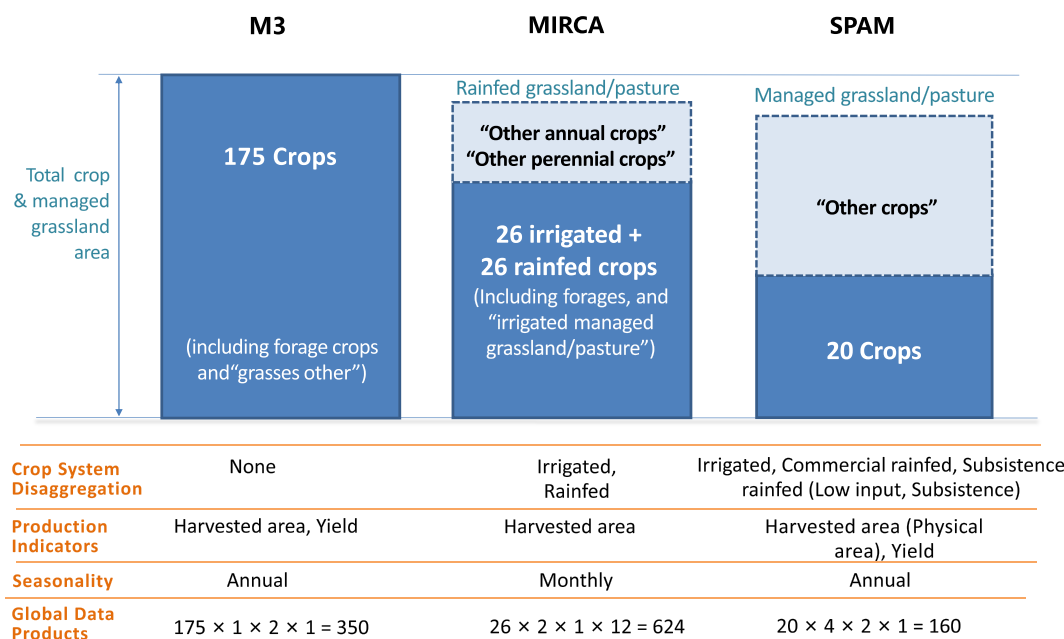



Fig. 1. A diagrammatic comparison of global gridded crop production datasets 2000. Note: M3 – Monfreda et al. (2008), MIRCA –Portmann et al. (2010), SPAM – You et al. (2014). Data products: No. of crops × No. of systems × No. of attributes × No. of products per year.

Table 1
Comparison of major global gridded crop data products.

Feature	M3	MIRCA	SPAM
Statistical Reporting/ Inventory Units (SRU) ^a	22106 (56-2299-19751) ^a	402 (calendar units)	24507 (251-2758-21498) ^a
Cropland extent/Area intensity sources	SAGE/Ramankutty	SAGE/Ramankutty GMIA (Siebert)	SAGE/Ramankutty
Irrigated extent/Area intensity source	n/a	GMIA (Siebert) (and M3 crop distribution)	GMIA (Siebert)
Cropping calendars/ Cropping intensity	n/a	Assembled by MIRCA team	Assembled by SPAM team (for all systems)
Crop-specific area downscaling	Crop-specific share of pixel cropland area = crop area share of total SRU cropland	In proportion to share of irrigated/rainfed crop area in calendar unit (cu) and as f (cropping intensity by cu, and M3 total crop area allocation by pixel)	Crop systems-specific share of pixel cropland area = f (system share of SRU crop production, pixel system-specific area and yield potential, SRU crop value/ha all SRU pixel values)
Allocation algorithm	Crop by crop	Crop by crop	All crops simultaneously
Yield downscaling	Reported SRU crop yield assigned uniformly to all downscaled crop area in SRU	n/a	Yields vary by system and by pixel in proportion to potential system yield, scaled and area weighted to derive reported SRU yield
Data and methods	 <p>More data, more assumptions, more modeling</p>		

Note: M3 – Monfreda et al. (2008), MIRCA –Portmann et al. (2010), SPAM – You et al. (2014).

^a SRU format: (C - 1–2) C - Country-level only, 1 - 1st level admin units, 2 - 2nd level admin units.

calendars, cropping intensities, irrigation, and crop water use. Within each calendar unit the grid cell level apportioning of total crop area into irrigated and rainfed shares, and further into monthly area allocations were made according to these unit-specific data sources. As a source of irrigated areas globally both MIRCA and SPAM used the GMIA Ver 4.0 released in 2007 (Siebert et al., 2005), but the MIRCA and SPAM teams compiled information in the national and sub-national shares of different production systems and cropping intensities independently.

Perhaps the most important methodological differences among M3, MIRCA, and SPAM are in the approaches used to downscale statistical data reported at the administrative unit level into gridcell specific values. M3 uses the most straightforward method. It allocates the total harvested area into each gridcell within the statistical reporting unit using the same proportion of that gridcell cropland area. Crop yield in each grid cell is assigned as being the same as the yield reported for the statistical unit on average. This approach implicitly assumes both environmental conditions and production systems are homogeneous across the cropland extents of each statistical reporting unit, making no allowance for local variation in temperature, rainfall, and soil condition. This approach does not, furthermore, acknowledge the very significant differences between the yield levels of irrigated and rainfed production systems, nor of commercial and smallholder producers within these sometimes large and highly diverse statistical reporting units. MIRCA deals only with the harvested area and essentially uses the relative share of rainfed and irrigated cropland within each grid cell to break out M3 total crop areas into gridcell specific rainfed and irrigated areas (Anderson et al., 2015).

With regard to more recent data products, unfortunately, only SPAM team has continued their effort after producing global crop maps for the year 2000. They expanded their crop coverage to more than 40 crops and covered over 90% of all crop productions in the world. The SPAM team has produced global crop maps for 2005, 2010, and 2017 (for Africa only) so far. The datasets are open access on a dedicated website www.mapspam.info (Yu et al., 2020).

3. Existing problems and challenges

Over the years, the global cropping system data products have made remarkable progress in various perspectives, including the number of crops covered and the accuracy of the spatial downscaling. Yet, many

problems and shortcomings can also be found in the current data products, which hinder the quality and applications of these data products.

First, most data products on global cropping systems are merely snapshots, e.g. the year 2000 of the global cropping systems. As aforementioned, only SPAM has been updated quasi-regularly. This is because the global crop system mapping still relies on official agricultural statistics. For many developing countries particularly African countries, agricultural production statistics even for key staple crops come out late and are problematic, if they exist at all. This lack of timely and reliable information on production volumes impedes food-crisis relief efforts, performance, and accentuates seasonal price volatility. Timely crop information will produce more effective responses to food crises, superior food and agricultural policies, a more informed citizenry, and more efficient commodity pricing. Lack of prospects for regular updates also erodes the trust and confidence in regular users.

Second, both spatial and temporal resolutions can be improved. While 5 arc minute spatial resolution can reasonably support the global-scale analysis, it is not enough in many regional-scale analyses. As for temporal resolutions, even SPAM data produce every few years, mostly at a five-year interval. Many applications, for example, in climate change impacts analysis, requires annual or even monthly time resolution (e.g. Portmann et al., 2010). Without long- or medium-term annual datasets, applications are confined.

Third, most existing global cropping systems data can well reflect the spatial distribution and patterns of crops and their yields. But the data accuracy largely depends on the input data sources, i.e. the crop census and statistical data, which is known to be inconsistent and inaccurate. To make things worse, the statistical data are essentially collected on a national basis. That means the data quality and standard vary country by country. Thus, the data uncertainty is difficult to quantify. And in most cases, there are no data uncertainty assessment and systematic validation at all. Like any models, the results should ideally be validated by a systematic evaluation, however, no product could do such an evaluation because we simply don't know the true spatial distribution of crops at a global scale. Instead each product did some partial validation. For example, SPAM and M3 used additional regional level crop area information to do partial verification. All of them claim that they get feedback and comments from data users. SPAM keeps updating their product once feedbacks are available, and released updated products using different versions of SPAM (e.g. SPAM 2000 Version 2.03).

This may lead to the “garbage in, garbage out” symptoms of many data products. Researchers who are using these data products should be aware of the potential problems of data quality and potential biases resulting from limitations in data availability and quality.

Fourth, most of these cropping system data products were produced independently by a few groups. Although there are loose communications and collaborations among the teams, the data processing and production were carried out by individual team. Even the input data sources are the same, the results, due to the disparity of methods in spatial allocation, can be very different. Collective efforts are still largely lacking.

The reasons behind these problems are manifolds. Global data production requires enormous resources, such as financial support, human resources, computation infrastructure, etc. Most team members are researchers, who usually cannot afford to spend years of effort or simply do not have the long-term financial support to constantly produce and update the global datasets. The good-quality input data, mainly the national and sub-national statistical data, on a global or continental extent are extremely difficult to attain. And many inputs datasets lack reliable data quality assurance and uncertainty assessment. Little can be done from the producing team perspective. All these challenges are holding back the progress of the global cropping system mapping.

4. Opportunities and future perspectives

Having identified the existing weaknesses of the existing global data products of cropping systems and various challenges, we would also like to point out the looming and forthcoming opportunities, which include the increasing availability of data, noticeably the remote sensing data, the rapid development of computational infrastructure (e.g. affordable cloud storage and computation resources), the trending methods in artificial intelligence (e.g. machine learning), and the sprouting organization platforms (e.g. data alliance, for example, data.org) driven the increasing awareness and demand of the importance of the global cropping system data. Recently some quite good global, regional or national datasets on cropping intensity, crop phenology, and harvest areas have been published in peer-reviewed journals (e.g. [Song et al., 2017](#); [Zhang et al., 2021](#)). All of these new developments provide opportunities for future efforts in producing better global cropping systems data products. Innovations in satellite-based remote sensing present opportunities to access more agricultural production data than ever before. The recent rapid increase in public availability of fine to moderate resolution satellite imagery, namely from the Landsat (US NASA) and Copernicus Sentinel-1 and Sentinel-2 sensors (European Space Agency), provides a unique opportunity to improve agricultural production information. New technologies—such as high-resolution photography, unmanned aerial vehicles (UAVs), robust and low-cost solar-powered sensors, rapidly expanding cellular phone networks, combined with machine learning algorithms and vast data processing power—present enormous opportunities to collect data on agricultural production. At the same time, there has been remarkable progress in utilizing satellite-based production estimates in countries such as the United States, though these advances have not translated to less food-secure regions of the world.

Second, global data products on cropping systems need to shift focus on the developing countries, such as Africa. These regions face extreme challenges in food security and high-quality data on crops will greatly help the fight of alleviation of poverty and undernutrition and malnutrition in these regions. However, the data quality in these regions are particularly questionable. There is a pressing need to fill the gaps. In these regions, most production is from smallholder farms with diverse farming practices and widespread intercropping and sequential cropping ([Annan, 2018](#)). It is harder to estimate crop areas and yields in these spatially and temporally heterogeneous areas. Moreover, limited use of fertilizer, irrigation, and pesticides often results in large differences between actual and potential yields. Crop performance models need to be

able to capture these biophysical constraints on crop growth, and their complex and non-linear relationships.

Improving agricultural production information in low-income countries will require not only better utilization of satellite and on-the-ground technologies, but also strong partnerships with local research institutions and governments to ensure that this information is accurate and useful. The aim is to build flexible, practical, locally-adapted, robust, sustainable, and low-cost systems that dramatically improve production estimates in developing countries. This publicly-led data revolution for agriculture will have widespread benefits, particularly for poor people, and is a vital means for achieving the Sustainable Development Goals (SDGs).

Third, big data analysis and machine learning algorithms are promising tools for data processing, particularly on the spatial downscaling stage, during the production of global cropping systems data. All these new emerging methods can complement the current spatial analysis and entropy-based methods to improve the accuracy of the data downscaling. The rapid availability of multiple data sources makes these approaches even more appealing.

Fourth, we call for collective efforts to jointly produce high-quality and comprehensive cropping datasets. For that, an open-access platform that enables broad collaboration from researchers and even other non-academic partners, to work together to share and produce global cropping system data. The platform can draw experiences and inspirations from some existing platforms, for example, the Center for SAGE, and the Global Open Data for Agriculture and Nutrition (GODAN) initiative ([Musker and Schaap, 2018](#)). FAIR principles (Findable, Accessible, Interoperable, and Reusable) should be applied for the open platform. A GitHub-like platform for data producing, instead of collaborative coding, can be very useful to keep the continuous data production and ensure the data quality. A current attempt under this concept has been launched, i.e. Land Use Change Knowledge Information Network (LUCKiNet, www.luckinet.org), led by the German Centre for Integrative Biodiversity Research (iDiv) with a consortium of researchers from 12 countries on 4 continents. They are developing a new system—a collaborative, open-science infrastructure linking various input data streams with analytical pipelines for the continuous production, periodic updating, dynamic improvement, and free online delivery of gridded land-use data products. These next-generation data products will be broad in scope, high in detail, interoperable, quality-assured, and fit-for-purpose. With a recent initiative to develop a 25-year time series of global maps of multiple state variables that together represent the three dominant land-use classes (i.e. cropping, grazing, and forestry), the first steps towards realizing this vision have now been taken. Collaborative workflows to enable iterative improvements of data products and make the data alive. A working group focusing on global cropping systems could be established. A complete open platform allowing broader participation under the defined protocol and standard might be the ultimate solution for producing high-quality and timely global cropping system data.

5. Conclusion

Global cropping system data, providing critical information on areas and yield of crop cultivation, is imperative in tackling current grand challenges, such as global food security, climate change, environmental protection, and ecosystem conservation. Despite the increasing number of global and regional products have been available, many problems and challenges remain. The quality of global cropping system data is still questionable and not consistent across space. Agricultural census data, the main ingredients for crop mapping, are not produced in consistent ways and thus not comparable from country to country, and year to year. The assumptions and method in spatially allocating the crops have significant room to be improved. At the same time, increasing availability of the observation data and statistical data on fine spatial and temporal resolution, development of big data analysis methods, and computation

infrastructure provide unique opportunities for global crop system mapping. Thus, we call for collective efforts under a transparent and standard protocol to produce and gradually improve high-quality and public available global cropping system dataset in contributing to the achievement of the sustainable development goals.

Declaration of competing interest

The author declares that he/she has no competing interests. Given his role as Editors-in-Chief, Liangzhi You had no involvement in the peer-review of this article and has no access to information regarding its peer-review. Full responsibility for the editorial process for this article was delegated to Shaobing Peng.

Acknowledgments

This work received financial support from the German Federal Ministry for Economic Cooperation and Development (BMZ) commissioned and administered through the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) Fund for International Agricultural Research (FIA), grant number: 81260864, with additional support from Startup Fund for Newly Hired Advanced Talents of Huazhong Agricultural University (HZAU) (Grant No. 108-11042210003) and Grant for MARI of HZAU (Project No. 201-11900022403).

References

- Anderson, W., You, L., Stanley, W., Wood-Sichra, U., Wu, W., 2015. An analysis of methodological and spatial differences in global cropping systems models and maps. *Global Ecol. Biogeogr.* 24, 180–191.
- Annan, K., 2018. Data can help to end malnutrition across Africa. *Nature* 555, 7–8.
- Baldos, U.L.C., Hertel, T.W., 2014. Global food security in 2050: the role of agricultural productivity and climate change. *Aust. J. Agric. Resour. Econ.* 58, 554–570.
- Blesh, J., Hoey, L., Jones, A.D., Friedmann, H., Perfecto, L., 2019. Development pathways toward “zero hunger”. *World Dev.* 118, 1–14.
- Center for International Earth Science Information Network (Ciesin); Columbia University; International Food Policy Research Institute (IFPRI); The World Bank; and Centro Internacional de Agricultura Tropical (CIAT), 2004. Global Rural–Urban Mapping Project (GRUMP), Alpha Version. Socioeconomic Data and Applications Center (SEDAC). Columbia University, Palisades, NY.
- Fader, M., Gerten, D., Krause, M., Lucht, W., Cramer, W., 2013. Spatial decoupling of agricultural production and consumption: quantifying dependences of countries on food imports due to domestic land and water constraints. *Environ. Res. Lett.* 8, 014046.
- Fischer, G., Nachtergaele, F.O., Prieler, S., Teixeira, E., Tóth, G., van Velthuisen, H., Verelst, L., Wiberg, D., 2013. Global Agro-Ecological Zones (GAEZ v3.0). <http://gaez.fao.org/Main.html#>. (Accessed October 2013).
- Fischer, G., Shah, M., Tubiello, F.N., Van Velthuisen, H., 2005. Socio-economic and climate change impacts on agriculture: an integrated assessment, 1990–2080. *Phil. Trans. Biol. Sci.* 360, 2067–2083.
- Foley, J.A., Ramankutty, N., Brauman, K.A., Cassidy, E.S., Gerber, J.S., Johnston, M., Mueller, N.D., O’Connell, C., Ray, D.K., West, P.C., Balzer, C., Bennett, E.M., Carpenter, S.R., Hill, J., Monfreda, C., Polasky, S., Rockström, J., Sheehan, J., Siebert, S., Tilman, D., Zaks, D.P.M., 2011. Solutions for a cultivated planet. *Nature* 478, 337–342.
- Golan, A., Judge, G., Miller, D., 1996. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. John Wiley and Sons, New York.
- Kastner, T., Rivas, M.J.I., Koch, W., Nonhebel, S., 2012. Global changes in diets and the consequences for land requirements for food. *Proc. Natl. Acad. Sci. U.S.A.* 109, 6868–6872.
- Kim, K.-H., Kwang-Hyung, K., Yasuhiro, D., Ramankutty, N., Iizumi, T., 2021. A review of global gridded cropping system data products. *Environ. Res. Lett.* 16, 093005.
- Leff, B., Ramankutty, N., Foley, J.A., 2004. Geographic distribution of major crops across the world. *Global Biogeochem. Cycles* 18, GB1009.
- Levers, C., Müller, D., 2019. Mapping export-oriented crop production. In: *Telecoupling*. Springer, pp. 89–113.
- 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. *Global Biogeochem. Cycles* 22, GB1022.
- Musker, R., Schaap, B., 2018. Global open data in agriculture and nutrition (GODAN) initiative partner network analysis. *F1000Research* 7, 47.
- Portmann, F.T., Siebert, S., Döll, P., 2010. MIRCA2000–global monthly irrigated and rainfed crop areas around the year 2000: a new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochem. Cycles* 24, GB1011.
- 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. *Global Biogeochem. Cycles* 22, GB1003.
- Ramankutty, N., Foley, J.A., 1998. Characterizing patterns of global land use: an analysis of global croplands data. *Global Biogeochem. Cycles* 12, 667–685.
- Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F.S., Lambin, E.F., Lenton, T.M., Scheffer, M., Folke, C., Schellnhuber, H.J., Nykvist, B., de Wit, C.A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P.K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R.W., Fabry, V.J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., Foley, J.A., 2009. A safe operating space for humanity. *Nature* 461, 472–475.
- Roser, M., Ritchie, H., 2019. Hunger and undernourishment. <https://ourworldindata.org/hunger-and-undernourishment>.
- Sacks, W.J., Deryng, D., Foley, J.A., Ramankutty, N., 2010. Crop planting dates: an analysis of global patterns. *Global Ecol. Biogeogr.* 19, 607–620. <https://ourworldindata.org/hunger-and-undernourishment>.
- Siebert, S., Döll, P., Hoogeveen, J., Faures, J.M., Frenken, K., Feick, S., 2005. Development and validation of the global map of irrigation areas. *Hydrol. Earth Syst. Sci.* 9, 535–547.
- Song, X.-P., Potapov, P.V., Krylov, A., King, L., Di Bella, C.M., Hudson, A., Khan, A., Adusei, B., Stehman, S.V., Hansen, M.C., 2017. National-scale soybean mapping and area estimation in the United States using medium resolution satellite imagery and field survey. *Rem. Sens. Environ.* 190, 383–395.
- Tubiello, F.N., Salvatore, M., Ferrara, A.F., House, J., Federici, S., Rossi, S., Biancalani, R., Condor Golec, R.D., Jacobs, H., Flammini, A., Prosperi, P., Cardenas-Galindo, P., Schmidhuber, J., Sanz Sanchez, M.J., Srivastava, N., Smith, P., 2015. The contribution of agriculture, forestry and other land use activities to global warming, 1990–2012. *Global Change Biol.* 21, 2655–2660.
- You, L., Wood, S., Wood-Sichra, U., Wu, W., 2014. Generating global crop distribution maps: from census to grid. *Agric. Syst.* 127, 53–60.
- Yu, Q., Yu, Q., You, L., Wood-Sichra, U., Ru, Y., Ru, Y., Joglekar, A.B., Fritz, S., Xiong, W., Wei, X., Xiong, W., Wei, X., Miao, L., Lu, M., Miao, L., Wu, W., Yang, P., Yang, P., Peng, Y., Peng, Y., 2020. A cultivated planet in 2010 – Part 2: the global gridded agricultural-production maps. *Earth Syst. Sci. Data Discuss.* 12, 3545–3572.
- Zhang, M., Wu, B., Zeng, H., He, G., Liu, C., Tao, S., Zhang, Q., Nabil, M., Tian, F., Bofana, J., Beyene, A.N., Elnashar, A., Yan, N., Wang, Z., Liu, Y., 2021. GCI30: a global dataset of 30 m cropping intensity using multisource remote sensing imagery. *Earth Syst. Sci. Data* 13, 4799–4817. <https://doi.org/10.5194/essd-13-4799-2021>.