International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

Address: IIASA, Schlossplatz 1, A-2361 Laxenburg, Austria
Email: repository@iiasa.ac.at **Telephone:** +43 (0)2236 807 342

YSSP Report
Young Scientists Summer Program

Small-scale farms' contribution to food security and sustainable water use:

The development of a 5 arcmin farm-size specific crop map of harvested area and its implication for water management

Han Su (h.su@utwente.nl)

Approved by

Supervisor: Barbara Willaarts, Diana Luna Gonzalez
Program: Biodiversity and Natural Resources Program (BNR)
28th Sept, 2021

This report represents the work completed by the author during the IIASA Young Scientists Summer Program (YSSP) with approval from the YSSP supervisor.

It was finished by ____28/Sept/2021___ and has not been altered or revised since.

This research was funded by IIASA and its National Member Organizations in Africa, the Americas, Asia, and Europe.

Table of Contents

Abstract

Small-scale farms play a critical role in current agricultural production and food security. Previous studies have developed various maps to estimate the geographical distribution of small-scale farms or the geographical distribution of crop production; however, current maps have limited capacities to support climate change and water resources studies where a grided farm size- and crop-specific map is needed. This study aimed to develop a 5 arcmin farm-size specific crop map of harvested area and to explore the implication for the management of water as a global resource. We developed such a map by downscaling a global database that directly measures the crop area and/or crop production per farm size for 42 crops and 56 countries. We downscaled the national data to grid-level by solving an optimization problem, where the objective function aimed to maximize the consistencies between the downscaled map and information from other databases on farm size structure, crop distribution, and field size distribution. We validated the developed map with empirical data from satellite images for oil palm, expert knowledge for coffee in Costa Rica, household surveys on irrigation, and similar estimated maps. Validations show an overall acceptable error. We then estimated small-scale farms' contribution to total agricultural water consumption (blue and green water) using the output of the Global Crop Water Model (GCWM). Results show, under the 2 ha threshold, small-scale farms contribute to 25.8% of total agricultural water consumption globally, and this number is significantly higher in developing countries. Their contribution is also higher in labor-intensive crops (e.g. sweet potato, banana, rice, coconut) and domestic market-oriented crops. This means that the water consumed by small-scale farms may not be virtually exported to other countries. Future work will focus on map development (further validation, inclusion of more countries, estimation of crop production) and the assessment of water use sustainability, productivity (per net weight and per nutrient), and equity.

Acknowledgments

I want to thank my supervisors, Barbara Willaarts and Diana Luna Gonzalez, for their time and efforts, for their support and trust even before the starting of the YSSP program, for their insightful comments and encouragements during the program, for their offers of collaboration on future work. Without them, my study would not have been possible.

I want to thank the Water Security Research Group (WAT) for providing chances and supporting YSSPers to share and discuss research ideas, and for constructive suggestions on my study. A special thanks to Sylvia Tramberend for organizing these activities. A special thanks to Peter Burek and Mikhail Smilovic for the valuable suggestions on the hydrological model and crop model.

I want to thank many excellent IIASA scholars who give me many valuable comments, especially Steffen Fritz, Tamas Krisztin, and Olga Danylo for their availability and interesting exchanges.

I want to thank YSSP Team for organizing all the events. Thanks to Brian Fath and Fabian Wagner for the guidance on YSSP research and knowledge on system analysis. Thanks to Tanja Huber and Aleksandra Cofala for reminding YSSP events every week and making each YSSPer involved. I also want to thank my colleagues and YSSPers for their sharing, being kind, and being open.

I also want to thank my supervisors from the University of Twente, Maarten Krol and Rick Hogeboom, for their support when I first came across the research ideas on small-scale farms and valuable feedback on progress.

Last but not least, I want to thank my wife and my daughter for their understanding and support with a nice working environment at home.

About the authors

Han Su is a Ph.D. candidate at the Multidisciplinary Water Management group of the University of Twente. His primary research interest is the sustainability of water as a global resource. His Ph.D. project focuses on the sustainability, equity, and productivity of water use related to the food supply. Before he started his Ph.D. project, he worked as a researcher and project manager to develop an urban recycled water and rainwater management supporting system for a local government. He holds a Bachelor’s degree in Environmental Engineering from Southwest Jiaotong University and a Master’s degree in Environmental Science from Peking University. He wrote tens of articles in Chinese social media to introduce the advances in environmental research and sustainability science.

Introduction

Small-scale farmers are of great importance for agricultural production and food security. Their exact contribution to agricultural production and food security depends on the definition of small-scale farmers in various contexts. At the global level, studies often use the 2 ha threshold to outline the picture of small-scale farmers despite its limitations (FAO, 2014). Under this threshold, small-scale farms account for 84% of the total number of global farms and support the livelihood of more than 2 billion people (Bosc et al., 2013; FAO, 2014, 2015; Gomez y Paloma et al., 2020). Recent global empirical study shows small-scale farmers use 24% of agricultural land, produce around 34% of global cereals, 17% of vegetables, 50% of pulses, 42% of fruits, and supply 30% of global food (Ricciardi et al., 2018a).

The scale of agricultural production does not directly dictate agricultural practices (e.g., planting, harvesting, irrigating), but it can be said that farmers with similar farm-size share similar patterns of agriculture production and social-economic conditions. Studies show that small-scale farmers plant more cereals, fruits, pulses, and roots and tubers; medium-scale farmers plant more vegetables and nuts; and large-scale farmers plant more oil crops and other cash crops (Ricciardi et al., 2018a). Small-scale farmers tend to increase the use of non-fixed inputs, such as fertilizers and pesticides to increase agriculture production, while large-scale farmers tend to increase fixed inputs, such as machinery (Ren et al., 2019). This helps to explain the overuse of fertilizers among small-scale farmers in some countries (Sheahan and Barrett, 2017; Wu et al., 2018). In the water scarcity region, small-scale farmers have less irrigation than non-small-scale farmers (Ricciardi et al., 2020). In developing countries, they are often found in subsistence conditions, featured by low profit and yield, limited market access (Meemken and Bellemare, 2020).

The characteristics of small-scale farmers are also highly context-dependent. For example, small-scale farmers in Europe can be both weak and strong market-oriented depending on their location and type of crops (Guarín et al., 2020; Rivera et al., 2020). Even in developing countries where they often have limited market access, small-scale farmers may play a critical role in export-oriented crops. For example, in Indonesia, the world's largest palm oil producer and exporter, small-scale farmers plant 38% of palm oil (Bakhtary et al., 2021). In Vietnam, the world's second-largest coffee producer and exporter, small-scale farmers plant 54% of coffee (ICO, 2019). Small-scale farmers still perform different agriculture practices compared to large-scale producers in these countries (Bakhtary et al., 2021).

Given the need to differentiate farm size in agriculture studies, significant progress has been made in mapping small-scale farmers at the global level. Samberg et al. (2016) put the first effort in mapping the geographic distribution of small-scale farmers, where they used the Mean Agricultural Area (MAA) from agriculture census and household survey to classify farm size for subnational administrative units. They overlapped their map with crop map, Monfreda et al. (2008), to estimate crop production per farm size. The MAA is calculated by dividing the total cropland area by the number of households. Since the number of small-scale farms is always large, using MAA as farm size to classify subnational administrative units will overestimate small-scale farms (Ricciardi et al., 2018a). Fritz et al. (2015); Lesiv et al. (2019) developed a grided global field size dominant map using manually labeled field size data on the satellite image and spatial interpolation. In agriculture census, the classification of a small-scale farm is based on the total operated or cultivated land by a household, thus, using field size to identify small-scale farmers overestimate the number of small-scale farmers, too. Herrero et al. (2017) used the country level farm size data from Lowder et al. (2016) and Fritz et al. (2015) to develop a dominant farm size map and overlapped the map with a crop map from Ray et al. (2013) to estimate the farm-size specific crop area. This map was updated by Mehrabi et al. (2020) using Lesiv et al. (2019) to providing a global grided dominant farm size map at 10km2 resolution. Mehrabi et al. (2020) further overlapped the map with Monfreda et al. (2008) to estimate crop production per farm size. Since the focus is on the dominant farm size, potential overestimation issues remain unsolved in Mehrabi et al. (2020). Different from the above maps, which estimated farm-size specific crop production indirectly and have potential overestimations, Ricciardi et al. (2018a); Ricciardi et al. (2018b) established an empirical global database using agriculture census and household survey that directly measure crop production or area and farm size. Ricciardi et al. (2018a); Ricciardi et al. (2018b) cover half of the global cropland for 56 countries[1] – with subnational data for 46 countries.

Ricciardi et al. (2018a); Ricciardi et al. (2018b) products, however, have limited abilities to fulfill the need of global climate change and water resources studies where the hydrological model requires a grided crop map. The farm size has a close link to export. The water embodied in the international food trade makes water is a global resource: one country could reduce water stress by importing water-intensive products (Hoekstra, 2020); thus, distinguishing the farm size contributes to estimating to what extent the local water scarcity is one part of global water management issues or local water management issues.

This study aimed to develop a 5 arcmin grided farm-size specific crop map of harvested area by downscaling Ricciardi et al. (2018a). The reason to use Ricciardi et al. (2018a) is that this dataset is the most complete empirical dataset that directly measures crop area or production and farm size. We achieved the downscaling by solving an optimization problem for each administrative unit to maximize the consistencies between the downscaled map and different cropping maps. Based on the downscaled map, we also estimated the small-scale farmers' contribution to total agricultural water consumption. We discussed the reliabilities of the developed map and the implications of results for future studies on water as a global resource.

Data and methods

We developed the 5 arcmin farm-size specific crop map of harvested area, across 42 crops, 11 classes of farm size, and 4 farming systems circa 2010, by downscaling the crop-specific farm size structure from Ricciardi et al. (2018a) with the crop distribution from SPAM2010 (Yu et al., 2020) and field size distribution from Lesiv et al. (2019) (Fig. 1, Table 1). Here, farm size is defined as the total operated or cultivated area by a holding or household. Since the definition of small-scale farms depends on research and policy context and an overview of these definitions can be found in Khalil et al. (2017), to increase the flexibility of our map to various definitions of the small-scale farmer, we distinguished 11 classes of farm size in our map defined by the World Census of Agriculture (WCA). For illustration purposes, in this study, we used the 2 ha threshold for small-scale farms which is widely adopted as a threshold by the global studies.

Given the inconsistencies between different datasets, we tackled the development of a crop-specific harvested area map per farm size as an optimization problem, aiming to use information from these datasets as much as possible and maximizing consistencies with them.

[1] In their paper, they claim to have data for 55 countries. In the dataset they published, it contains the 56th country, Czechia.

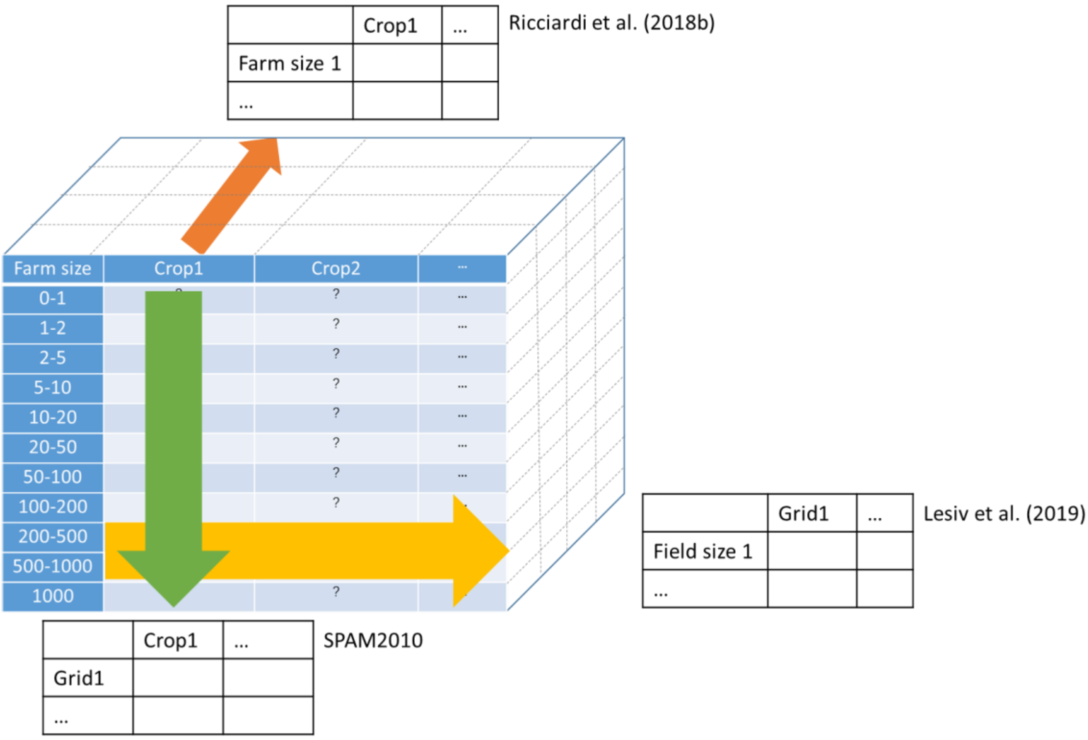


Fig. 1 The three databases used in this study to develop a 5 arcmin farm-size specific crop map of harvested area

Table 1 Detailed information on the three databases used in this study to develop a 5 arcmin farm-size specific crop map of harvested area

Dataset	**Indicator**	**Spatial resolution**	**Time**	**Crops**	**Note**
SPAM2010	Harvested area [ha]	5 arcmin	2010	Reclassify 161 FAO crops into 42 SPAM crops	Include 4 farming systems
Ricciardi et al. (2018a)	Cropping area [ha]	National or subnational administrative unit	The data source for each country ranging from 2001 to 2015	154 FAO crops	Include 11 farm sizes
Lesiv et al. (2019)	Dominant field size	1*1 km	The data source for each location ranging from 2000 to 2017	Not crop-specific	Include 5 field sizes

Data sources and preprocessing

SPAM2010 maps physical area, harvested area, and production for 42 crops in 2010 under 4 farming systems (irrigated, high input, low input, and subsistence rainfed farming systems) with a resolution of 5 arcmin. It was developed based on the CAAS-IFPRI cropland extent map (Lu et al., 2020) and national and subnational agricultural census on crop-specific data. SPAM2010 downscaled the national and subnational level data into a grid level considering social-ecological factors including crop suitability and market accessibility. SPAM2010 also reported the country-level data as FAO reports it. The 42 crops include 33 individual crops and 9 aggregated crops to cover all the FAO crops, except fodder crops (for crop classification, please find it in Yu et al. (2020)). In this study, we used the grided harvested area of each crop and crop classification from SPAM2010.

Ricciardi et al. (2018a) collected agricultural inventories and household surveys that directly measure the cropping area and/or production per farm size across 154 FAO crops for 56 countries covering half of the global cropland (Fig. 2). The cropping area in this dataset means either crop area, planted area, harvested area, or cultivated area. In this dataset, farm size was classified into 11 categories according to WCA: 0-1 ha, 1-2 ha, 2-5 ha, 5-10 ha, 10-20 ha, 20-50 ha, 50-100 ha, 100-200 ha, 200-500 ha, 500-1000 ha, and >1000 ha. We used the lower bound of each farm size classification to represent the 11 respective classes. The reported year is different from country to country ranging from 2001 to 2015 with the median year as 2013.

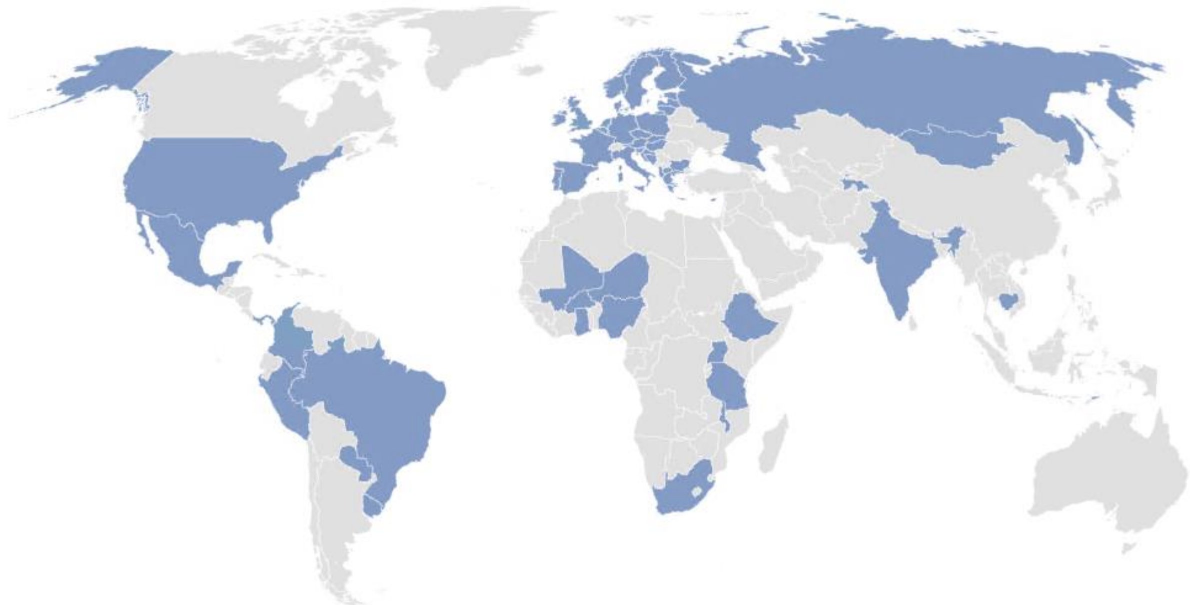


Fig. 2 The 56 countries covered by Ricciardi et al. (2018a) and this study

Data from Ricciardi et al. (2018a) was preprocessed to estimate the farm size structure, i.e. the percentage of 11 farm sizes in the total harvested area for each crop, at the administrative level in 2010. Among 42 SPAM crops, 38 crops are estimated by aggregating FAO crops in Ricciardi's dataset. The remaining 4 SPAM crops - pearl millet and small millet, and arabica coffee and robusta coffee - used the data on crop millet and coffee (green) from FAO. We basically relied on cropping area to calculate the percentage of 11 farm sizes in the total harvested area. Since the cropping area in this dataset include 4 items (crop area, planted area, harvested area, or cultivated area), when more than one item is available for the cropping area, we used the item with a larger overall area (after aggregation). If all 4 items are not available, we used production data to estimate the percentage of 11 farm sizes in the total harvested area. In this case, the underlying assumption is constant yield across farm sizes.

The data from Lesiv et al. (2019) indicates the dominant field size in a 1*1 km grid. Field sizes were classified into 5 categories: < 0.64 ha, 0.64-2.56 ha, 2.56-16 ha, 16-100 ha, and >100 ha. Here, field size is different from farm size since several fields may belong to one farm. This map was developed based on manually labeled crowdsourcing data with the time reference ranging from 2000 to 2017. We used it to estimate the lower bound area of farm sizes in 2010. We interpret "dominant" field size as the respective field size accounting for at least 50% of the cropland of the grid. To keep consistent with SPAM2010, we first extracted available field size information to the CAAS-IFPRI cropland extent map with a spatial resolution of 500*500 m. We then calculated the lower bound area for farm sizes by summing field area up to the 5 arcmin grid using cropland extent, lower bound of dominant field size type, and 50% threshold at the resolution of 500*500 m.

Prioritizing datasets

Among the three datasets, we prioritized SPAM2010 and Ricciardi et al. (2018a) over Lesiv et al. (2019).

We prioritized SPAM2010 because the development of SPAM2010 is based on very detailed agricultural censuses. It was developed for the year 2010. Also, SPAM2010 is widely used in various global models and it has been adjusted to FAOSTAT. Prioritizing SPAM2010 can maximize the potential application of our map in global agricultural studies. We ensured the total harvested area across 11 farm sizes would be the same in SPAM2010 for each crop and grid.

We prioritized Ricciardi et al. (2018a) dataset because this dataset provides the crop-specific proportion of 11 farm sizes at the administrative level based on agricultural census. We acknowledge that some of this information may not be close to the situation in 2010 and bias from the estimated data based on household surveys. Thus, we only ensured the relative differences between the developed map and Ricciardi et al. (2018a) are within 10%.

Data from Lesiv et al. (2019) is worth being included since it tells us where the large farms are located. This holds only for large farms because large fields indicate large farms but relative small fields could belong to both small and large farms. The spatial resolution of this map is 1*1 km, however, the grid-level data was estimated from several labeled data points within its neighborhood. The distance between the farthest labeled data and a grid ranges from 3 to 20 km. Considering our spatial resolution is about 10 km, field size information from this map may be uncertain, thus, we prioritized the other two datasets over this one when we came across inconsistencies.

Establishing optimization problems to and maximize consistencies

For each administrative unit, a, defined in Ricciardi's dataset, we solve the following optimization problem.

Sets:

c	Crops	$\lvert c \rvert = 42$
f	Farm size	$\lvert f \rvert = 11, f = list(0, 1, 2, 5, 10, 20, 50, 100, 200, 500, 1000)$
e	Field size	$\lvert e \rvert = 5, e = list(0, 0.64, 2.56, 16, 100)$
s	Farming system	$\lvert s \rvert = 4$
a	Administrative unit	$\lvert a \rvert = 3421$
g	Grid	$\lvert g \rvert = 832827$
l	Elastic factor	$\lvert l \rvert = 8, l = list(1, 1/2, 1/4, 1/8, 1/16, 1/32, 1/64, 0)$ by order

Parameters:

$ha.R_{c,f,a}$ percentage of harvested are of farm size f in crop c from administrative unit a estimated from Ricciardi's dataset

$ha.S_{c,s,g}$ harvested area of crop c for farming system s at grid g from SPAM2010

$ha.L_{e,g}$ the lower bound cropland area of field size e at grid g

Variables:

$ha_{c,f,s,g}$ harvested area for crop c, farm size f, farming system s at grid g estimated in our model.

Objective function:

The objective is to maximize the consistencies between the downscaled map and Ricciardi's dataset:

$$min \sum_{c,f} abs\left(ha.R_{c,f,a} \sum_{s,g \in a} ha.S_{c,s,g} - \sum_{s,g \in a} ha_{c,f,s,g}\right) \tag{1}$$

Constraints:

The first constraint ensures that the total harvested per crop per grid in our map equals the harvested area per crop per grid indicated in SPAM2010. In this way, we kept perfect consistency with SPAM2010:

$$\sum_{f} ha_{c,f,s,g} = ha.S_{c,s,g}, \forall c, s, g \tag{2}$$

The second constraint ensures that the relative differences between our map and Ricciardi's dataset are within 10%:

$$90\% * ha.R_{c,f,a} \sum_{s,g \in a} ha.S_{c,s,g} \leq \sum_{s,g \in a} ha_{c,f,s,g} \leq 110\% * ha.R_{c,f,a} \sum_{s,g \in a} ha.S_{c,s,g}, \forall c, f \tag{3}$$

Thirdly, since f is the lower bound of each farm size category, we also applied a minimum allocated area for each farm size (Note 1):

Hard form

$$ha_{c,f,s,g} \geq f, \forall c, f, s, g \tag{4}$$

Soft form

$$\sum_{s} ha_{c,f,s,g} \geq l \times f, \forall c, f, g \tag{5}$$

Fourthly, using Lesiv's dataset, we applied a lower abound for some farm sizes (Note 2).

For farms larger than 100 ha:

$$\sum_{c,s,f\geq 100} ha_{c,f,s,g} \geq ha.L_{100,g}, \forall g \tag{6}$$

For farms larger than 10 ha:

$$\sum_{c,s,f\geq 20} ha_{c,f,s,g} + \frac{20-16}{20-10}\sum_{c,s} ha_{c,10,s,g} \geq ha.L_{100,g} + ha.L_{16,g}, \forall g \tag{7}$$

For farms larger than 2 ha:

$$\sum_{c,s,f\geq 5} ha_{c,f,s,g} + \frac{5-2.56}{5-2}\sum_{c,s} ha_{c,2,s,g} \geq ha.L_{100,g} + ha.L_{16,g} + ha.L_{2.56,g}, \forall g \tag{8}$$

For all farms:

$$\sum_{c,s,f\geq 1} ha_{c,f,s,g} + \frac{1-0.64}{1-0}\sum_{c,s} ha_{c,0,s,g} \geq ha.L_{100,g} + ha.L_{16,g} + ha.L_{2.56,g} + ha.L_{0.64,g}, \forall g \tag{9}$$

Last but not least, we have non-negative area constraints:

$$ha_{c,f,s,g} \geq 0, \forall c, f, s, g \tag{10}$$

Note 1: This constraint is not necessarily required by the definition of farm size because farm size is defined based on the total operated or cultivated area that does not need to be a single crop area and single farming system. We think this constraint is still highly reasonable because we applied it to the grid level which is far larger than a single farm. However, due to inconsistencies, this constraint especially the hard form may make the optimization infeasible. These inconsistencies mainly come from SPAM2010 because it does not consider farm size in its algorithm leading to the grid-level crop-specific harvested area being less than a certain size sometimes. Due to it, when the optimization is infeasible, we first relaxed this constraint instead of the constraint derived from Lesiv's dataset. For the relaxation, we first tried soft form which ignores farming system requirements. If the optimization is still infeasible, we further relaxed the constraints by using the elastic factor from 1 to 0 in order.

Note 2: We relaxed the constraints from large farms to small farms gradually until the optimization is feasible. This case does not happen often in our calculation. We think these infeasible cases reflex the uncertainties to determine the farm location of difference scale among datasets.

Solving procedure

The optimizations were solved by Gurobi v9.1 using the dual simplex method with a time limit of 150s for each administrative unit. Gurobi v9.1 is a fast commercial optimization solver (Gurobi Optimization, 2021). Most of the optimization problems in this study could be solved within 60s with the optimal solutions. The above optimization always had multiple optimal solutions because of the existence of free variables (we do not have enough information for every crop and every farm size). To avoid potential bias of single optimal solutions, we calculated up to 80 (sub)optimal solutions for each optimization and averaged these solutions to get the final one. There may be still bias on the final averaged solution because the number and quality of solutions depend on the searching process of the dual simplex method.

For the administrative units containing more than 300 5-arcmin grids, the optimization problem becomes extremely large posing a great challenge for the solver. The number of decision variables would be more than half-million (11*42*4*300). In this case, we applied a two-tier optimization. We first randomly divided all grids into several groups. Each group includes around 100 grids (for Russia, it is 200 to keep the number of groups less than 300). We first solved the optimization problem at the group level. Then, we solved the grid-level optimization for each group. Of 3421 administrative units, 244 units need to be dealt with in this way -- they cover 89.4% of grids in this study.

Finally, to increase the reliability of our results, we masked the farm size of crops as unknown if these crops are not covered by Ricciardi's dataset. For these crops, the optimization could estimate their farm size components, but the uncertainties are significantly larger than those covered by Ricciardi's dataset.

Estimating water consumption per farm size per crop

We used the output of the Global Crop Water Mode (GCWM) to estimate the water consumption per farm size per crop. GCWM estimate the crop-specific water consumption (mm/ha) at 5 arcmin grid level under 1998-2002 climate conditions. It considers the soil water balance, crop calendar, and crop growth stages (Siebert and Döll, 2010). Here, we used the total water consumption from GCWM including the blue water (surface water, groundwater, and irrigated water) and green water (rainwater). The information on crop-specific farm size is provided by our downscaled map.

Results and discussion

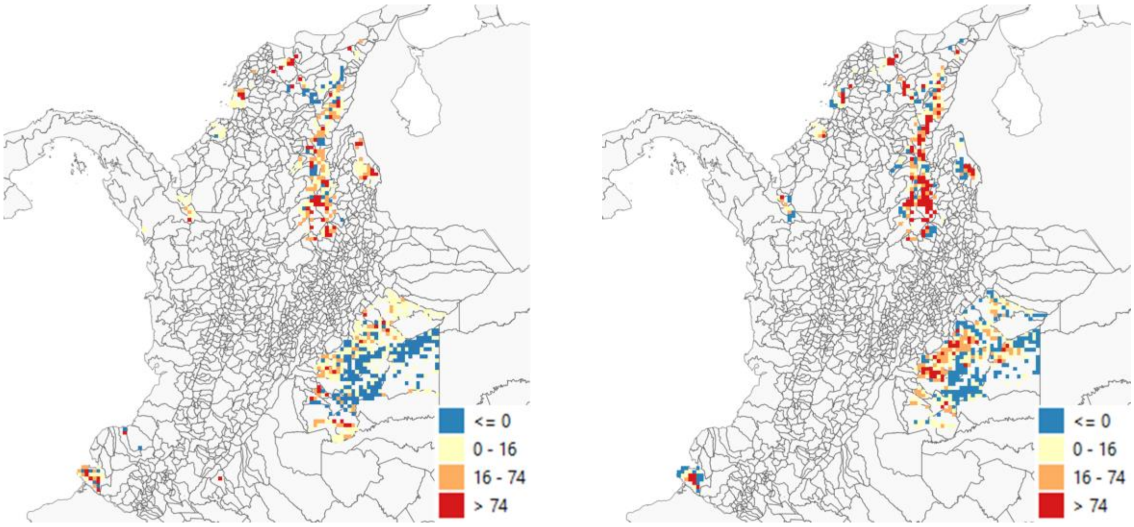
Validation

Comparison with empirical data

Oil palm from satellite images

We compared our map with the farm size distribution for oil palm from Descals et al. (2020). This dataset used deep learning and satellite image to identify the global small-scale oil palm farms and industrial oil palm farms in 2019. They distinguished the small-scale farm and large-scale farm based on landscape, e.g. whether the regular road can be found in the satellite image. Since we only have the farm size information, to have an equivalent definition with Descals et al. (2020), we adopted the Indonesian definition on small-scale oil palm farms and used 20 ha as the threshold, which is mentioned in Descals's paper. Descals et al. (2020) have a spatial resolution of 10 meters. We conducted Zonal Statistics in ArcGIS to calculate the total area of small-scale farms and large-scale farms for each 5 arcmin grid.

Since we do not have oil palm data for every country, we could only compare 5 countries (Table 2). We conducted comparisons at different resolutions and calculate the Pearson correlation coefficient. For 15 arcmin and 25 arcmin resolution, we compared the spatial average on 5 arcmin grids.



(a) (b)

Fig. 3 The small-scale oil palm farm distribution in our map (a) and validation map (Descals et al., 2020) (b) in Colombia. The color indicates the total area (in hectares) of small-scale farms at each grid cell.

Results show a significantly positive correlation with the validation map (Table 2). At the 5 arcmin level, the Pearson correlation coefficient is not large which means the geographical distribution of farm size is not exactly the same. Results from 15-arcmin and 25-arcmin spatial resolution present a stronger correlation which means the patterns of farm size distribution are similar (Fig. 3). For Tanzania, the validation results may be attributed to only a few oil palms are plant in this country, thus, we do not have enough data for validation. For Brazil, the correlation is not as strong as other countries because of 3 reasons. First, the validation map and SPAM2010 have inherent uncertainties. The validation map has an accuracy of 73.8% - 89.4%. SPAM2010 has a correlation coefficient between 0.05 and 0.94 with observed data depending on crops and countries. These errors lead to some inconsistencies as background for our validation. Second, we focus on different years. Our map was developed for 2010 and the validation map was for 2019. The farm size and oil palm distribution may change during the 9 years, especially in developing countries. Third, we adopt a different definition of small-scale farms which affects the identification of small-scale farms.

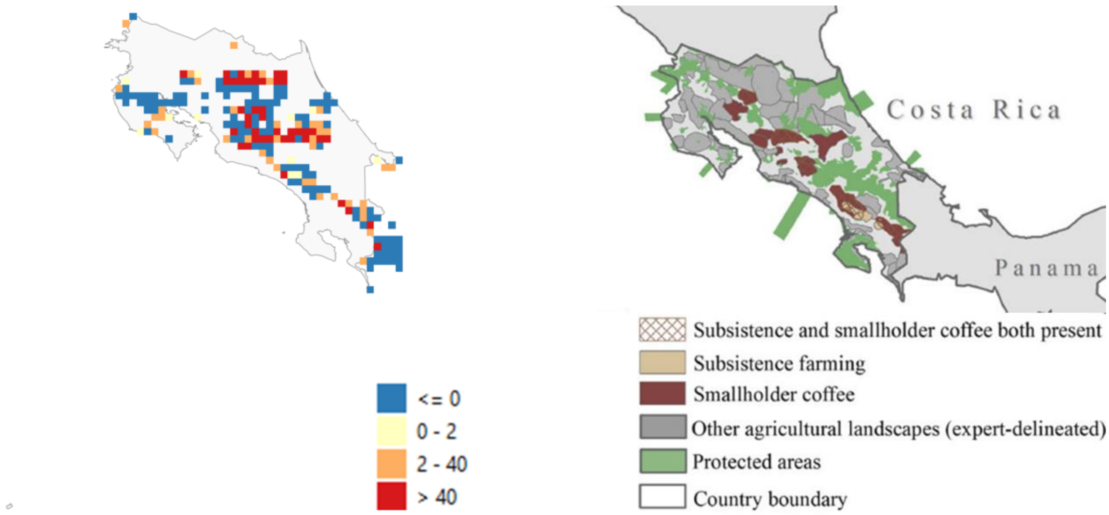
Table 2 Validation with oil palm farm size distribution at 5, 15, and 25 spatial resolution.

	Small-scale farms			**Large-scale farms**		
	5 arcmin	15 arcmin	25 arcmin	5 arcmin	15 arcmin	25 arcmin
Colombia	0.23***	0.56***	0.70***	0.38***	0.62***	0.70***
Costa Rica	0.10**	0.42***	0.62***	0.63***	0.85***	0.93***
Brazil	0.10***	0.21***	0.26***	0.12***	0.09***	0.08***
United republic of Tanzania	0	0.02	0.07	0		
Peru	0.36***	0.38***	0.41***	0.14***	0.21***	0.24***

*** $p<0.0001$
** $p<0.05$

Coffee from expert knowledge

We also compared with the small-scale coffee farm map for Costa Rica developed by Holland et al. (2016), where they applied local expert knowledge to map the small-scale coffee farms at the subnational level for the year around 2013. According to the comparison (Fig. 4), we have a similar small-scale farm distribution in the middle and south of Costa Rica. In the north of Costa Rica, our map indicates additional small-scale farms which are identified as other agricultural farms in Holland et al. (2016). Due to data availability, we are unable to quantify the comparison.



(a) (b)

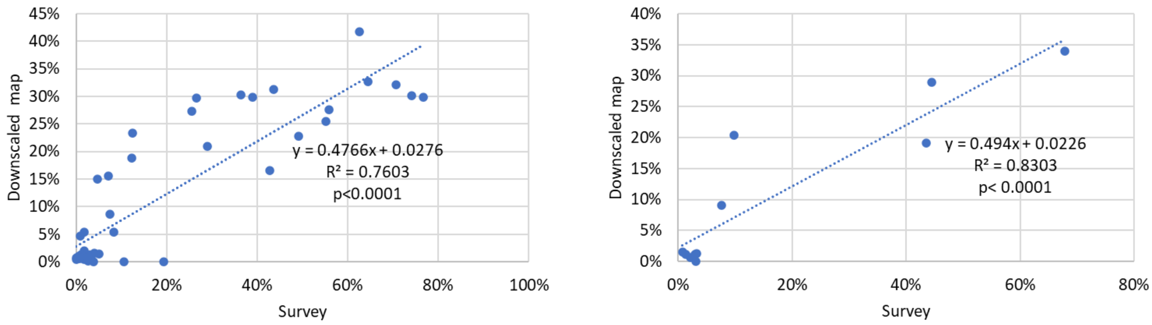
Fig. 4 The small-scale coffee farm distribution in our map (a) and validation map (b) (revised from Holland et al. (2016)) in Costa Rica. The color indicates the total area (in hectares) of small-scale farms at each grid cell.

Irrigation from household surveys

During the downscaling, we not only estimated the harvested area per farm size for each crop but also 4 farming systems. Here, we compared our estimation of the farm size-specific percentage of irrigated areas with observations from household surveys. We used the FAO RuLIS (Rural Livelihoods Information System) to get the micro-level data of household surveys where they provide 54 standardized household surveys from 32 countries. We selected the 11 countries that are covered by our downscaled map and provide crop area, cultivated area, and total irrigated area in the survey. If more than one surveys are available for one country, we used the one conducted around 2010. The list of these surveys can be found in Appendix A1.

With each household survey, we dropped samples that contain any Null value in crop area, cultivated area, and total irrigated area. Then, we classified samples into the 11 farm sizes based on the crop area and calculated the total cultivated area and irrigated area per farm size. For each farm size, we required 5 minimum samples. The percentage of the irrigated area was calculated by dividing the total cultivated area by the total irrigated area.

The comparisons show an overall significant positive relationship between our downscaled map and household survey (Fig. 5 (a)). We notice the regression indicates our downscaled map only has half percentage of irrigated area compared to the household survey. We attribute these systematic inconsistencies to SPAM2010 since SPAM2010 provides the farming system information for our map. This can be seen from the comparison between SPAM2010 and the household survey (Fig. 5 (b)). In SPAM2010, the information on the farming system is directly from national and subnational agriculture censuses or experts. We think the systematic inconsistencies may come from the definitions of irrigated area and cultivated area in the respective dataset. The systematic inconsistencies may also be attributed to the inclusion of industrial farms in SPAM2010 and the exclusion of these farms in the household survey. Besides the systematic inconsistencies, our downscaled map captures the same trends for irrigation in many countries where small-scale farms have a greater percentage of the irrigated area (detailed results can be found in Appendix A2).



(a) (b)

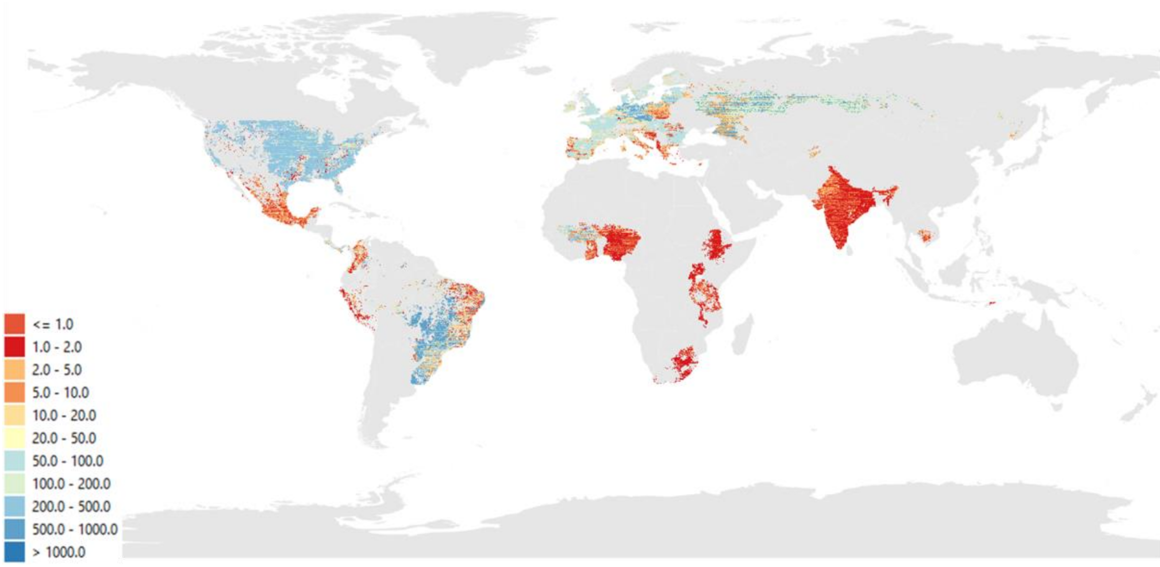
Fig. 5 The comparison of the percentage of irrigated area: (a) differentiates the 11 farm sizes at the country level; (b) only compares the country-level data without differentiating farm sizes.

Comparison with previous estimated farm size distribution map

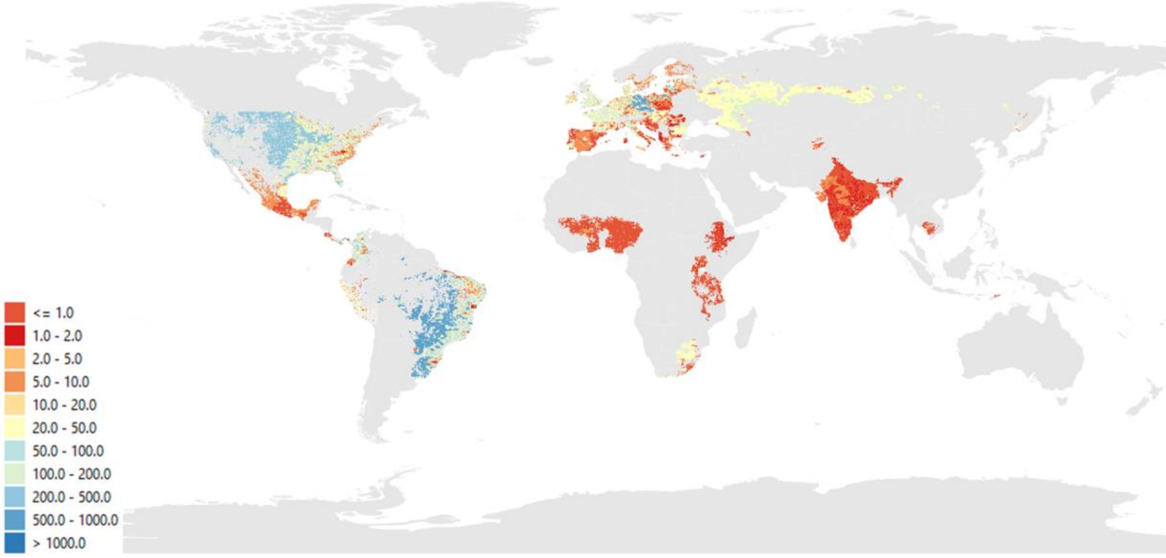
Previous studies focused on developing the farm size distribution map and overlapped the map with the crop map to estimate the crop area per farm size. Among these studies, the map from Mehrabi et al. (2020) represents the state-of-art status which was developed based on the work of Lowder et al. (2016), Herrero et al. (2017), and Lesiv et al. (2019). This map provides the dominant farm size at the 5 arcmin grid level.

To compare with Mehrabi's map, we first estimated the dominant farm size using our downscaled map at the grid level. In our downscaled map, we have multiple farm sizes in one grid. Here, we used the farm size that accounts for the largest area in each grid as the dominant farm size. Then, we compared the two maps pixel-to-pixel. To quantify the comparison, considering the uncertainties in the maps, if the dominant farm sizes in the two maps are next to each other or the same, we counted the comparison as "similar". Otherwise, we compared if the farm size in our downscaled map is larger or smaller than that in Mehrabi's map.

Overall, 53.4% of grids have similar dominant farm sizes; 27.0% of grids have a larger dominant farm size in our downscaled map, and 19.4% of grids have a smaller dominant farm size in our downscaled map (Appendix A3). The results vary from country to country (Fig. 6). We have more similar dominant farm sizes in developing countries. This means, in terms of small-scale farm identification, the two maps have more consistencies. In developed countries where farm sizes are often non-small, our downscaled map tends to indicate a larger dominant farm size than Mehrabi's map. Since our map is developed for the year 2010 and Mehrabi's map basically developed based on the date around 2000. Farm sizes may change a lot within 10 years.



(a)



(b)

Fig. 6 The dominant farm size in our downscaled map (a) and Mehrabi's map (b). Only the grid cells that are covered by both maps are shown.

Some differences between our downscaled map and Mehrabi's map also come from inconsistencies in data sources used to develop the respective map. We obtained the farm size information from Ricciardi et al. (2018a) and kept the relative differences between our map and Ricciardi's data within 10%. Mehrabi's map got the farm size information from Lowder et al. (2016) and remains a good consistent with it at the country level. We compared the proportional farm size distribution per country from our downscaled map and Lower's database (Fig. 7). The comparison shows an overall positive relationship with some disparities (Appendix A4). This means the farm size information from data sources is not exactly the same. These background inconsistencies lead to some differences in dominant farm size in our downscaled map and Mehrabi's.

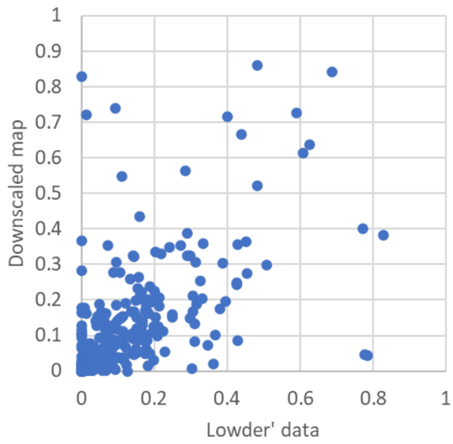


Fig. 7 The proportional farm size distribution per country from our downscaled map and Lower's database.

The overall distribution of small-scale farms

Through downscaling, we estimated the geographical farm size distribution for 42 crops across 56 countries. With the 2 ha threshold, we map the distribution of small-scale farms (Fig. 8). Small-scale farms are mainly distributed in developing countries, such as India, Africa, Central America, South America, and South European countries. The total harvested area of small-scale farms for 42 crops can be found in Appendix A5.

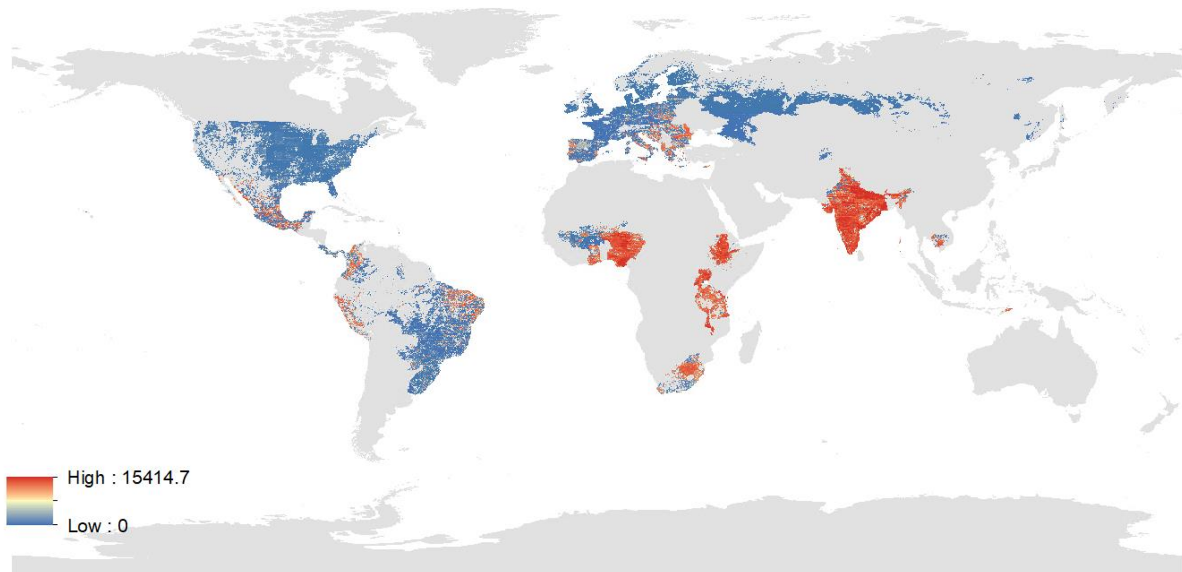


Fig. 8 The total harvested area (in hectares) of small-scale farms and their distribution estimated by this study.

Small-scale farms' contribution to total water consumption

Under the 2 ha threshold, small-scale farms contribution to 25.8% of total agricultural water consumption. This number is significantly higher in developing countries than in developed countries. In developing countries, this number is often higher than 40% while in developed countries (India and African countries), this number is often lower than 5% (Fig. 9).

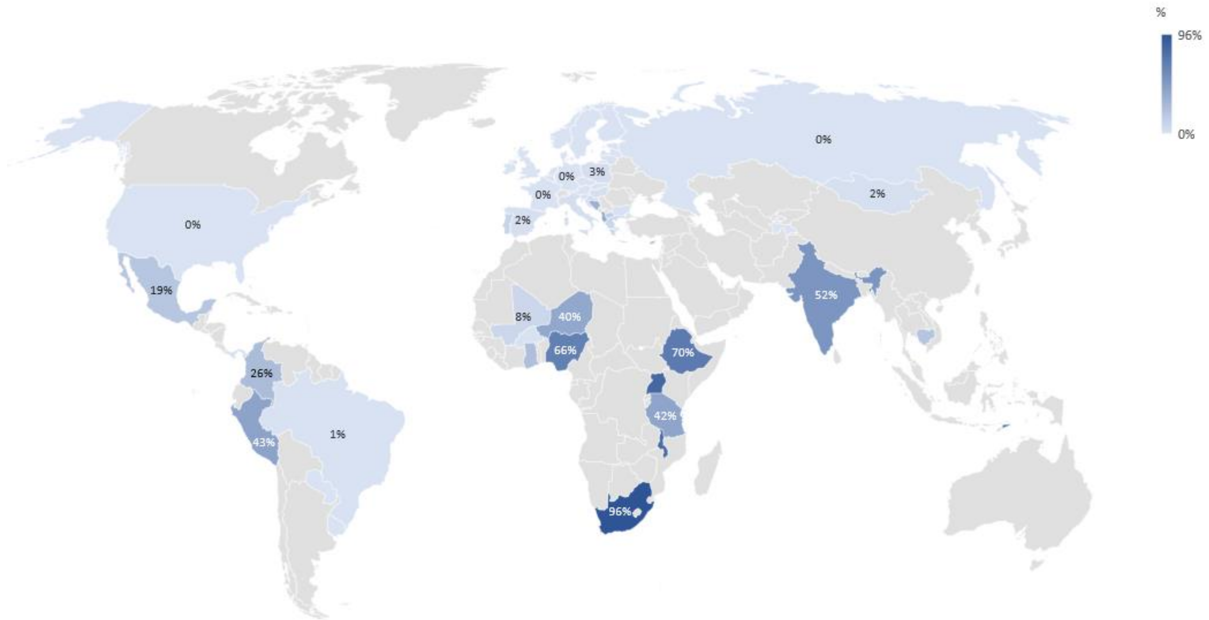


Fig. 9 Small-scale farms' contribution to total local agricultural water consumption in each country.

Small-scale farms' contribution to total water consumption depends on the type of crops (Fig. 10). Their contribution can reach over 60% in some crops, such as sweet potato, robusta coffee, coconut, plantain, cassava, yams, banana, and some fiber crop and roots. Most of these crops are labor-intensive. In some crops, their contribution is below 5%. These crops include soybean, oil palm, barley, sugarbeet, and rapeseed.

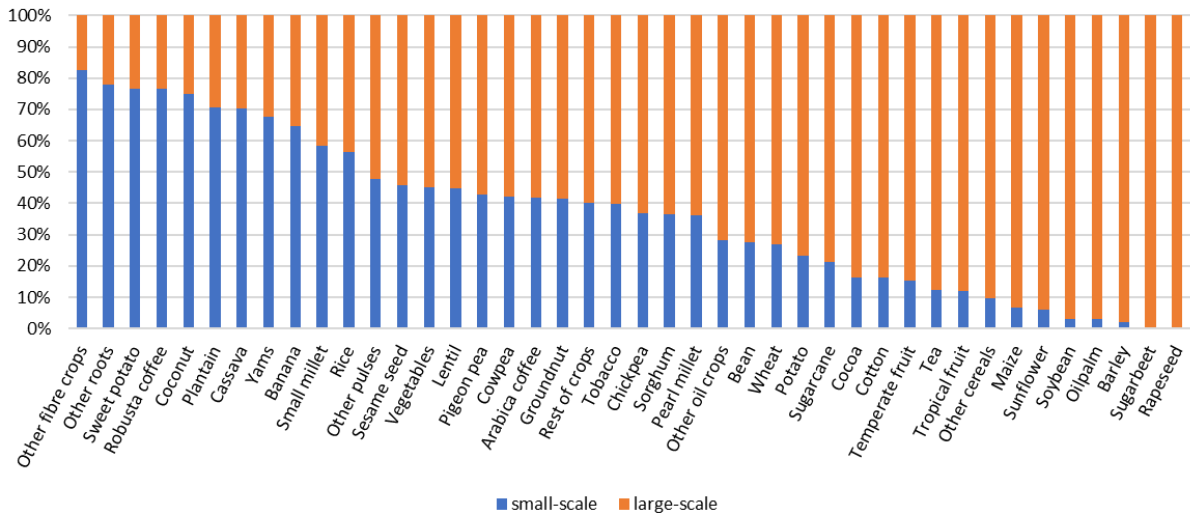


Fig. 10 Small-scale farms' contribution to total water consumption per crop.

Small-scale farms' contribution to total water consumption also depends on whether the crop is export-oriented in the country. We investigated the relationship between small-scale farms' contribution to total water consumption and export orientation of crops. We quantified the export orientation of a crop by dividing total export by the total production of the crop in FAOSTAT. Since the total export includes not only domestic production but also imported crops, we excluded the samples whose export divided by production or import divided by production is larger than 1.0. Due to data availability, we were able to investigate this relationship for 15 crops.

Of 15 crops, 7 crops are found to have a significant negative relationship between the small-scale farms' contribution to total water consumption and the export orientation, which means small-scale farms consume less water in export-oriented agriculture and more water in domestic market-oriented agriculture (Fig. 11). This also indicates the water embodies in international crop trade is mainly contributed by large-scale farms. The 7 crops are barley, maize, potato, rice, soybean, vegetable, wheat. Of 8 remain crops, bean has a significant positive relationship due to Ethiopia and Peru. No significant relationship is found for banana, cassava, coconut, groundnut, sorghum, sweet potato, and yams.

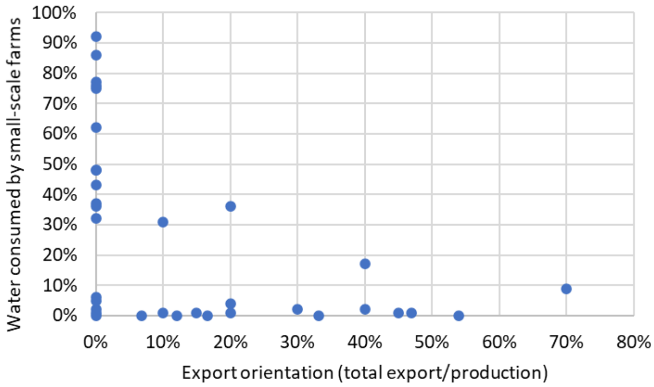


Fig. 11 Relationship between the small-scale farms' contribution to total water consumption and the export orientation for vegetables.

Limitations and next steps

As a global map, our farm-size specific crop map is limited to 56 countries while other estimated farm size distribution map have a larger coverage, e.g. Mehrabi et al. (2020). Since Ricciardi et al. (2018a), more data is available for some countries that directly measure the crop area per farm size. One example is the third agricultural census from China. We are considering adding these countries, e.g. China, in our future work.

Further validations will help us to better understand the reliability of our downscaled map, especially for crop area per farm size. We validated our map with the data from the satellite image, expert knowledge, household survey, and similar maps. The validations show an overall acceptable error. Due to data availability, we are unable to validation our map with some observations that provide the farm-size specific crop area at a large scale. If some data is available in the future, we could gain more understanding of the reliability of our downscaled map.

We need to update water consumption data in order to estimate crop production in our future work. Currently, we use GCWM to estimate water consumption. GCWM provides water consumption and yield under the 2000 climate conditions. It does not matter if we only focus on water consumption, however, it would be unreasonable to use the yield from GCWM to estimate crop production for the year 2010. To keep consistencies between the harvested area in our downscaled map, water consumption, and crop production in future work, we need to use the crop model that estimates the water consumption and yield under 2010 climate conditions. These data are expected available in the coming months from the global crop model ACEA (Mialyk et al., 2021).

For the small-scale farms' contribution to water consumption, we need to include more dimensions on the source of water (blue or green, surface water or groundwater). Tracking the source of water helps to assess water use sustainability and find solutions for water scarcity. It will be achievable to track the source of water by incorporating agricultural water consumption into the global hydrological model, e.g. CWatM (Burek et al., 2020).

Our downscaled map will make it possible to investigate the water sustainability assessment and food security from various perspectives. One of the important issues is to assess the water productivity per farm size, including nutritional water productivity which helps to inform food security. Current studies indicate small-scale farms have higher land productivity because of relative land scarcity (Ren et al., 2019). Little is known for water productivity per farm size. The changes in farm size always happen along with urbanization in rapidly developing regions. Estimating water productivity and water consumption per farm size helps to understand how water interacts with land and food security during social development. This topic is doable using the global nutrition database after updating water consumption data.

Another issue is to assess the sustainability of water as both a local resource and a global resource. It is possible with our downscaled map since farm size indicates the relationship between agricultural production and domestic and export-oriented market. We could identify whether domestic consumption or global consumption drives local water scarcity in a certain country by using the Environmentally Extended Input-Output Table, such as Bruckner et al. (2019). For water allocation, based on our downscaled map, we could estimate the equity among agricultural water use in response to some social concerns. Based on these, hopefully, we could find pathways towards sustainable, productive, and equitable water use for the global food supply.

Conclusions

In this study, we developed a farm-size specific crop map of harvested area at 5-arcmin spatial resolution for 42 crops and 56 countries. Validations show an overall acceptable error for the geographical farm size distribution of oil palm in 5 countries, coffee in Costa Rica, and irrigation across 11 countries. Part of the errors can be attributed to the data source used to develop the map.

We find small-scale farms, under the 2 ha threshold, contribute more to water consumption in labor-intensive crops, e.g. sweet potato, banana, rice, coconut. Our results also indicate most of the water consumed by small-scale farms is not virtually exported. We find an overall 25.8% contribution of small-scale farms to total water consumption. This number is significantly higher in developing countries, e.g. India, Africa, and South America.

Further work will focus on further validation on the downscaled map, increasing the map coverage, the estimation of crop production, the assessment of water productivity including nutritional water productivity, and the pathways towards sustainable, efficient, and equitable water use for global food supply.

References

Bakhtary H, Haupt F, Luttrell C, Landholm D, Jelsma I (2021) Promoting sustainable oil palm production by independent smallholders in Indonesia.

Bosc P-M, Berdegué J, Goïta M, van der Ploeg JD, Sekine K, Zhang L (2013) Investing in smallholder agriculture for food security. HLPE

Bruckner M, Wood R, Moran D, Kuschnig N, Wieland H, Maus V, Borner J (2019) FABIO-The Construction of the Food and Agriculture Biomass Input-Output Model. Environ Sci Technol 53:11302-11312. doi:10.1021/acs.est.9b03554

Burek P, Satoh Y, Kahil T, Tang T, Greve P, Smilovic M, Guillaumot L, Zhao F, Wada Y (2020) Development of the Community Water Model (CWatM v1.04) – a high-resolution hydrological model for global and regional assessment of integrated water resources management. Geoscientific Model Development 13:3267-3298. doi:10.5194/gmd-13-3267-2020

Descals A, Wich S, Meijaard E, Gaveau D, Peedell S, Szantoi Z (2020) High-resolution global map of smallholder and industrial closed-canopy oil palm plantations. Earth System Science Data.

FAO (2014) The state of food and agriculture. Innovation and family farming. FAO, Rome

FAO (2015) The state of food insecurity in the world (SOFI) 2015: Meeting the 2015 international hunger targets: Taking stock of uneven progress. FAO, Rome

Fritz S, See L, McCallum I, You L, Bun A, Moltchanova E, Duerauer M, Albrecht F, Schill C, Perger C, Havlik P, Mosnier A, Thornton P, Wood-Sichra U, Herrero M, Becker-Reshef I, Justice C, Hansen M, Gong P, Abdel Aziz S, Cipriani A, Cumani R, Cecchi G, Conchedda G, Ferreira S, Gomez A, Haffani M, Kayitakire F, Malanding J, Mueller R, Newby T, Nonguierma A, Olusegun A, Ortner S, Rajak DR, Rocha J, Schepaschenko D, Schepaschenko M, Terekhov A, Tiangwa A, Vancutsem C, Vintrou E, Wenbin W, van der Velde M, Dunwoody A, Kraxner F, Obersteiner M (2015) Mapping global cropland and field size. Glob Chang Biol 21:1980-1992. doi:10.1111/gcb.12838

Gomez y Paloma S, Riesgo L, Louhichi K (2020) The Role of Smallholder Farms in Food and Nutrition Security. Springer Nature

Guarín A, Rivera M, Pinto-Correia T, Guiomar N, Šūmane S, Moreno-Pérez OM (2020) A new typology of small farms in Europe. Global Food Security 26. doi:10.1016/j.gfs.2020.100389

Gurobi Optimization L (2021) Gurobi Optimizer Reference Manual.

Herrero M, Thornton PK, Power B, Bogard JR, Remans R, Fritz S, Gerber JS, Nelson G, See L, Waha K, Watson RA, West PC, Samberg LH, van de Steeg J, Stephenson E, van Wijk M, Havlík P (2017) Farming and the geography of nutrient production for human use: a transdisciplinary analysis. The Lancet Planetary Health 1:e33-e42. doi:10.1016/s2542-5196(17)30007-4

Hoekstra AY (2020) The water footprint of modern consumer society (second edition). Routledge, London

Holland MB, Shamer SZ, Imbach P, Zamora JC, Medellin Moreno C, Hidalgo EJL, Donatti CI, Martínez-Rodríguez MR, Harvey CA (2016) Mapping adaptive capacity and smallholder agriculture: applying expert knowledge at the landscape scale. Climatic Change 141:139-153. doi:10.1007/s10584-016-1810-2

ICO (2019) Country coffee profile Vietnam. International Coffee Organization (ICO), London

Khalil CA, Conforti P, Ergin I, Gennari P (2017) Defining small scale food producers to monitor target 2.3 of the 2030 Agenda for Sustainable Development. FAO, Rome

Lesiv M, Laso Bayas JC, See L, Duerauer M, Dahlia D, Durando N, Hazarika R, Kumar Sahariah P, Vakolyuk M, Blyshchyk V, Bilous A, Perez-Hoyos A, Gengler S, Prestele R, Bilous S, Akhtar IUH, Singha K, Choudhury SB, Chetri T, Malek Z, Bungnamei K, Saikia A, Sahariah D, Narzary W, Danylo O, Sturn T, Karner M, McCallum I, Schepaschenko D, Moltchanova E, Fraisl D, Moorthy I, Fritz S (2019) Estimating the global distribution of field size using crowdsourcing. Glob Chang Biol 25:174-186. doi:10.1111/gcb.14492

Lowder SK, Skoet J, Raney T (2016) The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide. World Development 87:16-29. doi:10.1016/j.worlddev.2015.10.041

Lu M, Wu W, You L, See L, Fritz S, Yu Q, Wei Y, Chen D, Yang P, Xue B (2020) A cultivated planet in 2010 – Part 1: The global synergy cropland map. Earth System Science Data 12:1913-1928. doi:10.5194/essd-12-1913-2020

Meemken EM, Bellemare MF (2020) Smallholder farmers and contract farming in developing countries. Proc Natl Acad Sci U S A 117:259-264. doi:10.1073/pnas.1909501116

Mehrabi Z, McDowell MJ, Ricciardi V, Levers C, Martinez JD, Mehrabi N, Wittman H, Ramankutty N, Jarvis A (2020) The global divide in data-driven farming. Nature Sustainability 4:154-160. doi:10.1038/s41893-020-00631-0

Mialyk O, Schyns JF, Booij MJ, Hogeboom RJ (2021) Historical simulation of maize water footprints with a new global gridded crop model ACEA. Hydrology and Earth System Sciences Discussions:1-25.

Monfreda C, Ramankutty N, Foley JA (2008) Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Global Biogeochemical Cycles 22:n/a-n/a. doi:10.1029/2007gb002947

Ray DK, Mueller ND, West PC, Foley JA (2013) Yield trends are insufficient to double global crop production by 2050. PloS one 8:e66428.
Ren C, Liu S, van Grinsven H, Reis S, Jin S, Liu H, Gu B (2019) The impact of farm size on agricultural sustainability. Journal of Cleaner Production 220:357-367. doi:10.1016/j.jclepro.2019.02.151
Ricciardi V, Ramankutty N, Mehrabi Z, Jarvis L, Chookolingo B (2018a) How much of the world's food do smallholders produce? Global Food Security 17:64-72.
Ricciardi V, Ramankutty N, Mehrabi Z, Jarvis L, Chookolingo B (2018b) An open-access dataset of crop production by farm size from agricultural censuses and surveys. Data Brief 19:1970-1988. doi:10.1016/j.dib.2018.06.057
Ricciardi V, Wane A, Sidhu BS, Godde C, Solomon D, McCullough E, Diekmann F, Porciello J, Jain M, Randall N (2020) A scoping review of research funding for small-scale farmers in water scarce regions. Nature Sustainability 3:836-844.
Rivera M, Guarín A, Pinto-Correia T, Almaas H, Mur LA, Burns V, Czekaj M, Ellis R, Galli F, Grivins M, Hernández P, Karanikolas P, Prosperi P, Sánchez Zamora P (2020) Assessing the role of small farms in regional food systems in Europe: Evidence from a comparative study. Global Food Security 26. doi:10.1016/j.gfs.2020.100417
Samberg LH, Gerber JS, Ramankutty N, Herrero M, West PC (2016) Subnational distribution of average farm size and smallholder contributions to global food production. Environmental Research Letters 11. doi:10.1088/1748-9326/11/12/124010
Sheahan M, Barrett CB (2017) Ten striking facts about agricultural input use in Sub-Saharan Africa. Food Policy 67:12-25. doi:10.1016/j.foodpol.2016.09.010
Siebert S, Döll P (2010) Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation. Journal of Hydrology 384:198-217. doi:10.1016/j.jhydrol.2009.07.031
Wu Y, Xi X, Tang X, Luo D, Gu B, Lam SK, Vitousek PM, Chen D (2018) Policy distortions, farm size, and the overuse of agricultural chemicals in China. P Natl Acad Sci USA 115:7010-7015. doi:10.1073/pnas.1806645115
Yu Q, You L, Wood-Sichra U, Ru Y, Joglekar AKB, Fritz S, Xiong W, Lu M, Wu W, Yang P (2020) A cultivated planet in 2010 – Part 2: The global gridded agricultural-production maps. Earth System Science Data 12:3545-3572. doi:10.5194/essd-12-3545-2020

Appendix

A1 List of 11 household surveys to estimate the percentage of irrigated area per farm size

Country	Year	Survey
Albania	2005	Living Standard Measurement Survey
Cambodia	2009	Cambodia Social-Economic Survey
Ethiopia	2014	Ethiopia Socialeconomic Survey
India	2012	India Human Development survey
Malawi	2011	Integrated Household Survey
Mali	2014	Enquête Agricole de conjoncture integree aux Conditions de Vie des Menages

Niger	2011	National Survey on Household Living Conditions and Agriculture
Nigeria	2013	General Household Survey
Timor-Leste	2008	Survey of Living Standards
United Republic of Tanzania	2009	National Panel Survey
Uganda	2010	The Uganda National Panel Survey

A2 Percentage of irrigated area per farm size in the 11 countries based on household survey and downscaled map

Country		0 - 1	1 - 2	2 - 5	5 - 10	10 - 20	20 - 50	50 - 100	100 - 200	200 - 500	500 - 1000	> 1000
Albania	Survey	55.2%	36.4%	25.5%								
	Downscaled map	25.5%	30.2%	27.3%	32.0%	39.3%	3.9%	55.8%				
Cambodia	Survey	49.1%	43.6%	42.7%	29.0%							
	Downscaled map	22.8%	31.3%	16.6%	21.0%	16.4%	16.7%	21.7%	4.1%	6.9%		
Ethiopia	Survey	5.1%	3.1%	2.4%	2.9%							
	Downscaled map	1.4%	1.2%	1.1%	0.8%	3.3%	3.8%	6.6%	6.7%	3.2%	6.7%	12.6%
India	Survey	62.5%	64.5%	70.7%	76.8%	74.1%	55.9%					
	Downscaled map	41.8%	32.7%	32.1%	29.8%	30.1%	27.6%	26.9%	40.8%	35.2%	32.6%	30.5%
Malawi	Survey	1.8%	2.6%	2.6%								
	Downscaled map	0.5%	1.0%	0.1%			0.0%					
Mali	Survey	39.0%	26.6%	12.3%	7.3%	4.7%	1.7%					
	Downscaled map	29.9%	29.8%	18.9%	8.6%	15.0%	5.4%	8.9%	4.1%	2.5%	9.1%	20.6%
Niger	Survey	19.4%	10.5%	3.8%	2.0%	1.6%	0.1%					
	Downscaled map	0.0%	0.1%	0.0%			0.0%					

Nigeria	Survey	1.8%	1.8%	0.6%	0.0%							
	Downscaled map	2.0%	0.8%	0.6%	0.4%	2.6%	0.8%	0.3%	0.5%	0.5%		0.0%
Timor-Leste	Survey	12.5%	8.3%	7.1%								
	Downscaled map	23.4%	5.4%	15.5%	2.1%	15.1%		64.7%				
United Republic of Tanzania	Survey	4.0%	2.9%	3.0%	2.8%	1.0%						
	Downscaled map	1.6%	1.3%	1.1%	0.8%	1.3%	1.0%	1.5%	0.9%	1.4%	8.0%	1.3%
Uganda	Survey	0.6%	1.0%	0.9%	0.0%							
	Downscaled map	1.0%	0.7%	4.7%	0.7%	0.0%						

A3 Number and percentage of pixels that have similar, larger, and smaller dominant farm size in our downscaled map and Mehrabi's map

Country	Number			Percentage		
	similar	larger	smaller	similar	large	smaller
Albania	340	42	79	73.8%	9.1%	17.1%
Austria	367	240	162	47.7%	31.2%	21.1%
Belgium	401	99	85	68.5%	16.9%	14.5%
Burkina Faso	621	2033	59	22.9%	74.9%	2.2%
Bulgaria	1283	355	42	76.4%	21.1%	2.5%
Bosnia and Herzegovina	330	382	22	45.0%	52.0%	3.0%
Brazil	17976	4273	20216	42.3%	10.1%	47.6%
Colombia	962	404	2990	22.1%	9.3%	68.6%
Costa Rica	68	338	12	16.3%	80.9%	2.9%
Cyprus	17	86	0	16.5%	83.5%	0.0%
Czech Republic	876	290	63	71.3%	23.6%	5.1%
Germany	2619	2021	1445	43.0%	33.2%	23.7%
Denmark	692	130	18	82.4%	15.5%	2.1%
Spain	923	5768	91	13.6%	85.0%	1.3%
Estonia	469	354	12	56.2%	42.4%	1.4%
Ethiopia	4076	439	828	76.3%	8.2%	15.5%
Finland	123	1558	5	7.3%	92.4%	0.3%
France	6438	926	326	83.7%	12.0%	4.2%
United Kingdom	2431	178	140	88.4%	6.5%	5.1%
Ghana	473	1794	0	20.9%	79.1%	0.0%
Greece	820	532	296	49.8%	32.3%	18.0%
Croatia	255	472	50	32.8%	60.7%	6.4%

Hungary	718	532	23	56.4%	41.8%	1.8%
India	23084	8853	4194	63.9%	24.5%	11.6%
Ireland	713	222	106	68.5%	21.3%	10.2%
Italy	1554	1188	1093	40.5%	31.0%	28.5%
Cambodia	328	983	6	24.9%	74.6%	0.5%
Lithuania	381	737	178	29.4%	56.9%	13.7%
Luxembourg	25	1	0	96.2%	3.8%	0.0%
Latvia	523	490	157	44.7%	41.9%	13.4%
Mexico	7306	2878	2586	57.2%	22.5%	20.3%
Mali	247	3713	0	6.2%	93.8%	0.0%
Mongolia	27	17	20	42.2%	26.6%	31.3%
Malawi	703	111	0	86.4%	13.6%	0.0%
Niger	1808	1479	53	54.1%	44.3%	1.6%
Nigeria	6270	3589	1	63.6%	36.4%	0.0%
Netherlands	339	158	74	59.4%	27.7%	13.0%
Norway	88	236	15	26.0%	69.6%	4.4%
Panama	175	112	164	38.8%	24.8%	36.4%
Peru	631	297	2334	19.3%	9.1%	71.6%
Poland	2671	1754	1249	47.1%	30.9%	22.0%
Portugal	628	325	168	56.0%	29.0%	15.0%
Paraguay	871	91	760	50.6%	5.3%	44.1%
Romania	1613	1242	474	48.5%	37.3%	14.2%
Russia	27122	11516	8585	57.4%	24.4%	18.2%
Slovakia	489	141	2	77.4%	22.3%	0.3%
Slovenia	20	142	0	12.3%	87.7%	0.0%
Switzerland	633	1059	25	36.9%	61.7%	1.5%

Tajikistan	173	595	20	22.0%	75.5%	2.5%
Timor-Leste	125	7	0	94.7%	5.3%	0.0%
Tanzania	1646	3690	6	30.8%	69.1%	0.1%
Uganda	983	127	839	50.4%	6.5%	43.0%
Uruguay	1077	247	275	67.4%	15.4%	17.2%
United States of America	42431	15793	7447	64.6%	24.0%	11.3%
South Africa	640	59	4177	13.1%	1.2%	85.7%
Total	168602	85098	61972	53.4%	27.0%	19.6%

A4 Proportional farm size distribution per country from Lowder's dataset and the downscaled map in this study

Country		0_1	1_2	2_5	5_10	10_20	20_50	50_100	100_200	200_500	500_1000	1000_5000
Albania	Lowder	6.8%	10.5%	82.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	This study	16.1%	27.9%	38.3%	9.4%	6.4%	1.1%	0.8%	0.0%	0.0%	0.0%	0.0%
Austria	Lowder	0.0%	2.2%	5.3%	9.7%	17.8%	24.0%	9.6%	31.5%	0.0%	0.0%	0.0%
	This study	0.0%	0.5%	1.4%	3.5%	10.2%	34.8%	30.6%	18.7%	0.0%	0.0%	0.0%
Belgium	Lowder	0.0%	0.9%	2.2%	4.4%	10.8%	38.6%	29.8%	13.4%	0.0%	0.0%	0.0%
	This study	0.0%	0.0%	0.7%	2.6%	8.2%	30.3%	32.4%	25.8%	0.0%	0.0%	0.0%
Brazil	Lowder	0.1%	0.2%	0.7%	1.3%	2.8%	7.2%	7.8%	9.3%	14.3%	11.4%	45.1%
	This study	1.4%	1.8%	4.2%	4.1%	7.1%	10.5%	7.2%	7.3%	10.7%	9.5%	36.3%
Bulgaria	Lowder	6.6%	0.0%	8.3%	0.0%	0.0%	6.6%	78.5%	0.0%	0.0%	0.0%	0.0%
	This study	2.4%	0.0%	2.2%	1.7%	2.2%	4.1%	4.4%	83.0%	0.0%	0.0%	0.0%
Burkina Faso	Lowder	1.8%	7.4%	34.6%	36.7%	19.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	This study	0.3%	1.6%	7.3%	10.1%	11.0%	17.7%	10.5%	10.4%	11.5%	3.2%	16.4%
Colombia	Lowder	0.4%	0.8%	2.7%	4.0%	6.2%	13.6%	14.8%	14.9%	22.9%	19.8%	0.0%
	This study	6.9%	17.8%	1.7%	14.0%	15.3%	16.6%	8.8%	6.1%	5.4%	3.1%	4.1%

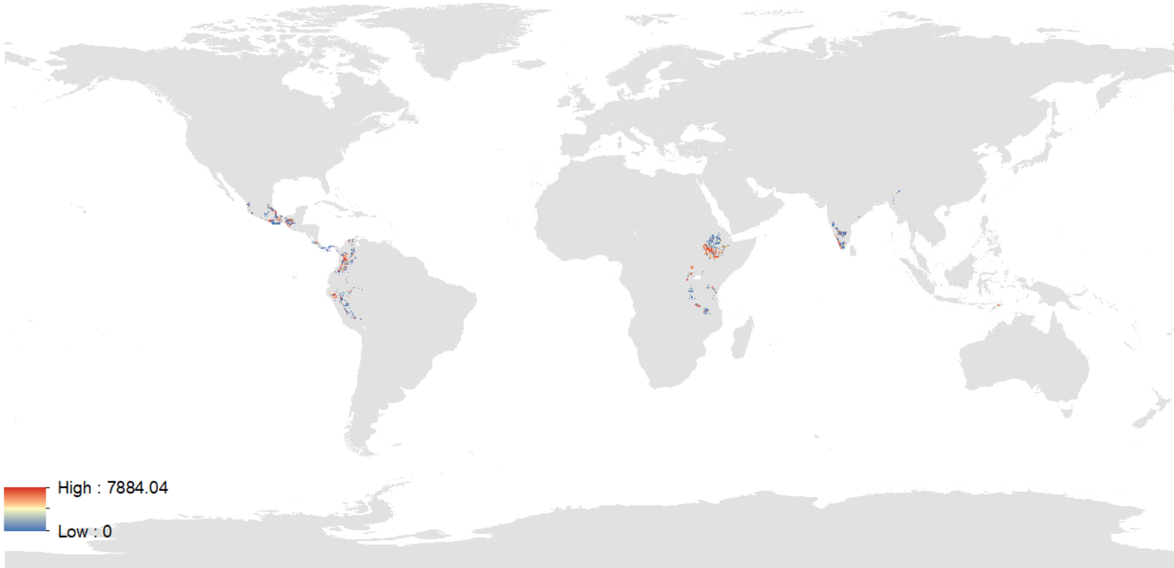
Croatia	Lowder	5.9%	7.4%	19.9%	21.1%	15.1%	30.5%	0.0%	0.0%	0.0%	0.0%	0.0%
	This study	2.7%	2.8%	12.4%	12.4%	12.0%	16.8%	12.7%	28.2%	0.0%	0.0%	0.0%
Cyprus	Lowder	6.4%	6.8%	14.4%	12.6%	13.6%	16.4%	11.5%	18.4%	0.0%	0.0%	0.0%
	This study	11.7%	11.8%	16.7%	10.2%	9.3%	14.1%	9.8%	16.5%	0.0%	0.0%	0.0%
Czech Republic	Lowder	0.1%	0.3%	0.8%	1.2%	2.0%	3.7%	3.5%	4.2%	8.0%	15.2%	60.8%
	This study	0.2%	0.1%	0.7%	0.7%	1.4%	3.6%	4.3%	4.4%	8.0%	15.2%	61.5%
Denmark	Lowder	0.0%	0.1%	0.2%	2.8%	6.4%	20.8%	29.7%	40.0%	0.0%	0.0%	0.0%
	This study	0.0%	0.0%	0.0%	0.9%	2.7%	9.8%	14.9%	71.7%	0.0%	0.0%	0.0%
Estonia	Lowder	0.5%	2.2%	6.0%	8.7%	12.3%	14.4%	7.6%	48.2%	0.0%	0.0%	0.0%
	This study	0.1%	0.1%	0.5%	1.0%	1.9%	4.5%	6.1%	85.9%	0.0%	0.0%	0.0%
Ethiopia	Lowder	27.1%	33.3%	32.6%	5.5%	1.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	This study	35.5%	35.9%	25.4%	2.5%	0.1%	0.1%	0.0%	0.0%	0.4%	0.0%	0.0%
Finland	Lowder	0.0%	1.1%	2.9%	7.1%	18.6%	42.6%	20.4%	7.3%	0.0%	0.0%	0.0%
	This study	0.0%	0.0%	0.1%	1.0%	4.9%	24.7%	33.5%	35.3%	0.1%	0.1%	0.1%
France	Lowder	0.0%	0.7%	1.3%	1.9%	4.2%	17.3%	30.6%	44.0%	0.0%	0.0%	0.0%
	This study	0.0%	0.2%	0.4%	0.9%	2.2%	8.5%	21.2%	66.6%	0.0%	0.0%	0.0%
Germany	Lowder	0.0%	0.3%	2.1%	3.7%	8.3%	22.4%	21.4%	12.2%	8.1%	6.2%	15.4%

	This study	0.3%	0.2%	0.8%	1.2%	3.8%	11.2%	18.3%	19.0%	12.5%	9.3%	23.4%
Greece	Lowder	0.0%	11.4%	20.6%	20.4%	19.1%	18.3%	5.8%	4.4%	0.0%	0.0%	0.0%
	This study	0.0%	11.2%	20.1%	20.3%	19.3%	20.2%	6.4%	2.6%	0.0%	0.0%	0.0%
India	Lowder	18.7%	20.2%	31.2%	16.7%	8.3%	4.9%	0.0%	0.0%	0.0%	0.0%	0.0%
	This study	23.7%	22.5%	30.6%	14.0%	5.9%	2.9%	0.3%	0.0%	0.0%	0.0%	0.0%
Ireland	Lowder	0.0%	0.1%	0.8%	3.0%	11.5%	39.6%	29.0%	16.0%	0.0%	0.0%	0.0%
	This study	0.4%	0.1%	0.1%	0.7%	3.1%	19.5%	32.5%	43.5%	0.1%	0.0%	0.0%
Italy	Lowder	2.4%	3.6%	8.5%	9.3%	11.2%	16.1%	10.9%	37.9%	0.0%	0.0%	0.0%
	This study	2.2%	3.2%	11.9%	13.3%	15.4%	21.6%	14.9%	17.5%	0.0%	0.0%	0.0%
Latvia	Lowder	0.0%	0.4%	3.3%	8.0%	17.5%	31.0%	18.0%	9.4%	5.7%	6.7%	0.0%
	This study	0.0%	0.4%	1.3%	2.5%	4.8%	8.2%	8.6%	73.9%	0.1%	0.1%	0.0%
Lithuania	Lowder	0.0%	1.1%	13.9%	15.1%	17.7%	17.2%	8.3%	6.2%	6.2%	14.3%	0.0%
	This study	0.0%	0.9%	4.9%	5.1%	6.3%	10.1%	12.5%	14.1%	13.8%	32.3%	0.0%
Luxembourg	Lowder	0.0%	0.3%	0.9%	1.6%	2.7%	15.3%	50.7%	28.6%	0.0%	0.0%	0.0%
	This study	0.8%	0.2%	0.5%	1.6%	2.2%	8.4%	29.7%	56.4%	0.2%	0.0%	0.0%
Malta	Lowder	33.1%	24.9%	29.0%	9.6%	3.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	This study	20.4%	15.2%	38.8%	18.6%	7.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

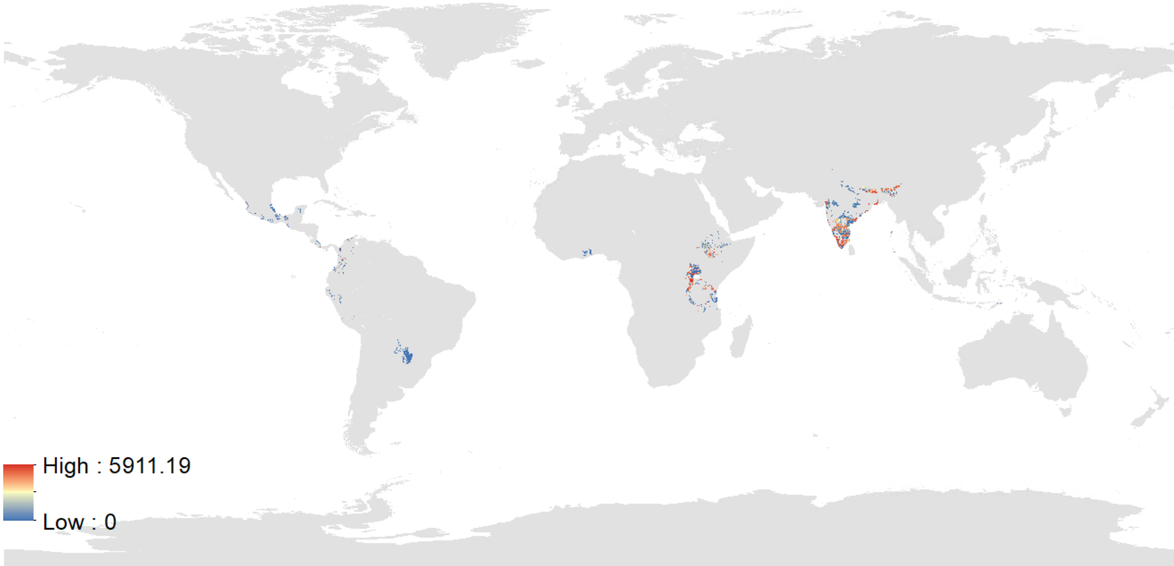
Netherlands	Lowder	0.0%	1.0%	2.7%	5.5%	12.2%	42.6%	21.9%	14.1%	0.0%	0.0%	0.0%
	This study	0.1%	0.1%	0.8%	2.5%	6.7%	24.3%	33.0%	32.5%	0.0%	0.0%	0.0%
Norway	Lowder	0.0%	0.3%	3.6%	11.7%	31.0%	42.8%	8.9%	1.4%	0.2%	0.0%	0.0%
	This study	0.3%	0.1%	0.6%	3.7%	13.4%	35.7%	27.8%	16.3%	1.8%	0.1%	0.1%
Panama	Lowder	0.6%	1.0%	2.7%	3.7%	7.2%	17.5%	17.8%	15.2%	15.0%	6.5%	12.7%
	This study	6.4%	1.9%	1.5%	3.0%	4.2%	16.9%	18.5%	15.1%	19.9%	12.5%	0.0%
Paraguay	Lowder	0.0%	0.1%	0.8%	1.8%	3.4%	3.6%	2.1%	2.4%	4.4%	4.2%	77.1%
	This study	0.2%	0.8%	2.5%	4.6%	5.8%	5.2%	4.9%	7.9%	15.4%	12.8%	40.0%
Peru	Lowder	0.0%	0.0%	5.4%	5.0%	4.4%	7.5%	77.6%	0.0%	0.0%	0.0%	0.0%
	This study	7.2%	36.5%	1.4%	18.9%	14.0%	12.0%	4.5%	1.7%	1.4%	0.7%	1.7%
Poland	Lowder	2.8%	4.6%	12.6%	18.1%	21.4%	15.5%	4.9%	2.7%	4.4%	4.2%	8.8%
	This study	1.1%	1.8%	9.9%	16.1%	20.7%	20.0%	9.5%	2.9%	4.7%	4.6%	8.8%
Portugal	Lowder	2.8%	6.4%	10.5%	8.6%	9.7%	9.8%	7.0%	45.3%	0.0%	0.0%	0.0%
	This study	3.1%	6.9%	14.9%	12.6%	11.7%	14.0%	9.4%	27.4%	0.0%	0.0%	0.0%
Romania	Lowder	4.9%	8.1%	20.2%	11.1%	3.9%	2.0%	1.7%	48.1%	0.0%	0.0%	0.0%
	This study	4.9%	8.1%	15.2%	7.8%	3.9%	4.1%	3.7%	52.1%	0.1%	0.0%	0.1%
Switzerland	Lowder	0.7%	1.0%	3.2%	9.3%	36.2%	42.7%	5.6%	1.4%	0.0%	0.0%	0.0%

	This study	0.1%	0.1%	0.1%	0.4%	1.9%	8.6%	16.7%	72.0%	0.0%	0.0%	0.0%
Uganda	Lowder	11.0%	15.8%	24.8%	18.2%	30.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	This study	54.8%	26.4%	15.9%	1.6%	0.6%	0.6%	0.0%	0.0%	0.0%	0.0%	0.0%
United Kingdom	Lowder	0.0%	0.3%	0.5%	1.3%	2.9%	10.1%	16.4%	68.6%	0.0%	0.0%	0.0%
	This study	0.0%	0.0%	0.0%	0.1%	0.5%	3.9%	11.4%	84.0%	0.0%	0.0%	0.0%
United States of America	Lowder	0.0%	0.0%	0.1%	0.4%	1.1%	3.9%	6.4%	9.0%	16.4%	62.6%	0.0%
	This study	0.2%	0.1%	0.1%	0.1%	0.2%	1.6%	3.8%	8.1%	22.1%	63.6%	0.1%
Uruguay	Lowder	0.0%	0.0%	0.1%	0.3%	0.6%	1.7%	2.9%	5.5%	13.2%	16.6%	59.1%
	This study	0.6%	0.0%	0.3%	0.0%	0.1%	0.3%	0.9%	2.7%	9.5%	13.0%	72.6%

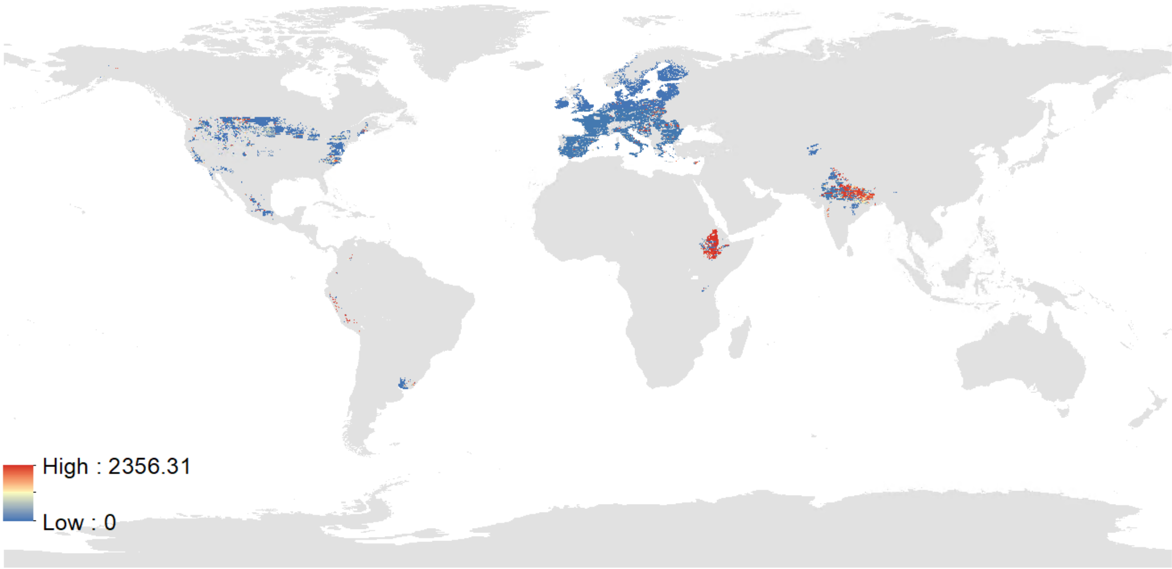
A5 The total harvested area (in hectares) of small-scale farms for 42 crops across 56 countries



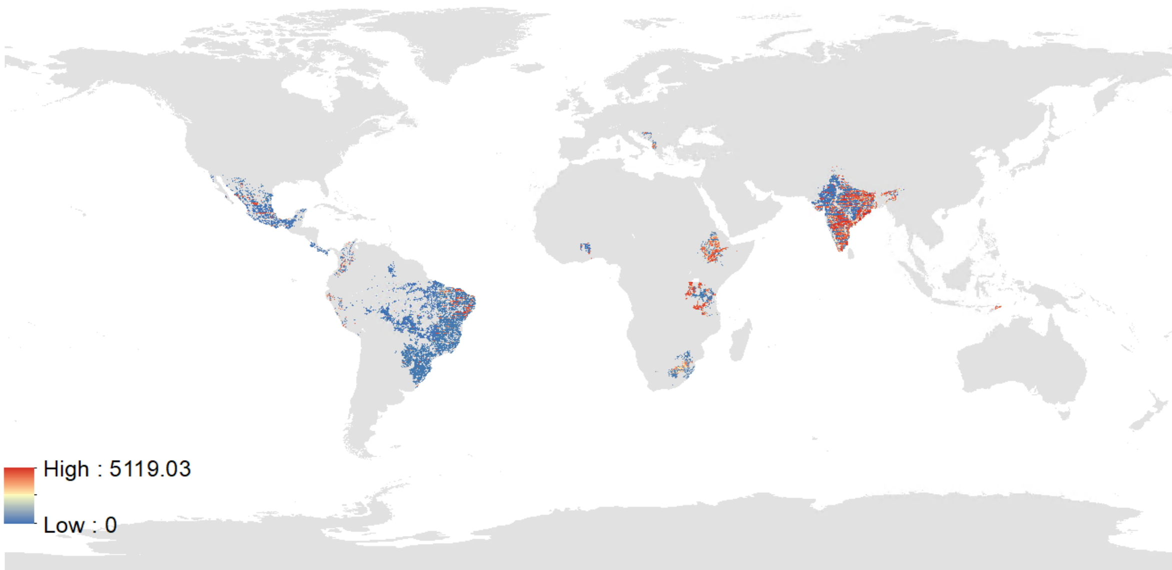
Arabica coffee



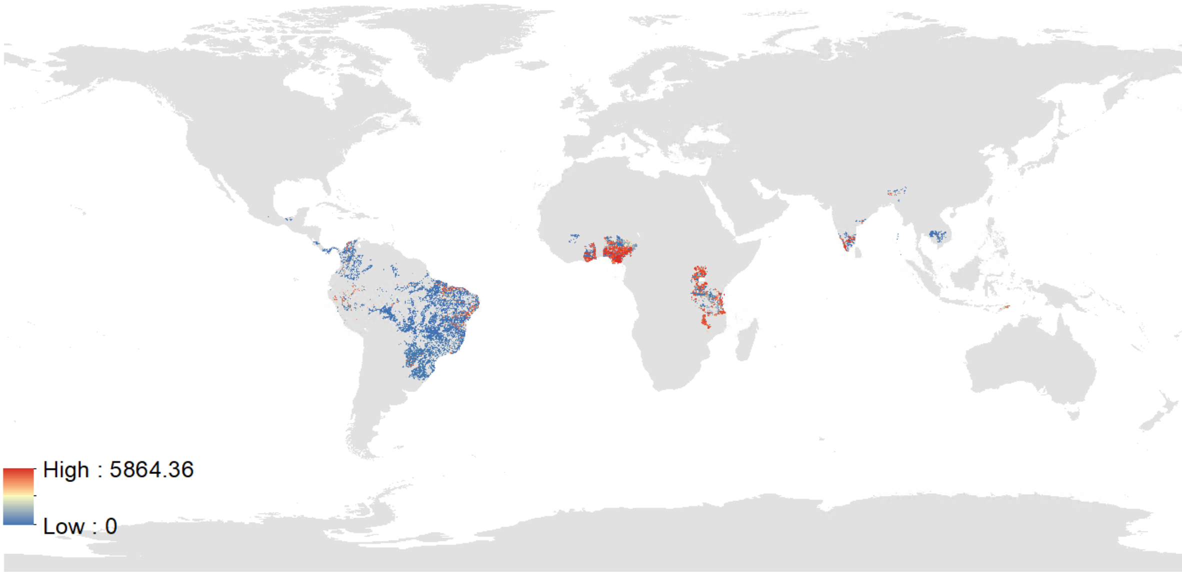
Banana



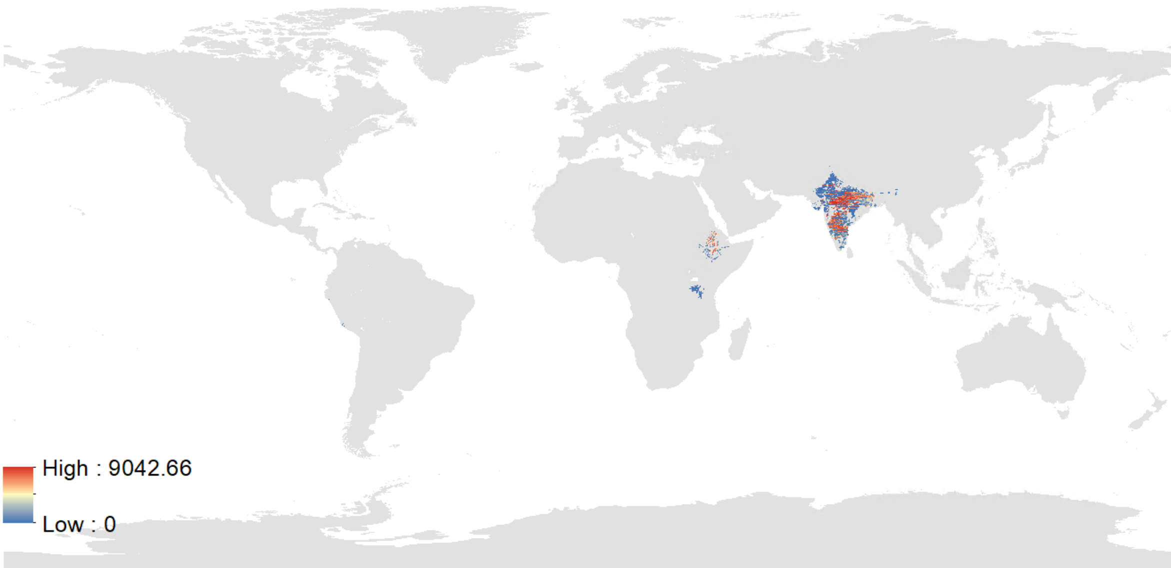
Barley



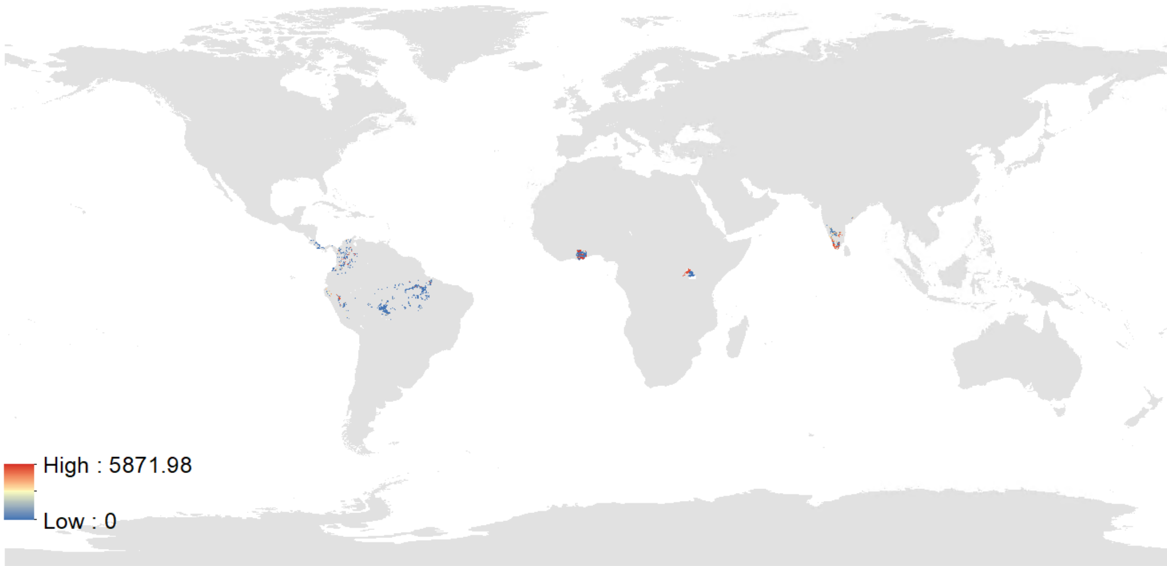
Bean



Cassava

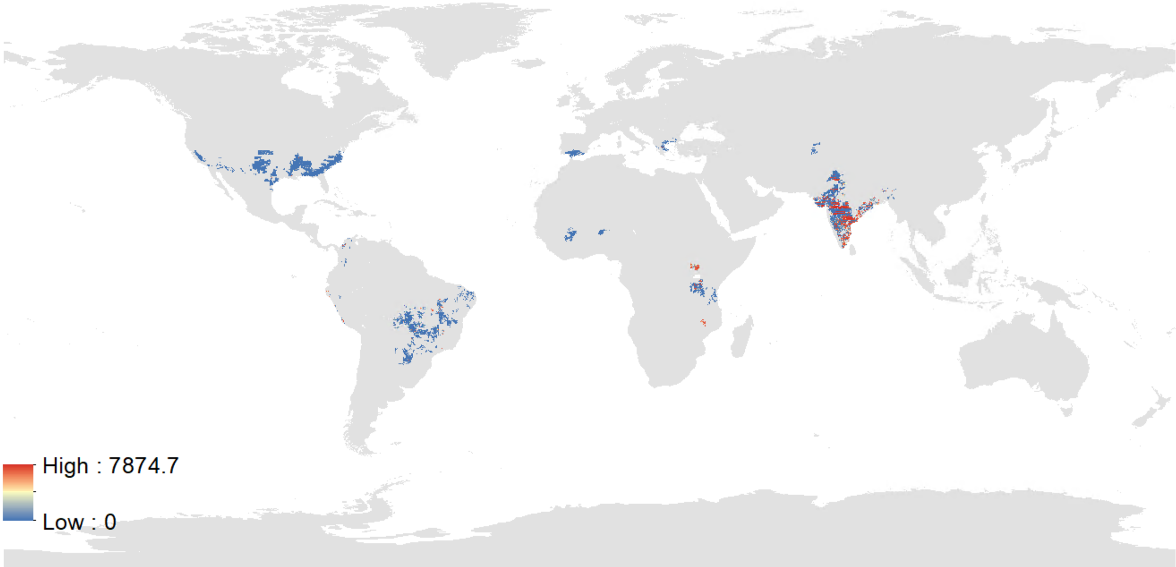


Chickpea

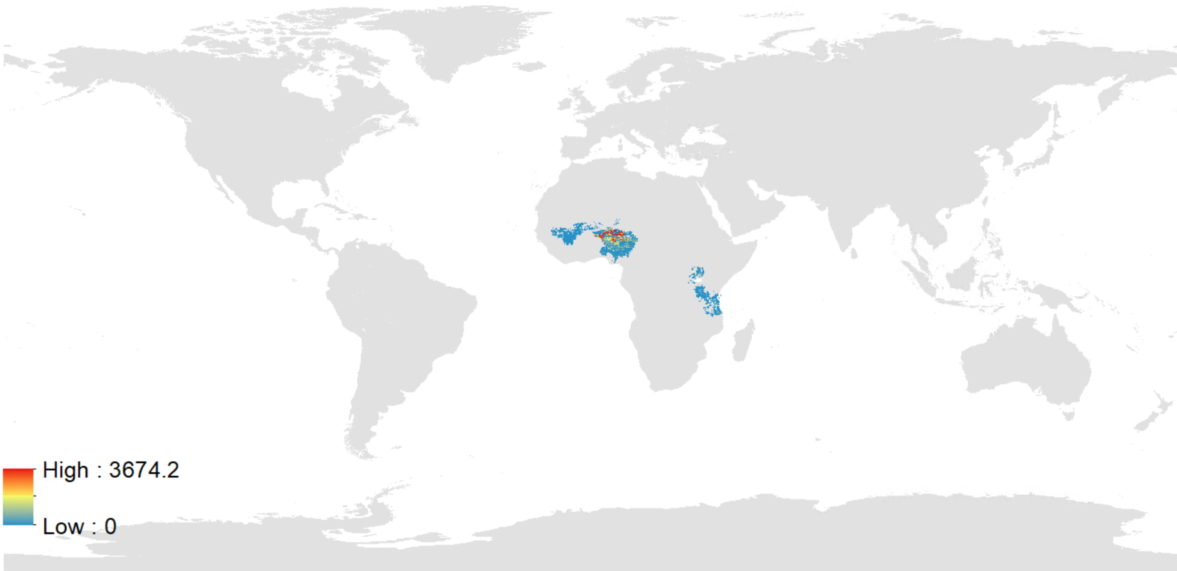


Cocoa

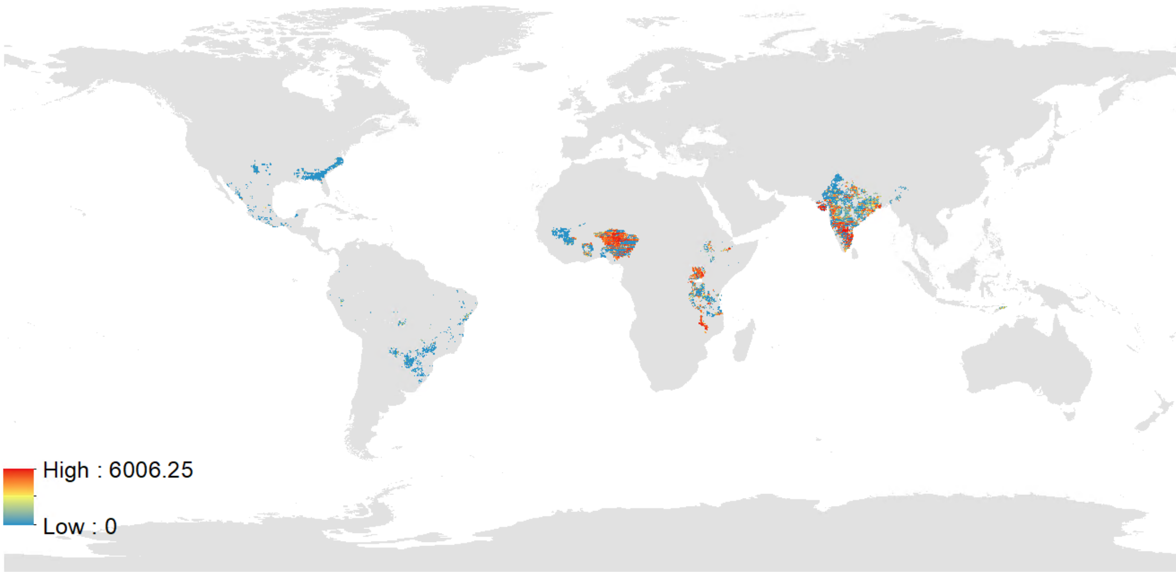
Coconut



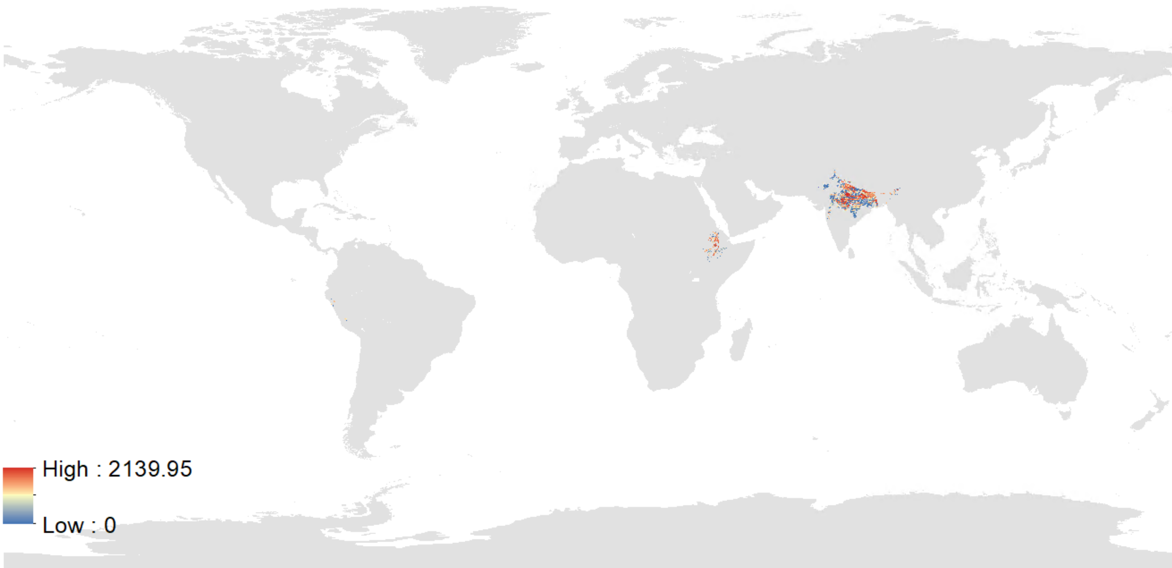
Cotton



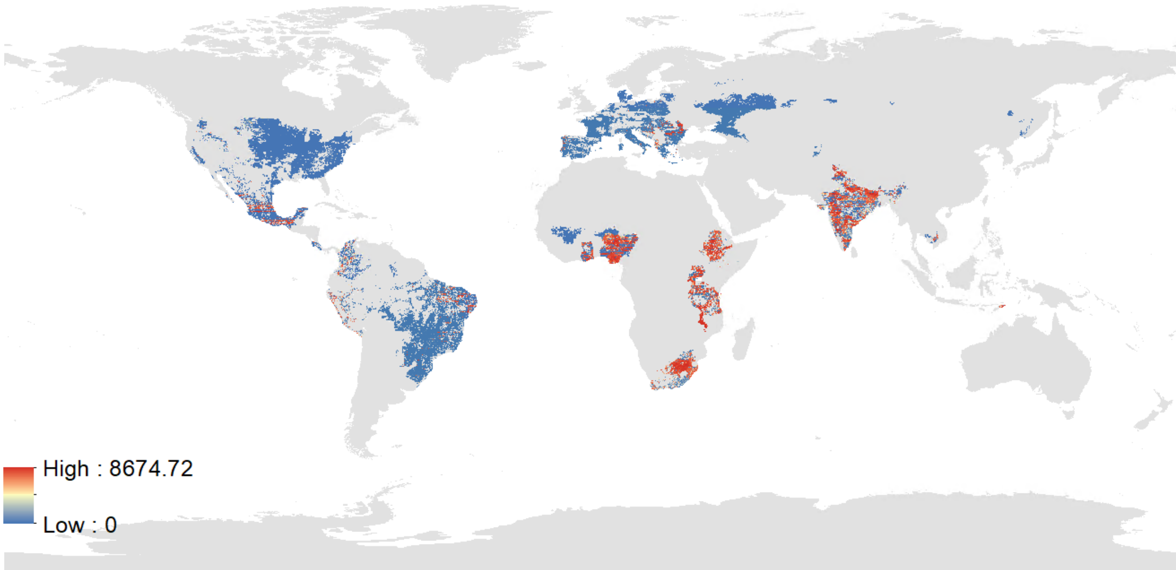
Cowpea



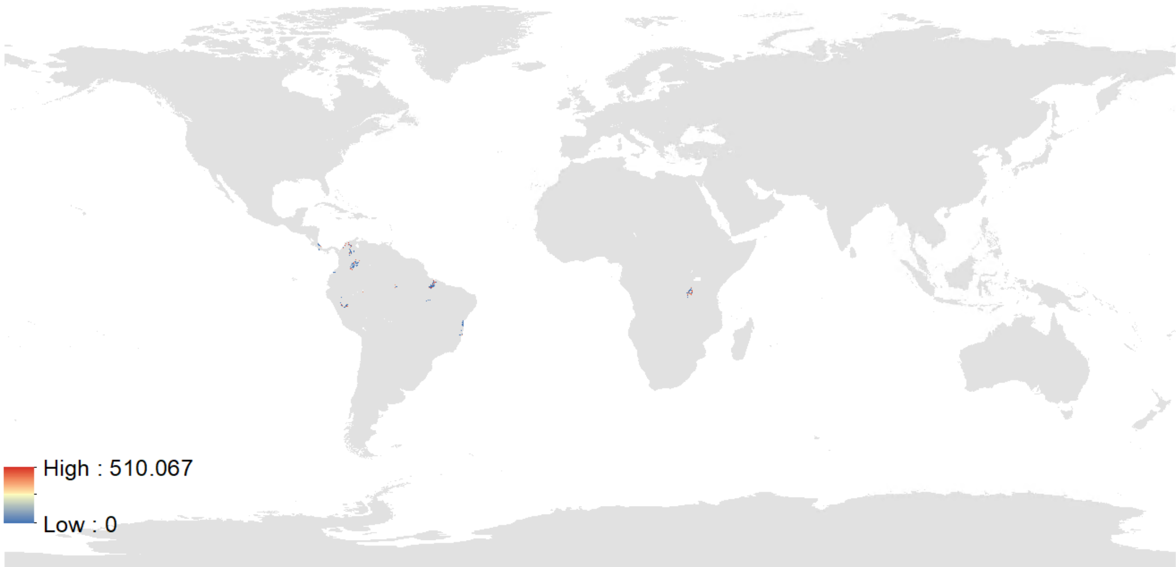
Groundnut



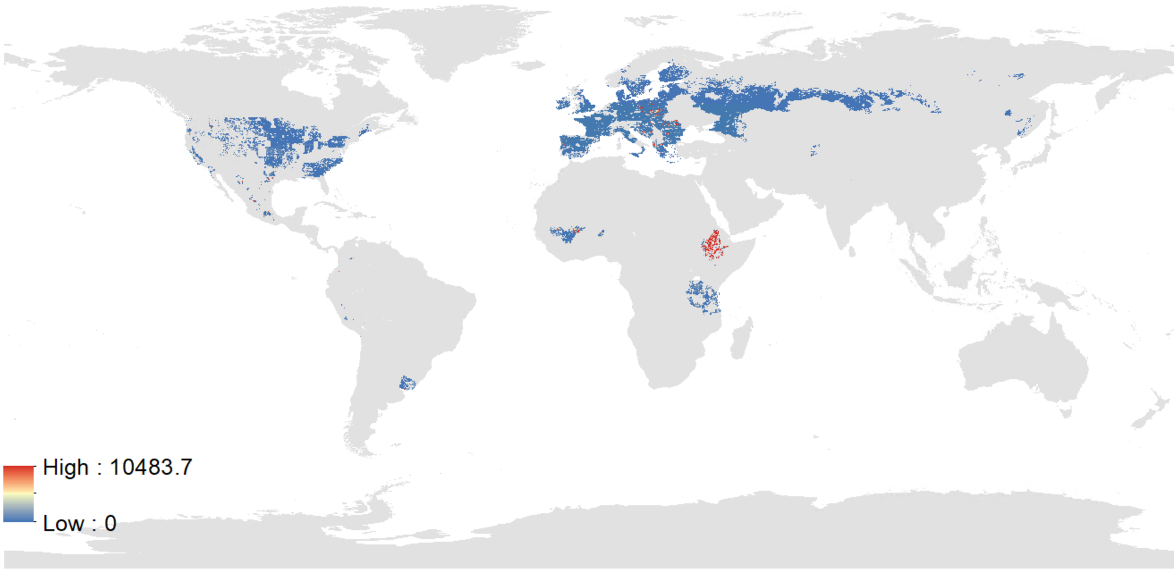
Lentil



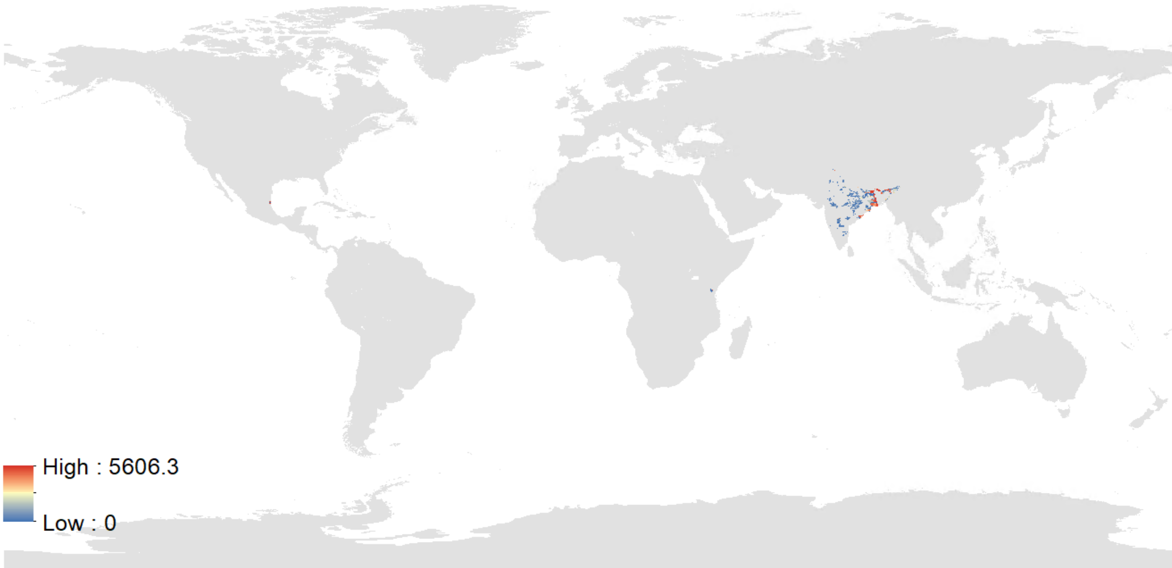
Maize



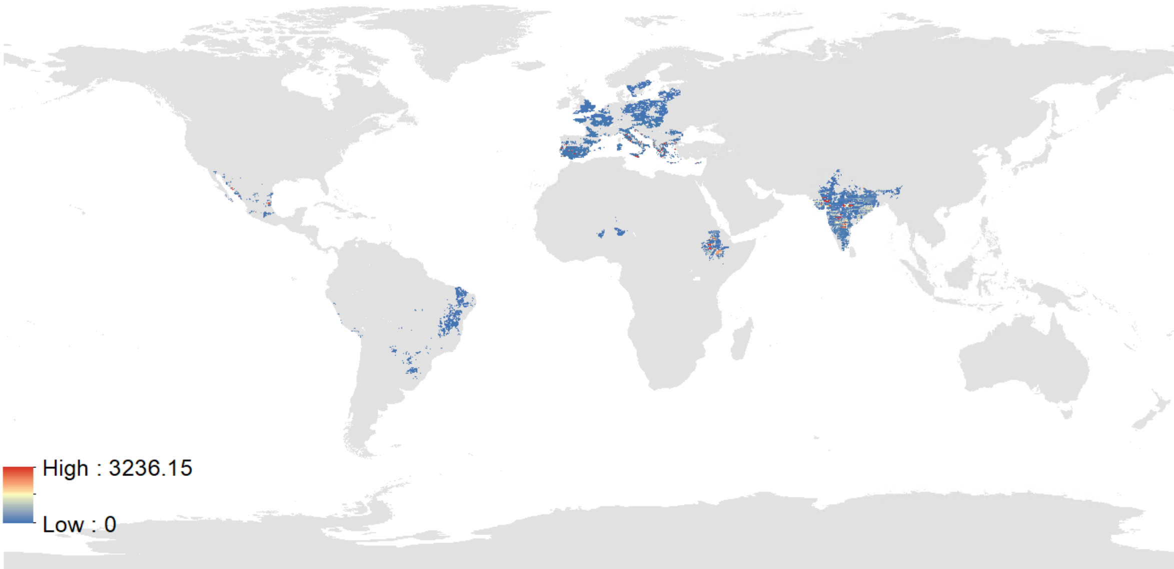
Oil palm



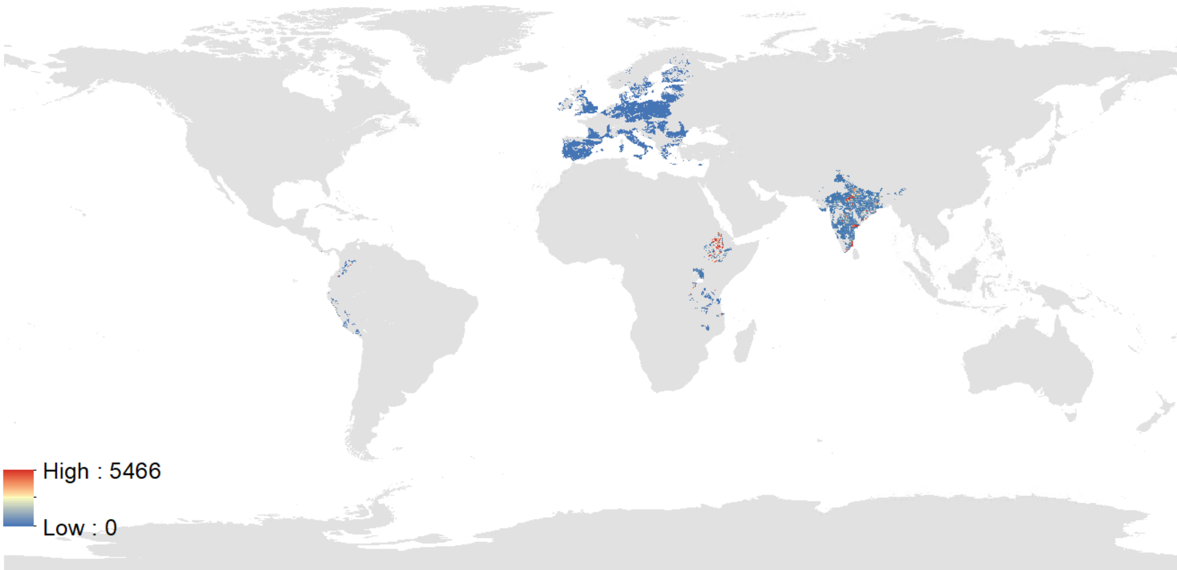
Other cereals



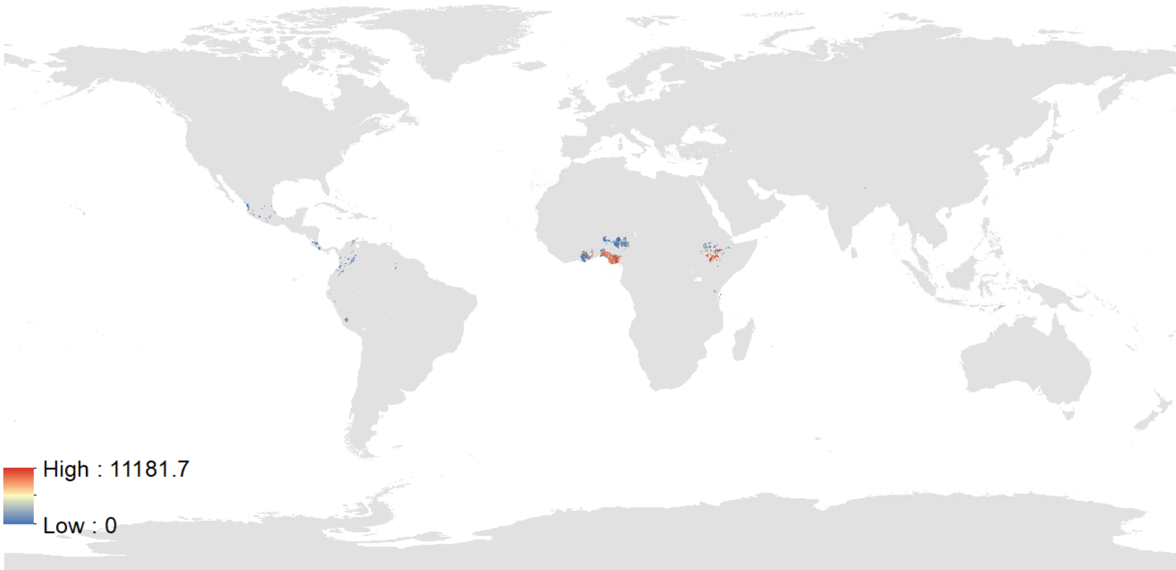
Other fibre crops



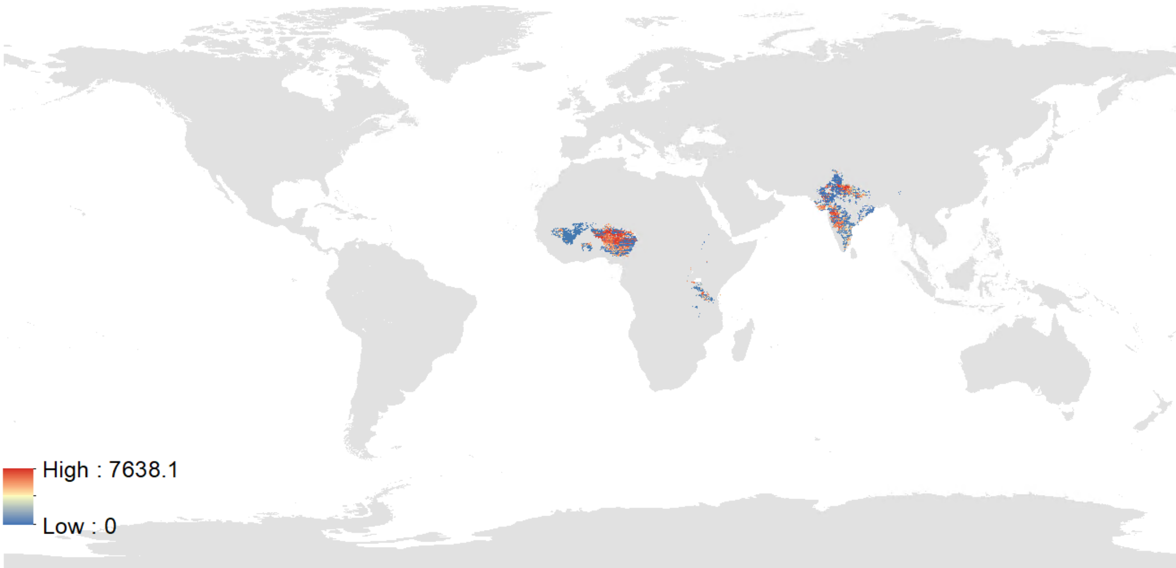
Other oil crops



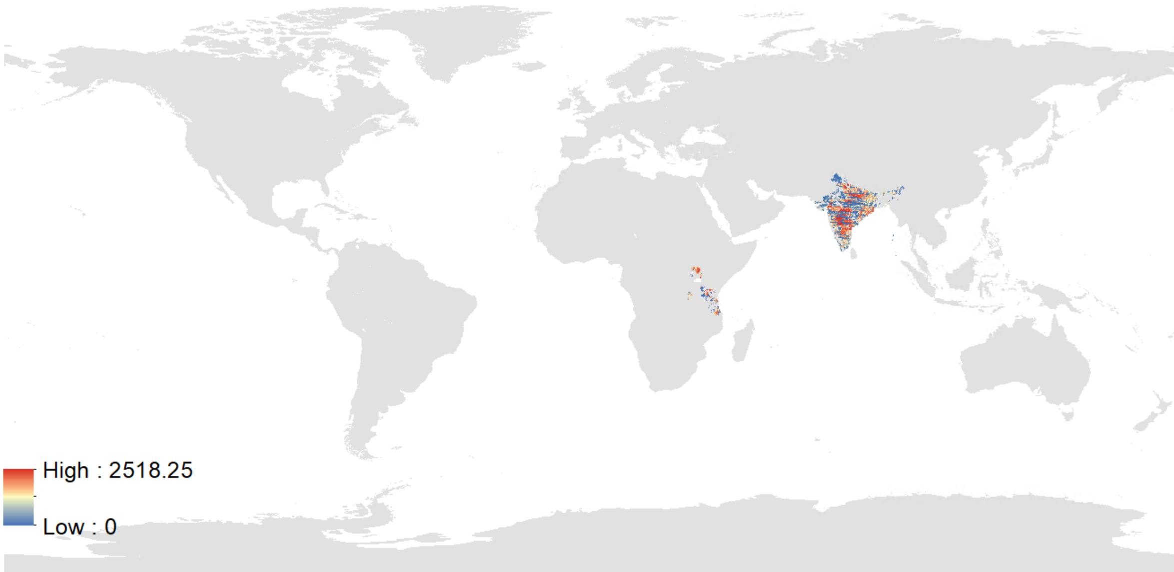
Other pulses



Other roots

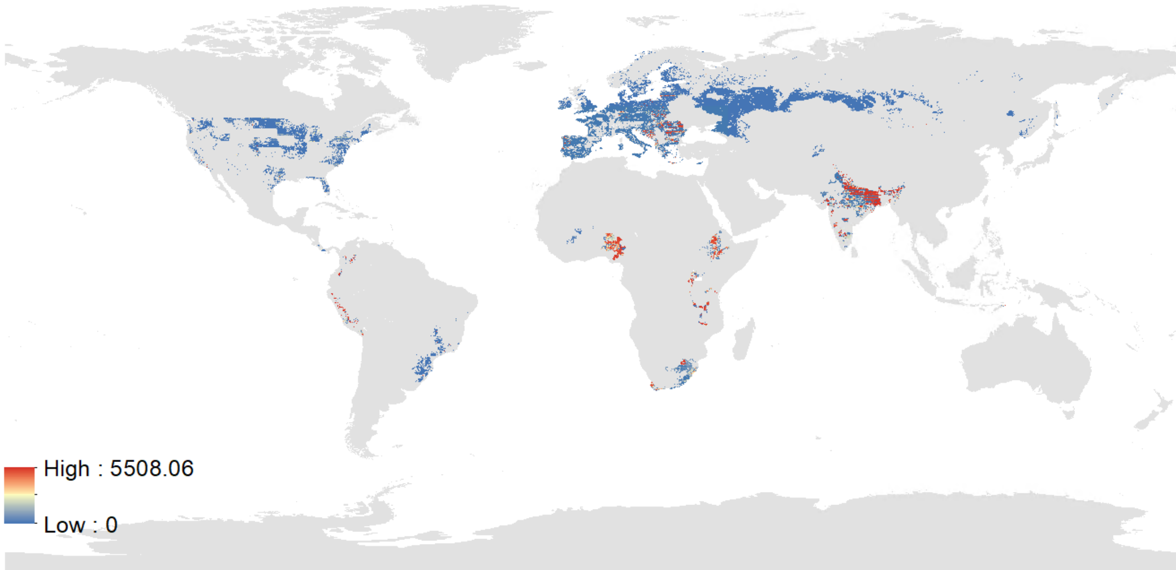


Pearl millet

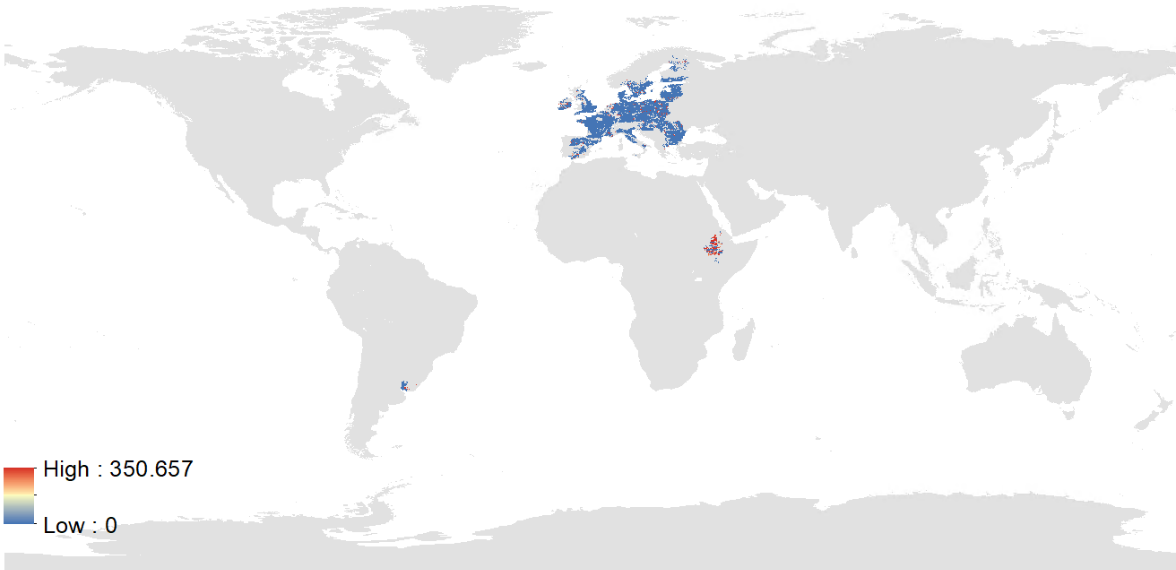


Pigeon pea

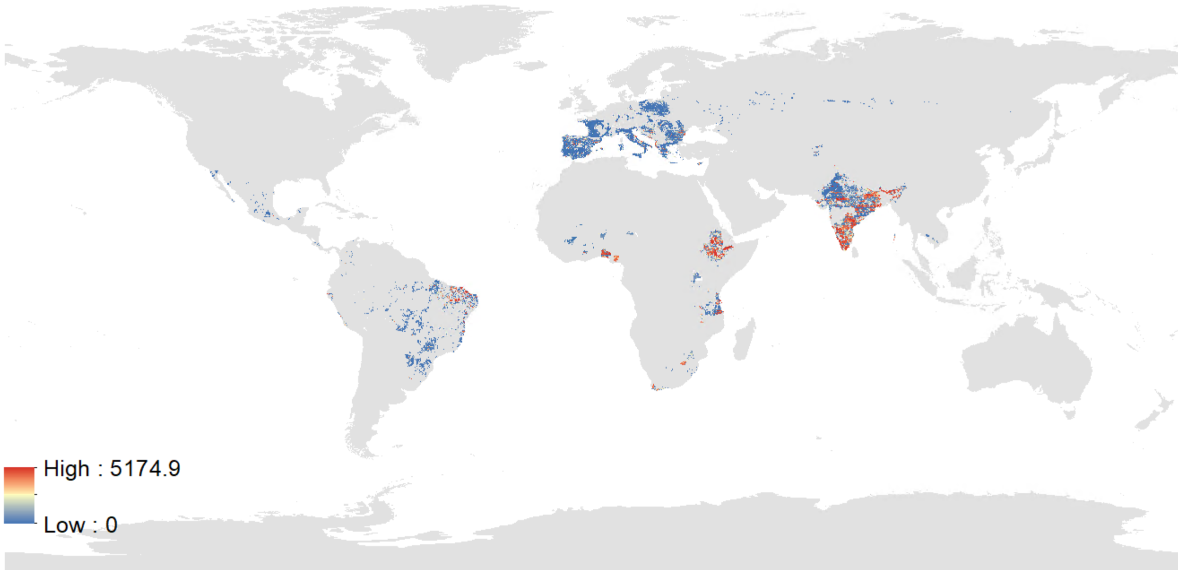
Plantain



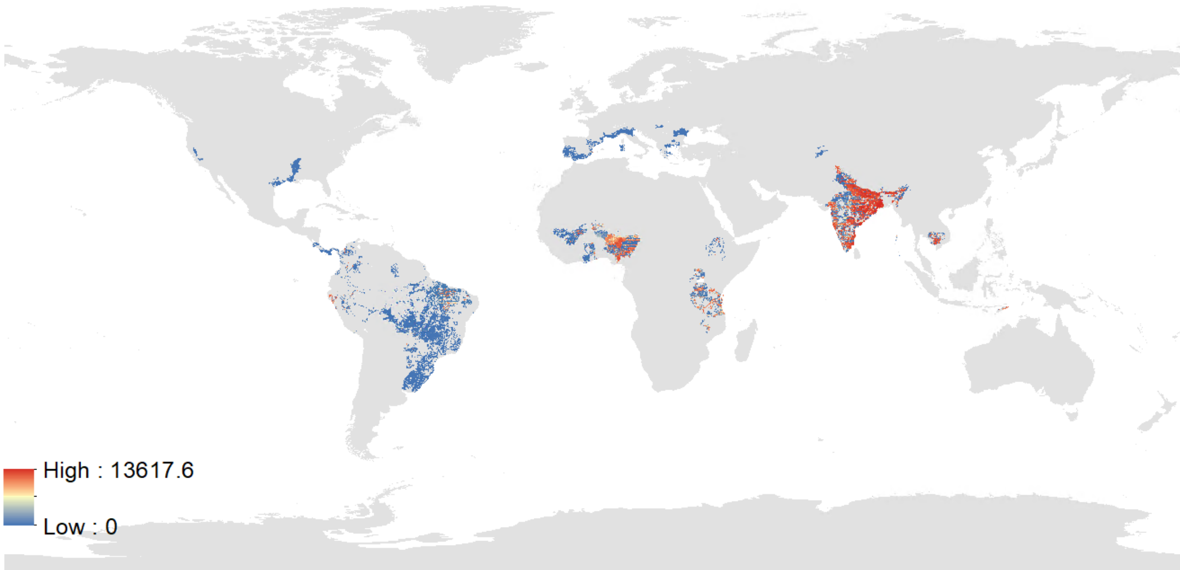
Potato



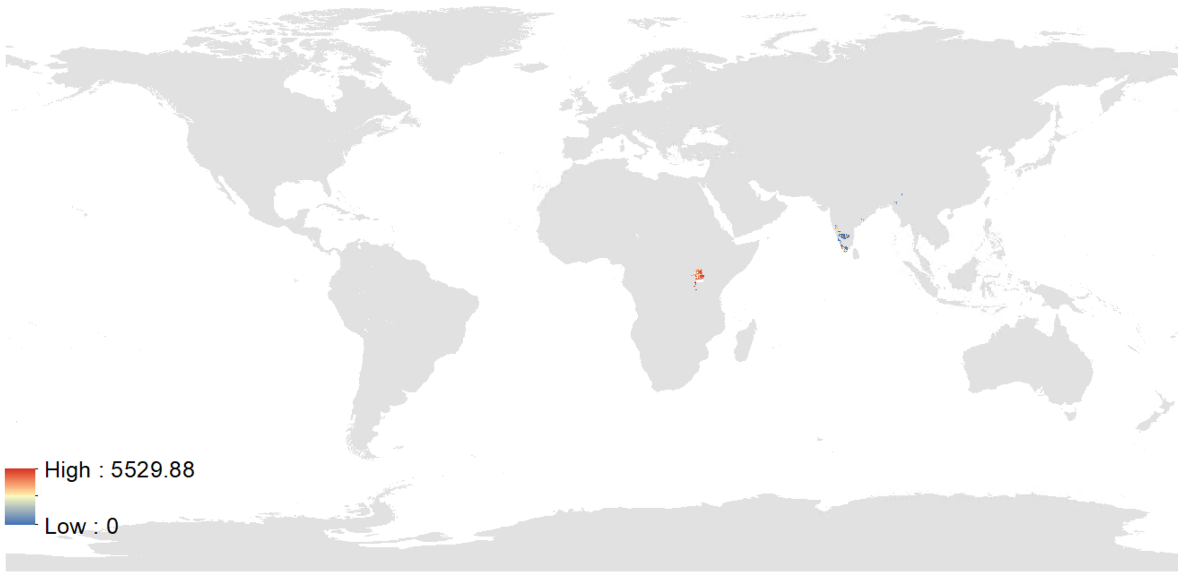
Rapeseed



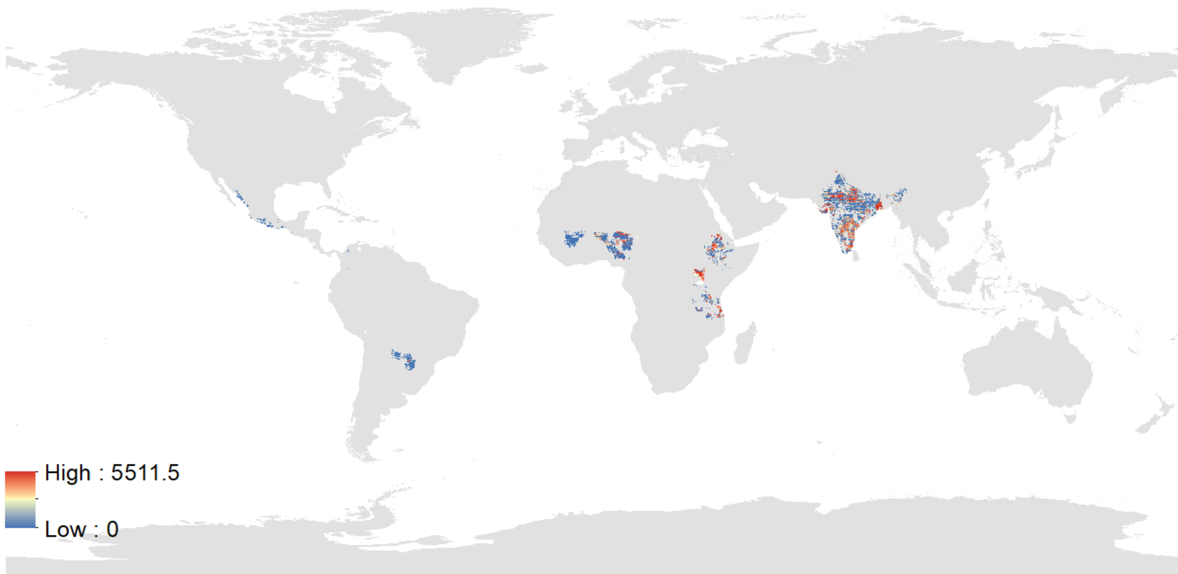
Rest of crops



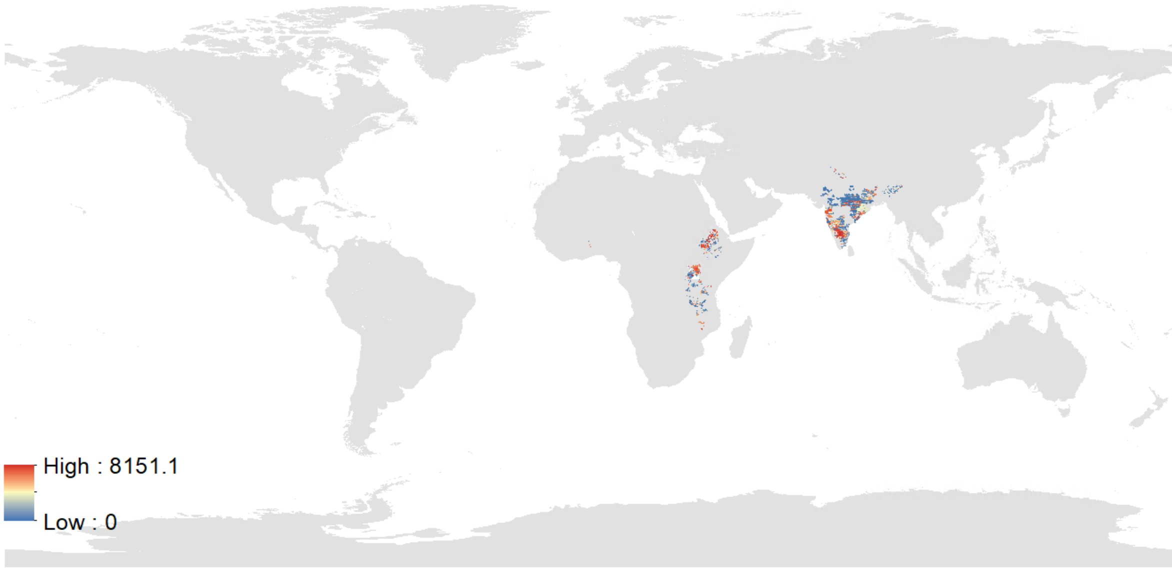
Rice



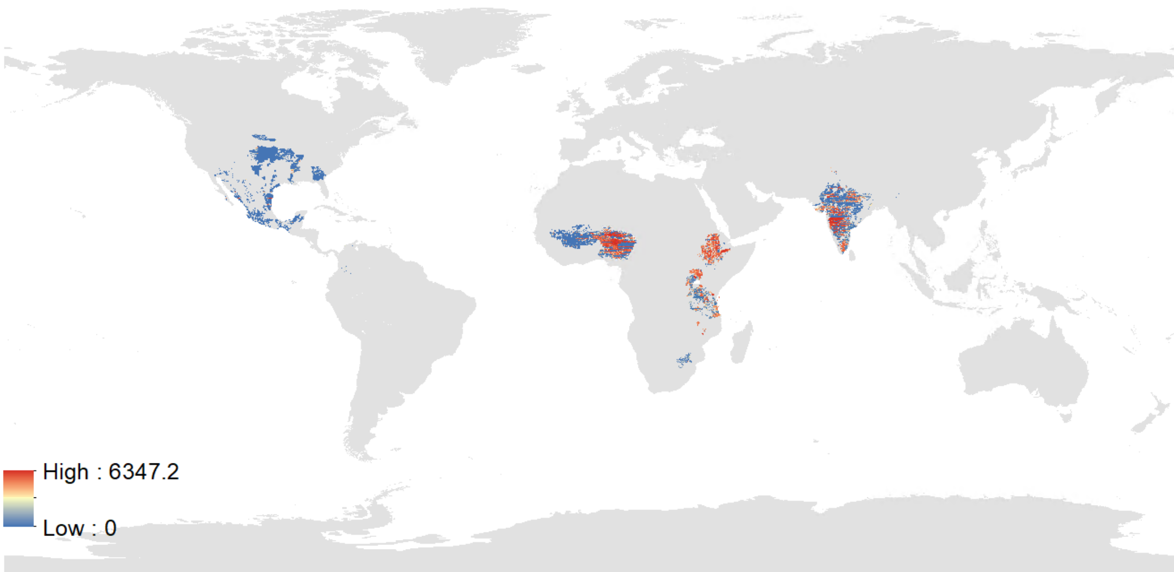
Robusta coffee



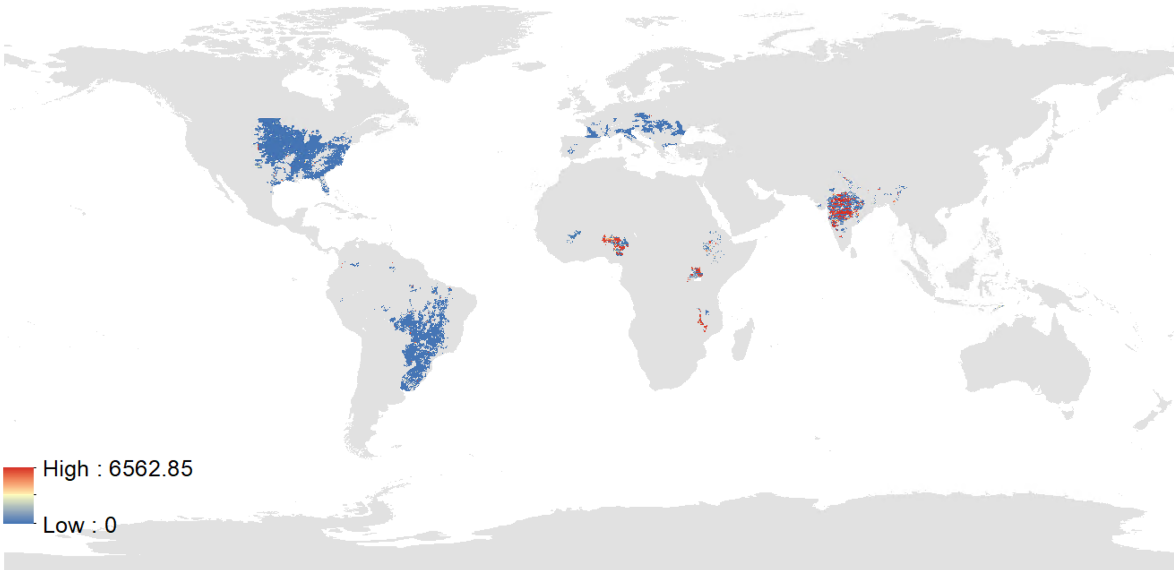
Sesame seed



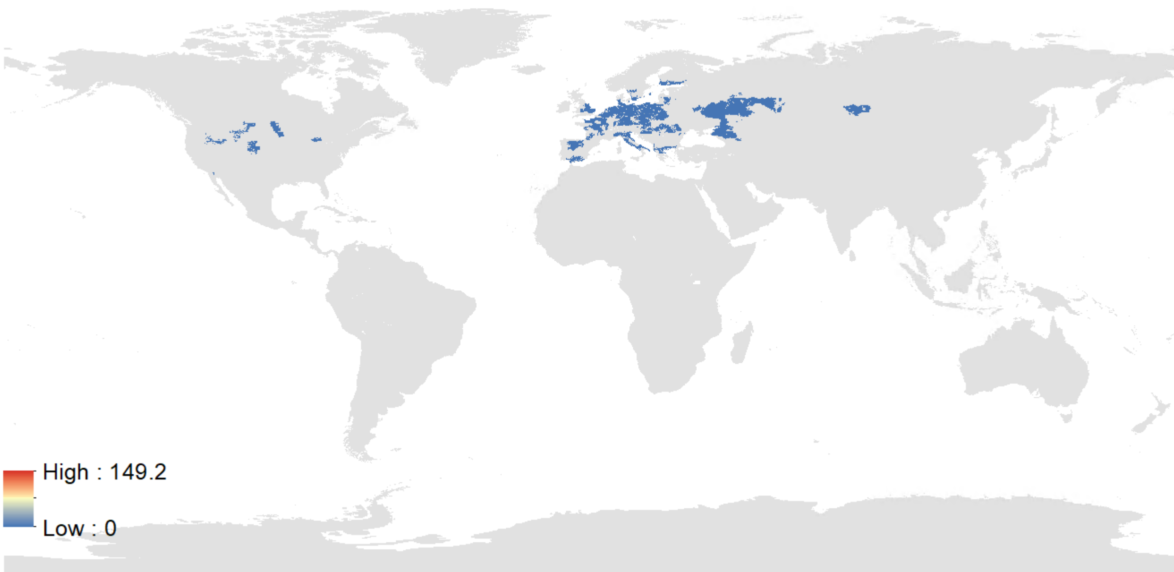
Small millet



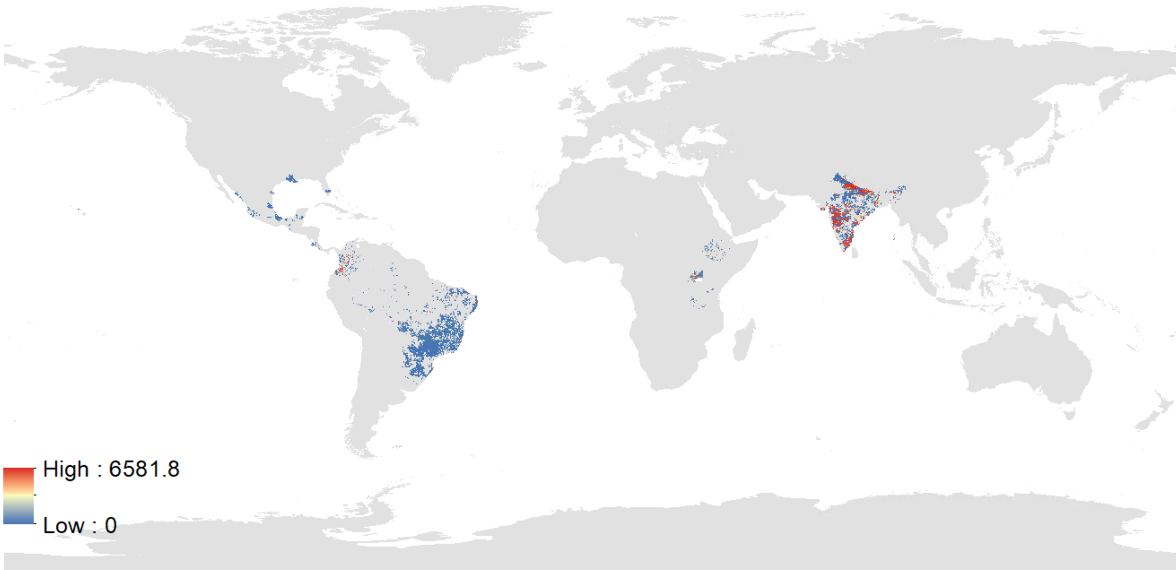
Sorghum



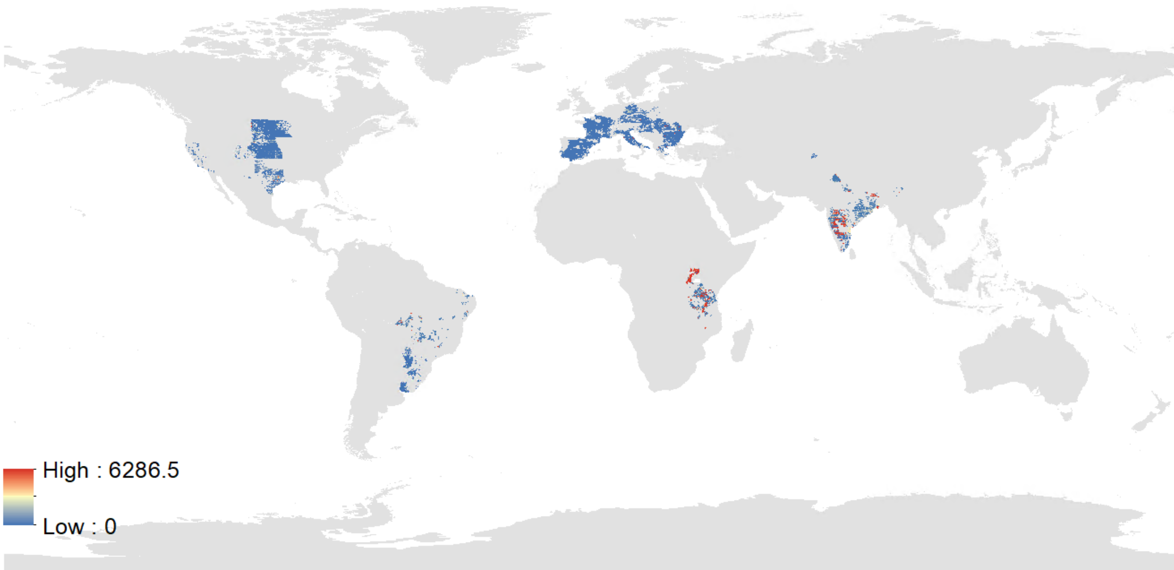
Soybean



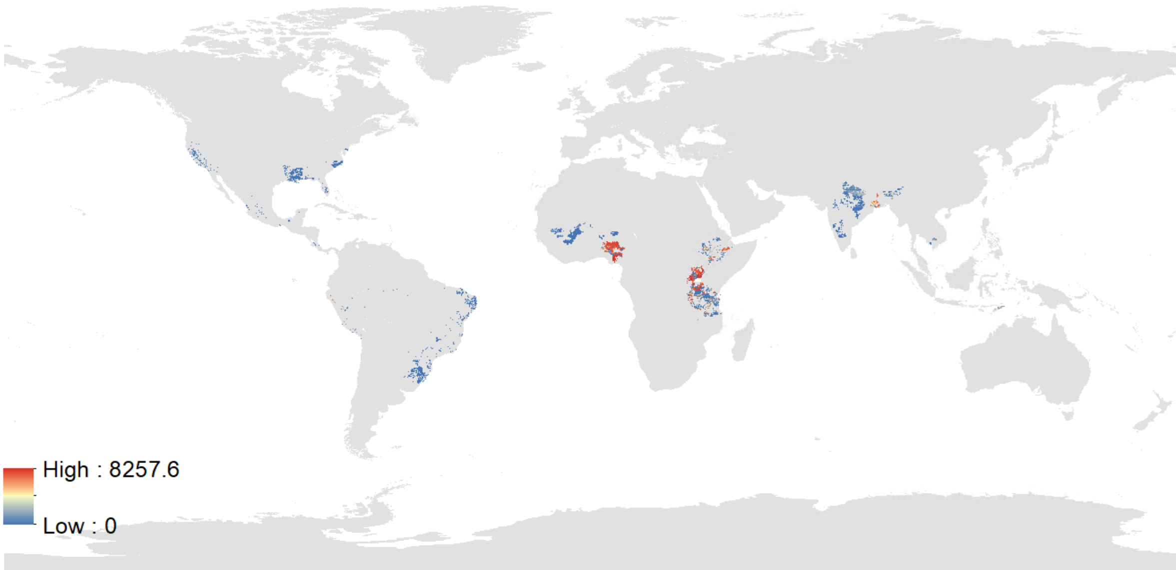
Sugarbeet



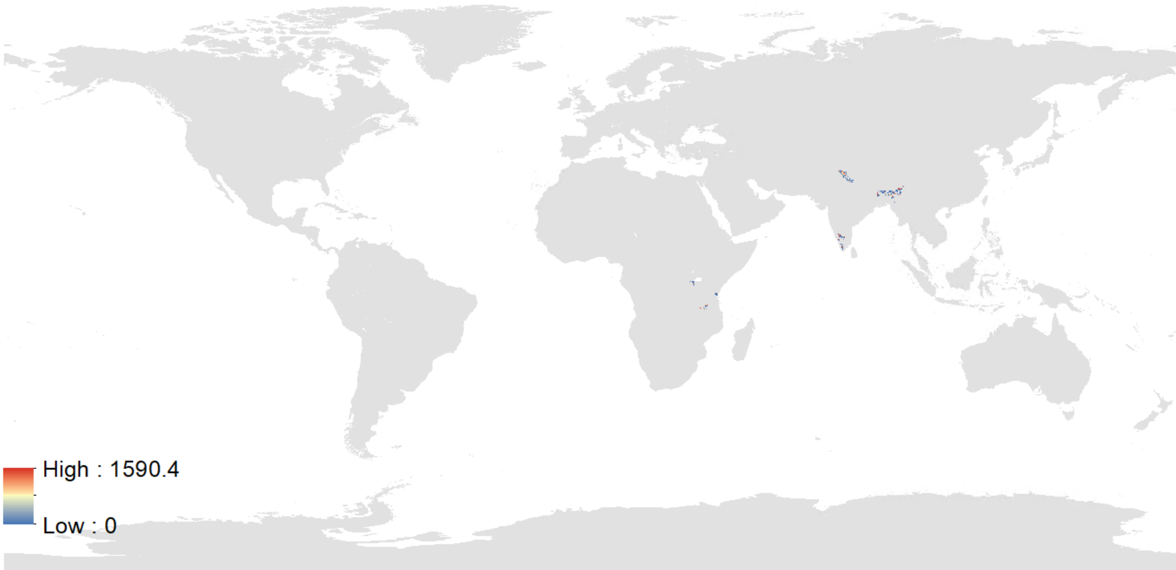
Sugarcane



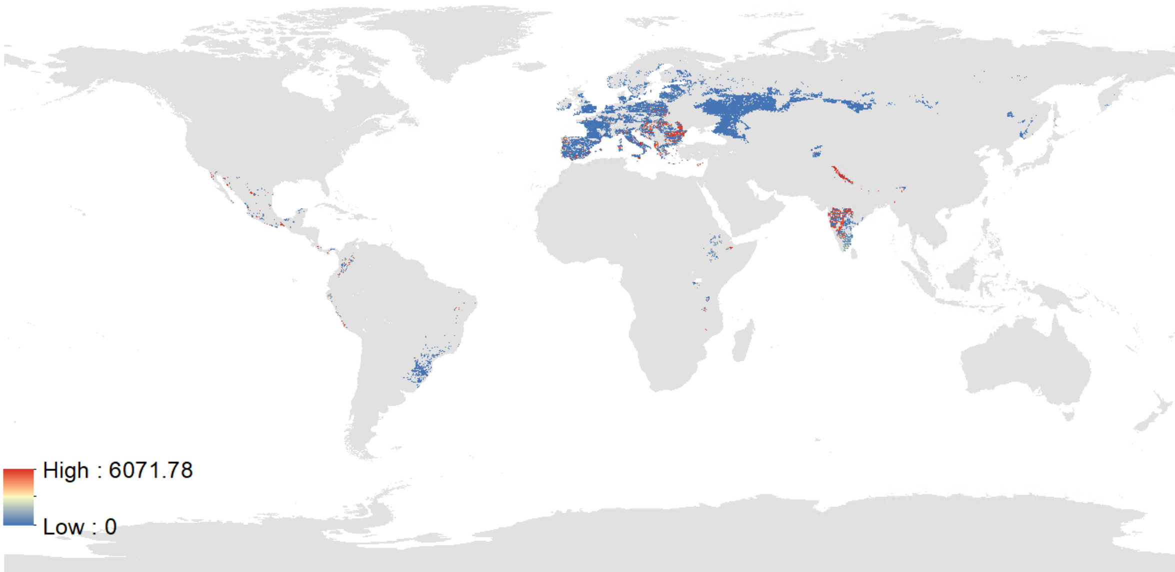
Sunflower



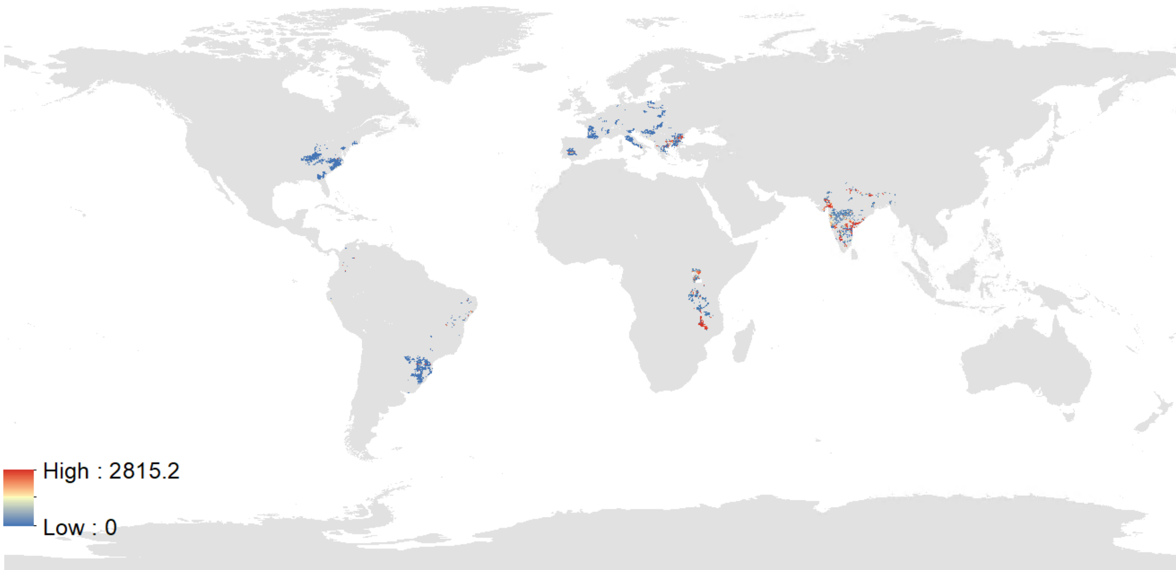
Sweet potato



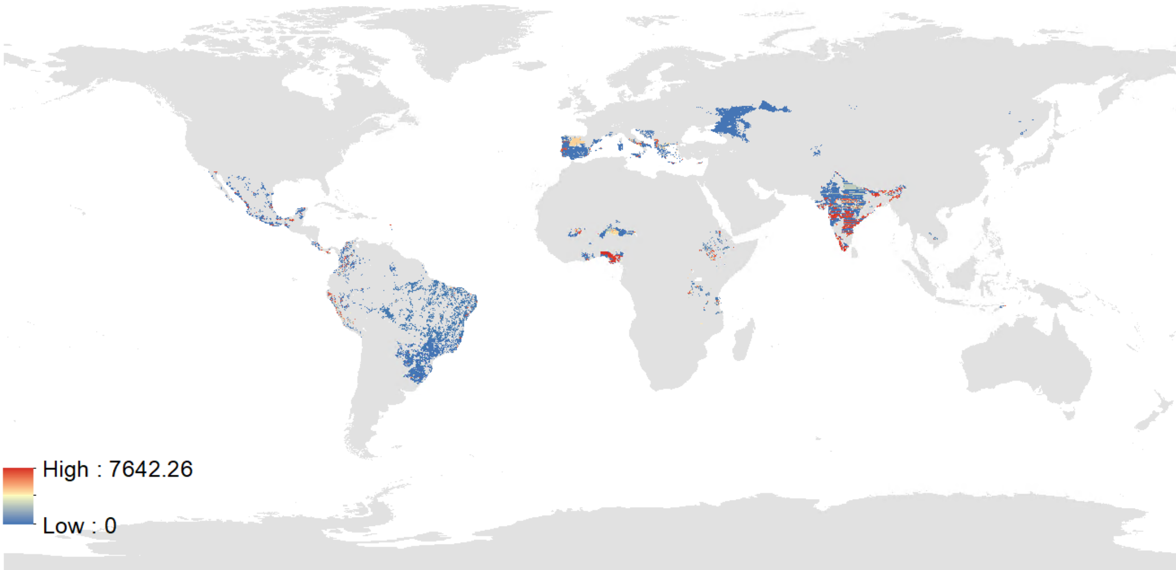
Tea



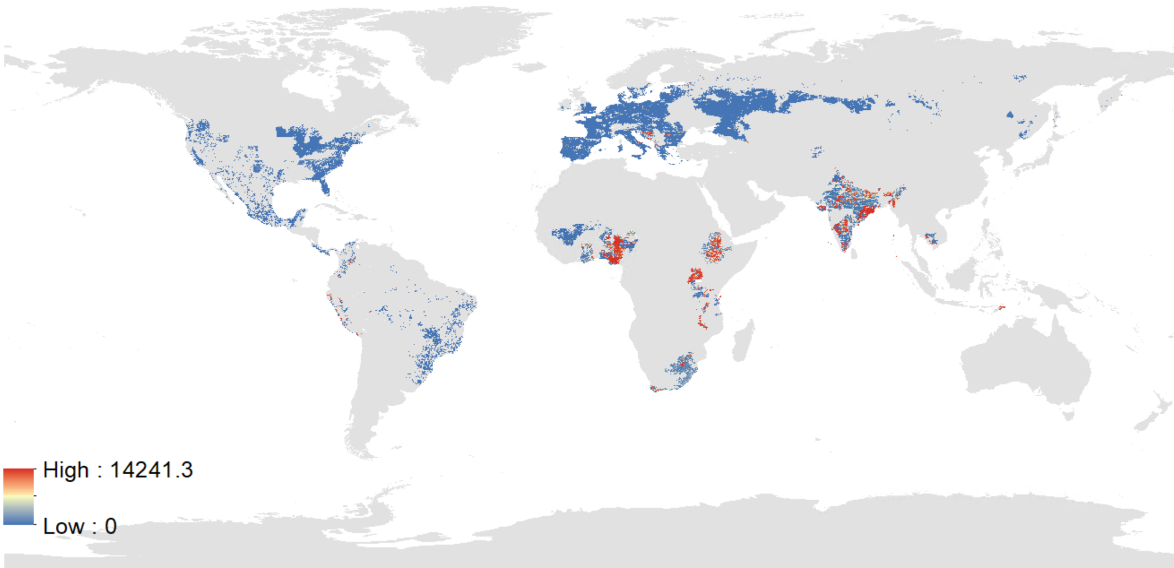
Temperate fruit



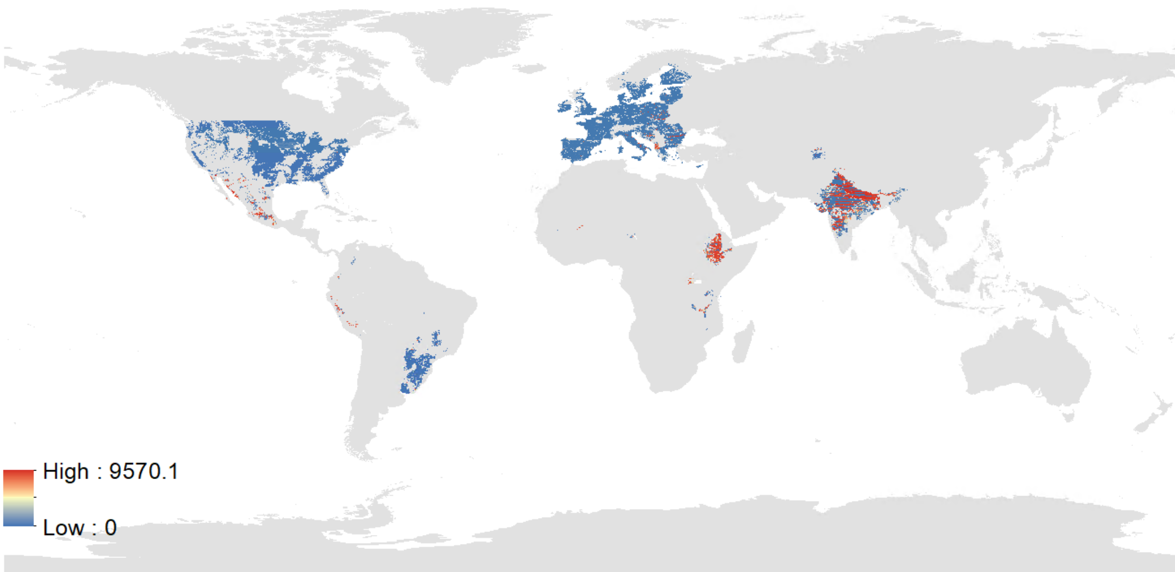
Tobacco



Tropical fruit



Vegetables



Wheat

Yams