ENVIRONMENTAL RESEARCH
LETTERS

LETTER • OPEN ACCESS

Potential, attainable, and current levels of global crop diversity

To cite this article: Fernando Aramburu Merlos and Robert J Hijmans 2022 *Environ. Res. Lett.* **17** 044071

IOP Publishing Environ. Res. Lett. 17 (2022) 044071 https://doi.org/10.1088/1748-9326/ac62ab

ENVIRONMENTAL RESEARCH
LETTERS

OPEN ACCESS

RECEIVED
24 December 2021

REVISED
22 February 2022

ACCEPTED FOR PUBLICATION
30 March 2022

PUBLISHED
8 April 2022

LETTER

Potential, attainable, and current levels of global crop diversity

Fernando Aramburu Merlos[1,2,*] **and Robert J Hijmans**[1]

[1] Department of Environmental Science and Policy, University of California Davis, Davis, CA, United States of America
[2] Instituto Nacional de Tecnología Agropecuaria (INTA, Unidad Integrada Balcarce), Ruta 226, km 73.5, CC 276, CP 7620 Balcarce, Buenos Aires, Argentina
[*] Author to whom any correspondence should be addressed.

E-mail: faramburumerlos@ucdavis.edu

Keywords: attainable diversity, potential diversity, diversity gap, crop distribution, crop suitability, crop diversification, agroecology

Supplementary material for this article is available online

Abstract

High levels of crop species diversity are considered beneficial. However, increasing diversity might be difficult because of environmental constraints and the reliance on a few major crops for most food supply. Here we introduce a theoretical framework of hierarchical levels of crop diversity, in which the environmental requirements of crops limit potential diversity, and the demand for agricultural products further constrain attainable crop diversity. We estimated global potential, attainable, and current crop diversity for grid cells of 86 km^2. To do so, we first estimated cropland suitability values for each of 171 crops, with spatial distribution models to get estimations of relative suitability and with a crop model to estimate absolute suitability. We then used a crop allocation algorithm to distribute the required crop area to suitable cropland. We show that the attainable crop diversity is lower in temperate and continental areas than in tropical and coastal regions. The diversity gap (the difference between attainable and current crop diversity) is particularly large in most of the Americas and relatively small in parts of Europe and East Asia. By filling these diversity gaps, crop diversity could double on 84% of the world's agricultural land without changing the aggregate amount of global food produced. It follows that while there are important regional differences in attainable diversity, specialization of farms and regions is the main reason for low levels of local crop diversity across the globe, rather than our high reliance on a few crops.

1. Introduction

High crop species diversity is considered important for agriculture sustainability (Jones *et al* 2021) because of its positive association with food production stability and resilience (Gaudin *et al* 2015, Renard and Tilman 2019). However, it is unclear how much diversity would be enough or desirable and how this varies between locations. Not all crops can grow everywhere, and what may be considered low diversity in one region could be beyond what is ecologically possible in another region. Moreover, only a few crops provide the vast majority of our food supply, which *a priori* imposes a severe demand-side constraint on diversification (Cassman and Grassini 2020, Renard and Tilman 2021). Even though it could be desirable to reduce the importance of the most dominant crops (Tilman and Clark 2014), vast areas would still be needed to produce staple crops such as wheat, rice, and cassava, but not for many other crops that are only useful in relatively small quantities. Thus, to understand current diversity patterns and assess opportunities for diversification, it is essential to consider both ecological constraints (which crops can be grown in a location) and economic constraints (how much demand is there for these crops).

To allow for such analysis, we need a framework to determine what levels of crop species diversity are possible to compare these with the actual situation. Here we provide and apply such a framework inspired by concepts from production ecology and the work on crop yield gaps (van Ittersum *et al* 2013). We first define different theoretical levels of crop diversity (maximal, potential, and attainable)

and then calculate their present values (circa 2010) for 86 km^2 grid cells for the entire world. The theoretical levels of diversity depend on estimates of crop-specific cropland suitability that we computed in two ways: using spatial distribution models (SDMs) (relative suitability) and rule-based crop models (absolute suitability). We then used a crop allocation algorithm to predict potential and attainable crop distributions and calculate the corresponding diversity level. Finally, we contrasted these levels to current patterns of diversity and computed diversity gaps.

2. Theoretical levels of crop diversity

Crop diversity (*D*) has been defined as the effective number of different crop species planted in a given area (Jost 2006). The term effective refers to the number of equally abundant virtual species that has the same entropy as the actual species when considering their relative abundance.

To better interpret patterns in current crop diversity (*cD*), we define three new theoretical concepts: maximum, potential, and attainable crop diversity (figure 1). Maximum diversity (*mD*) results from planting all crops that can be grown in an area in equal proportion. It has very limited relevance, and while we include it in our framework, we do not discuss it further. Potential diversity (*pD*) is reached when the area planted with each crop is a function of crop-specific cropland suitability (the proportion of land planted to the best-adapted crops is largest). Attainable diversity (*aD*) is obtained if all crops are planted to maximize *D* while considering crop-specific cropland suitability, as well as the demand for different crops and the interspecific competition for land. The main difference between *aD* and *pD* is that *aD* is constrained to meet crop-specific total demand (for any purpose, including food, feed, fiber, and industrial use). We define crop-specific 'demand' as equal to each crop's total current production (supply). Thus, *aD* can be reached without changing total consumption and crop diversity at the global level (or, more generally, in the entire study region in question, which could be a country). Therefore, by definition, at the global (study region) level, *aD* equals *cD*, but this is not true at lower levels of aggregation (areas within the study region). *pD* is directly proportional to the number of crops considered, as it assumes that any crop can take an equal amount of land. In contrast, *aD* is less sensitive to the omission of rare crops.

Lastly, we define the diversity gap (*Dg*) as the difference between the *aD* and *cD*, expressed as a percentage of *aD*. We refer to all factors reducing crop diversity from its attainable to its current level with the broad term specialization, which includes many factors not directly controlled by farmers, such as access to market, technology, and know-how (figure 1).

3. Methods

3.1. Data

We used crop distribution data from two sources: 'SPAM' (IFPRI 2019) and 'Monfreda' (Monfreda *et al* 2008). These data sets include gridded, at a 5 arc-minutes spatial resolution, crop-specific physical (SPAM) and harvested (SPAM and Monfreda) areas (the same physical area may be harvested more than once per year) that were generated by downscaling regional (national and subnational) crop statistics over the available cropland area. SPAM includes data for 42 crop categories (33 individual crops and nine crop groups), while Monfreda provides data for 175 crops. We merged both data sets prioritizing SPAM crop physical area to end with a total of 171 crops. See section S.3.1 (available online at stacks.iop.org/ERL/17/044071/mmedia) in the supplementary material for further details on these data sets, how they were merged, and an assessment of their quality.

Cropland suitability predictors were derived from Soil Grids (soil pH, Hengl *et al* 2017), AQUASTAT (irrigation availability, FAO 2016), and WorldClim (climatic and bioclimatic variables, Fick and Hijmans 2017). All variables were aggregated to 5 min spatial resolution (about 9 × 9 km at the Equator) to match the crop data, and crop suitability and allocation were computed at that spatial resolution.

3.2. Crop suitability

We applied two modeling approaches to estimate crop-specific cropland suitability. We used a SDM approach to compute 'relative' suitability and a rule-based model to compute 'absolute' suitability.

SDMs are commonly used to predict relative environmental suitability by assessing the similarity between the conditions at a site of interest and the conditions at locations of known occurrence or abundance (Elith and Leathwick 2009). We refer to this approach as 'relative' suitability because of the (implicit) effect of competition on species distributions: any crop observed abundance is a function of the suitability of a site for that crop, but also for other crops. Here, we predicted the suitability of all cropland for each crop using the crop distribution data as the response variable, bioclimatic, soil pH, and irrigation variables as predictors, and three algorithms: Maxent, Random Forest regression, and Boosted Regression Trees. See section S.3.2 in the supplementary material for details on SDMs methods.

We used the ECOCROP model (Hijmans and Graham 2006, Hijmans 2021) to predict 'absolute suitability', which indicates where a species can be grown without major environmental constraints. ECOCROP is a rule-based model that estimates absolute environmental suitability for each species or subspecies from a combination of dynamic (monthly)

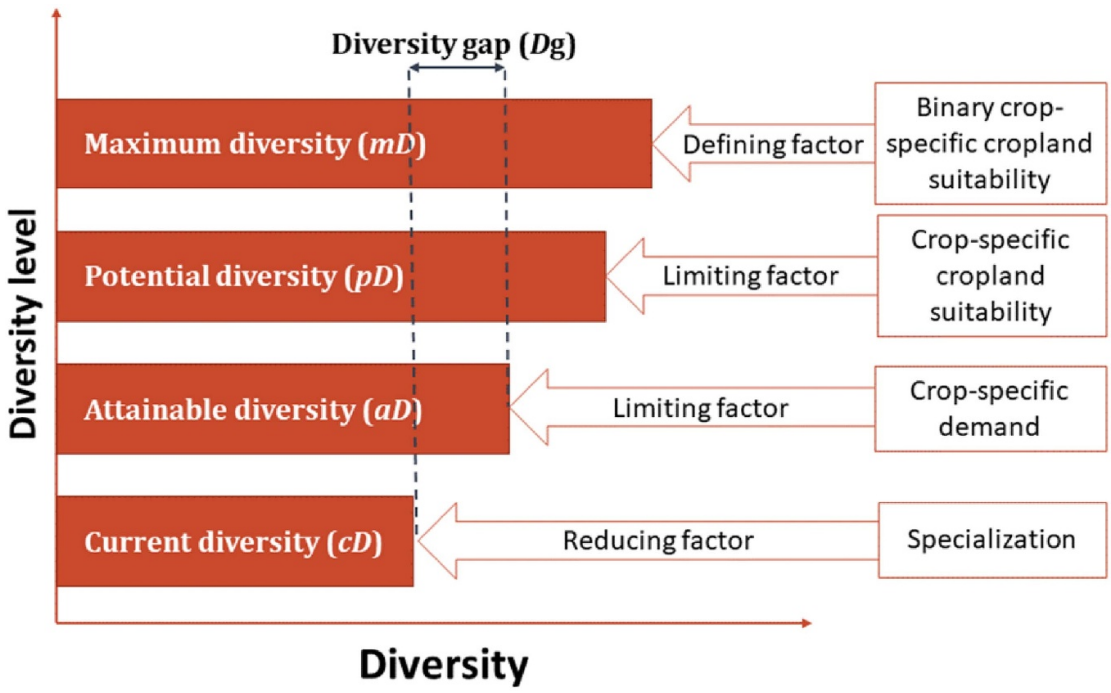

Figure 1. Crop diversity levels as determined by defining, limiting, and reducing factors. The defining factor for maximum diversity (*mD*) is the number of crops that can be grown to harvest. Potential diversity (*pD*) is limited by the unevenness in the environmental requirements of crops (better-adapted crops are more abundant, reducing diversity). Attainable diversity (*aD*) is further constrained by the unevenness in demand for different crops (crop supply should match demand). Current diversity (*cD*) is further reduced due to specialization. This figure and framework were inspired by concepts from the production ecology and yield gaps literature (van Ittersum and Rabbinge 1997, van Ittersum *et al* 2013).

and static predictors, including monthly average and minimum temperature, monthly precipitation, and soil pH. For all variables, default parameters indicate the extreme minimum and maximum value beyond which the crop cannot grow (suitability is zero) and a minimum and maximum optimal value within which suitability is one. Between extreme and optimal values, suitability is determined with linear interpolation between zero and one. See section S.3.3 in the supplementary material for ECOCROP model calibration and usage details.

3.3. Crop allocation

We developed a cross-entropy-based spatial allocation algorithm to compute potential and attainable crop distributions, using each crop's relative or absolute cropland suitability as priors. It is described in detail in section S.3.4 in the supplementary material. Similar algorithms have been used to downscale regional crop area data to generate current and historical global crop distribution maps (You *et al* 2014, Jackson *et al* 2019). Potential diversity only considers adaptation and does not consider demand; thus, areas allocated to each crop are proportional to the crop-specific cropland suitability. In contrast, for attainable diversity, the total global demand for each crop is considered, and we used our allocation algorithm to distribute required areas for each crop to their most suitable cropland.

3.4. Diversity and diversity gap calculation

We quantified crop diversity (D) as the effective number of crop categories, which is the inverse of the weighted average of their proportional abundances and indicates the number of equally-abundant virtual crops with the same entropy as the actual crops (Jost 2006, Tuomisto 2010). We computed these averages as the exponent of the Shannon entropy, using nominal weights (each crop affects the mean based on their relative proportion) to avoid over or underrepresenting rare crops (equation (1)).

$$^{1}D = \exp\left(-\sum_{j=1}^{n}\left(p_j \ln p_j\right)\right) \quad (1)$$

where p_j is the proportion of crop area occupied by crop j and n is the total number of crops.

We also computed diversity with the inverse of the Simpson index (^{2}D), which gives more weight to the most dominant species. The differences between these two approaches are described in section S.4.2 in the supplementary material.

Because of diversity scale-dependency (Aramburu Merlos and Hijmans 2020), we transformed the current and allocated crop areas to raster with equal-area grid-cells, using the Equal Earth map projection (Šavrič *et al* 2019). We chose a 9.26 × 9.26 km spatial resolution (ca 86 km^2) to match the largest grid cells of the original (longitude/latitude) raster data (i.e. at the Equator). We then

computed crop diversity for each 86 km^2 grid cell (local diversity) and at the country level (total national diversity).

Moreover, the local-diversity-average ($\overline{D\alpha}$) (Jost 2007, Tuomisto 2010) was computed for each country and diversity level with (equation (2)), in which m is the number of cells in a given country, and w_j is the weight of cell j, computed as the cropland area in cell j divided by the total cropland of that country. Note that this is the local (86 km^2 cell) average national diversity and different from the total national diversity, $D\gamma$, used in most global crop species diversity analyses (Renard and Tilman 2019, Aguiar *et al* 2020).

$$D\alpha = \exp\left[-\sum_{j=1}^{m}\left(w_j\sum_{i=1}^{n}\left(p_{ij}\ln p_{ij}\right)\right)\right]. \quad (2)$$

Diversity gaps (D_g) were computed as the difference between aD, averaged across the two methods used to compute aD, and cD, relative to aD and multiplied by 100 to express D_g as a percentage.

$$Dg\ (\%) = \frac{(aD - cD)}{aD} \times 100. \quad (3)$$

3.5. Software

All the analysis was done in R (R Core Team 2020), including data preparation, modeling, allocation algorithm, data analysis, and mapping, with the packages listed in section S.3.5 in the supplementary material. The code is available on GitHub (https://github.com/aramburumerlos/globcropdiv).

4. Results

4.1. Current diversity

There are large extents with high levels of current cD in East Asia (China, the Korean peninsula, and Japan), Sub Saharan Africa except for the driest regions, and the Mediterranean, especially Portugal, Italy, and western Turkey. cD is also high in other parts of Europe (such as the Netherland and Belarus), parts of India, New Zealand's North Island, Peru and Central Chile in South America, the Caribbean islands, and the west coast of the United States (figure 2 and supplementary figure S1). In contrast, cD is very low in most other parts of the Americas: Argentina, Uruguay, and Brazil (dominated by soybean), Mexico (maize), and the central United States (maize and soybean in the east, wheat in the west); central Asia: Afghanistan and Kazakhstan (wheat); and in parts of Southeast Asia: Thailand and Cambodia (rice), and Malaysia (oil palm) (figures 2 and 3(a)). These regions have one or two major crops covering more than 50% of the cropland area (supplementary figures S2 and S3).

cD is highest (around 20 effective crops) around the Equator and the Tropic of Cancer (Northern tropic, 23° N), from where it linearly decreases when going northwards (figures 2 and 4), dropping to four at around 64° N. In contrast, south of the Equator, crop diversity decreases rapidly with latitude and reaches eight around the Tropic of Capricorn (23° S). It then remains stable until 40° S, and there is not much cropland further south (supplementary figure S4). The difference in cD between the southern and northern hemispheres (figure 4) is strongly associated with the lower amount of cropland in the southern hemisphere, most of which is low diversity cropland in South America. In contrast, in the northern hemisphere, there is more high diversity cropland in Europe and East Asia than low diversity cropland in North America (supplementary figures S4 and S5).

The countries with the most extreme (high or low) local diversity average ($c\overline{D\alpha}$) are small countries with little cropland area (e.g. Grenada and Western Sahara, supplementary table S1), perhaps because crop diversity sample variance is higher at small sampling units (Aramburu Merlos and Hijmans 2020). When only considering countries with at least 0.1 Mha of cropland, Israel stands out for its high $c\overline{D\alpha}$ of 19.5. Lebanon, Italy, Taiwan, Portugal, Cuba, and Republic of Congo, also have high $c\overline{D\alpha}$ values (12–14.5). About half of these countries (88 out of 151) has a $c\overline{D\alpha}$ between 4 and 8, while only 6 have a $c\overline{D\alpha}$ lower than 3, including two countries with much cropland (>10 Mha), Kazakhstan and the USA. These values are considerably lower than the current total country-level diversity ($cD\gamma$). For instance, less than 5% of these countries have a $c\overline{D\alpha}$ greater than 12, but 40% have a $cD\gamma$ greater than 12 effective crops (supplementary table S1).

4.2. Potential and attainable diversity

The values for pD strongly depend on the suitability estimation method used (figures 4, 5 and supplementary figure S6). In contrast, the aD values derived from the two suitability methods are remarkably similar and considerably lower than pD (figures 4, 6 and supplementary figures S8 and S9). The differences in pD and aD due to the suitability estimation methods are described and discussed in supplementary section S.4.1.

Irrespective of the method used, aD is higher in the tropics than temperate regions, and outside the tropics, higher in coastal than continental regions. aD is highest in Sub Saharan Africa, southern India, some regions of Southeast Asia, eastern Brazil, northern South America, and the Caribbean. In subtropical and temperate areas, aD is high in East Asia, New Zealand's North Island, Chile, the US southeast and west coast, and parts of the Mediterranean region (e.g. Portugal and Italy). aD is very low in Kazakhstan, Mongolia, Russia, the Baltic States, Scandinavia, Canada, and the northern US (figure 6

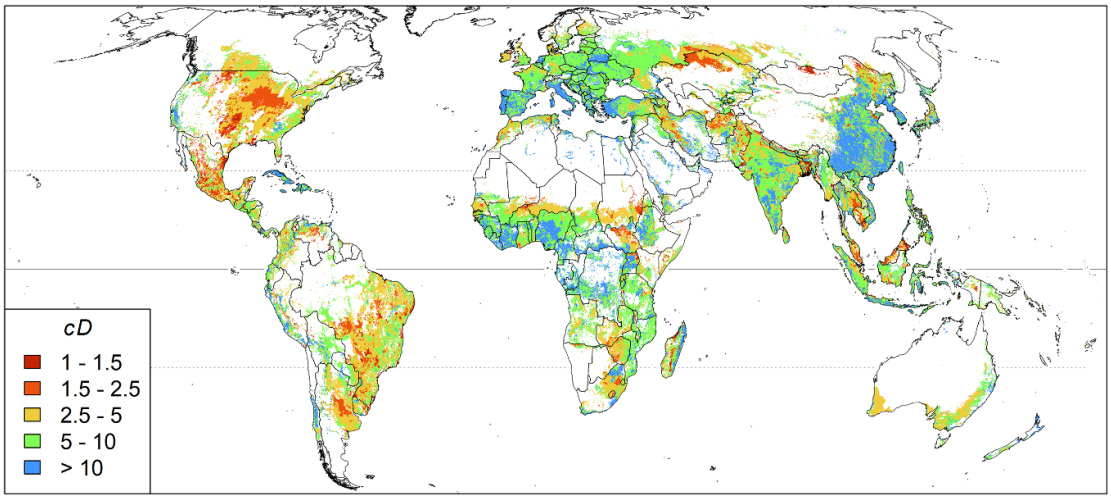

Figure 2. Global patterns of current local crop species diversity (*cD*) for 86 km^2 cells. White areas have less than 0.5% of cropland coverage.

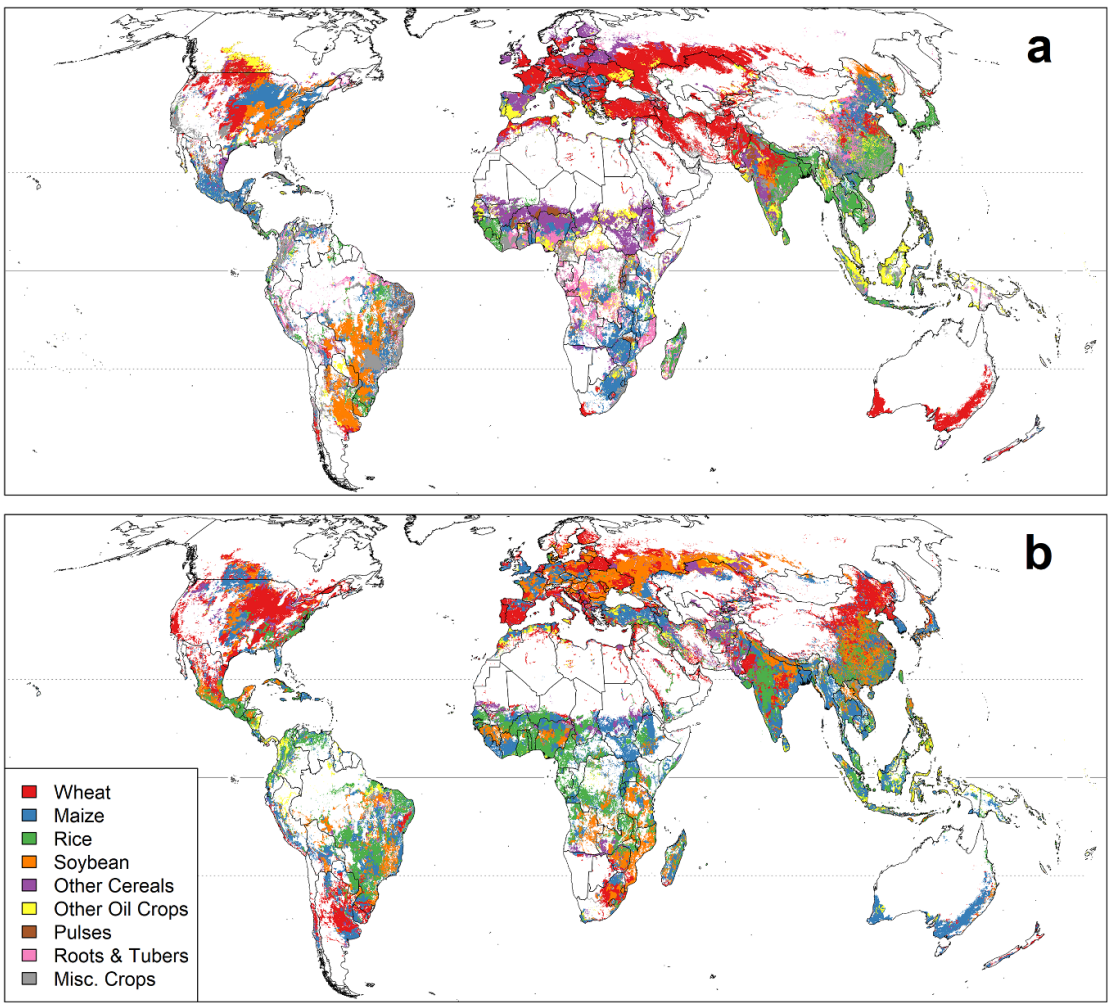

Figure 3. Most abundant (a) and most underutilized (b) crops for 86 km^2 grid cells. The most underutilized crop has the highest difference between attainable and actual crop area, where the attainable area is obtained after allocating current crop total areas to the most suitable cropland for each crop that is available considering the needs for all other crops. Miscellaneous crops include fiber crops, tree nuts, fruits, sugar crops, stimulants, vegetables and melons, and other crops. White areas have less than 0.5% of cropland coverage. The same color code was used in (a), (b).

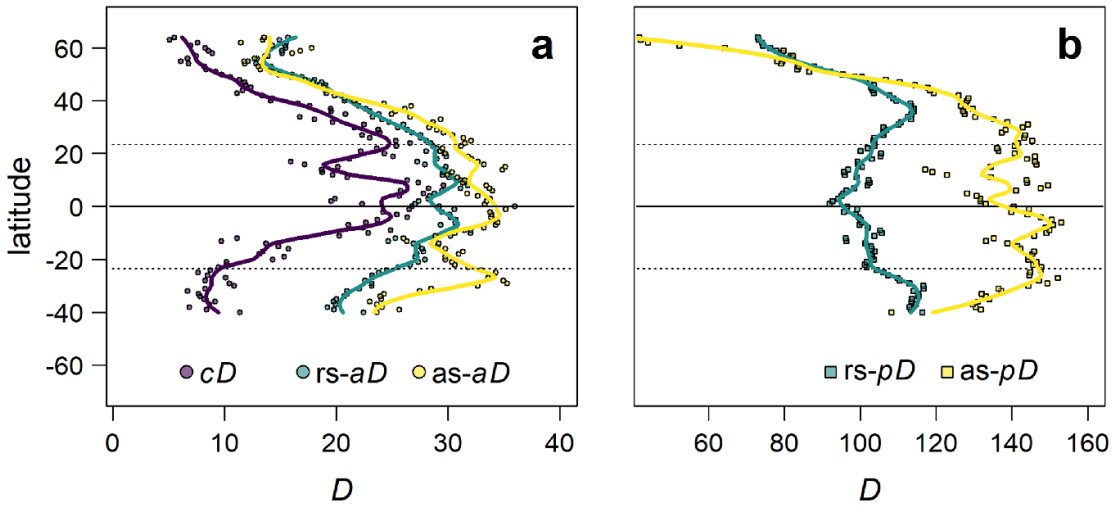

Figure 4. Crop species diversity levels by latitude. Each point represents the total crop species diversity in a band of 1° of latitude for different diversity levels: (a) current diversity (*cD*), relative-suitability-derived attainable diversity (rs-*aD*), and absolute-suitability-derived attainable diversity (as-*aD*); and (b) relative-suitability-derived potential diversity (rs-*pD*) and absolute-suitability-derived potential diversity (as-*pD*). Values for latitudinal bands with less than 100 000 ha of cropland were excluded. The horizontal dashed lines represent the tropics. The solid-colored lines are local regression lines.

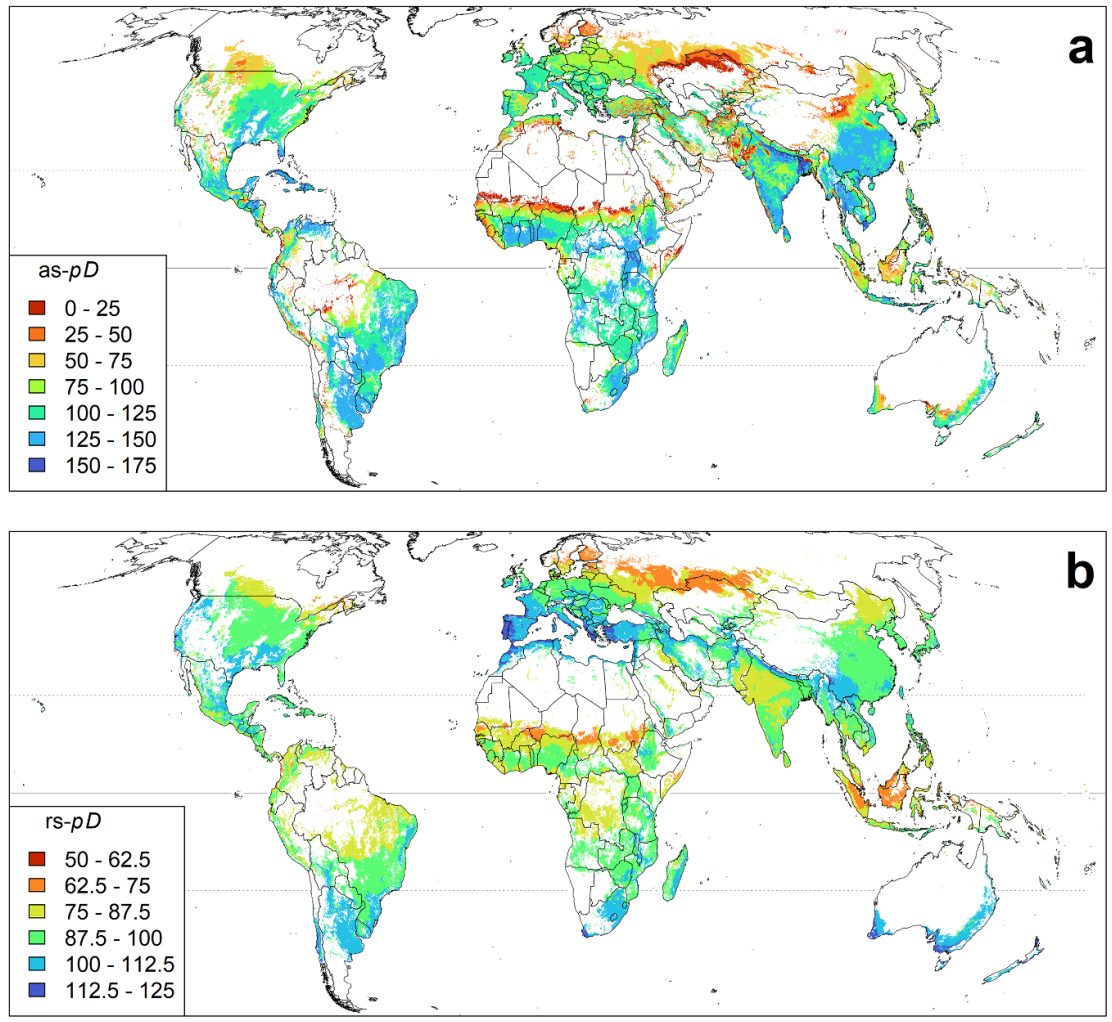

Figure 5. Global patterns of two estimates of potential crop species diversity: (a) the absolute-suitability-derived potential diversity (as-*pD*) and (b) the relative-suitability-derived potential diversity (rs-*pD*), for 86 km^2 cells. White areas have less than 0.5% of cropland coverage.

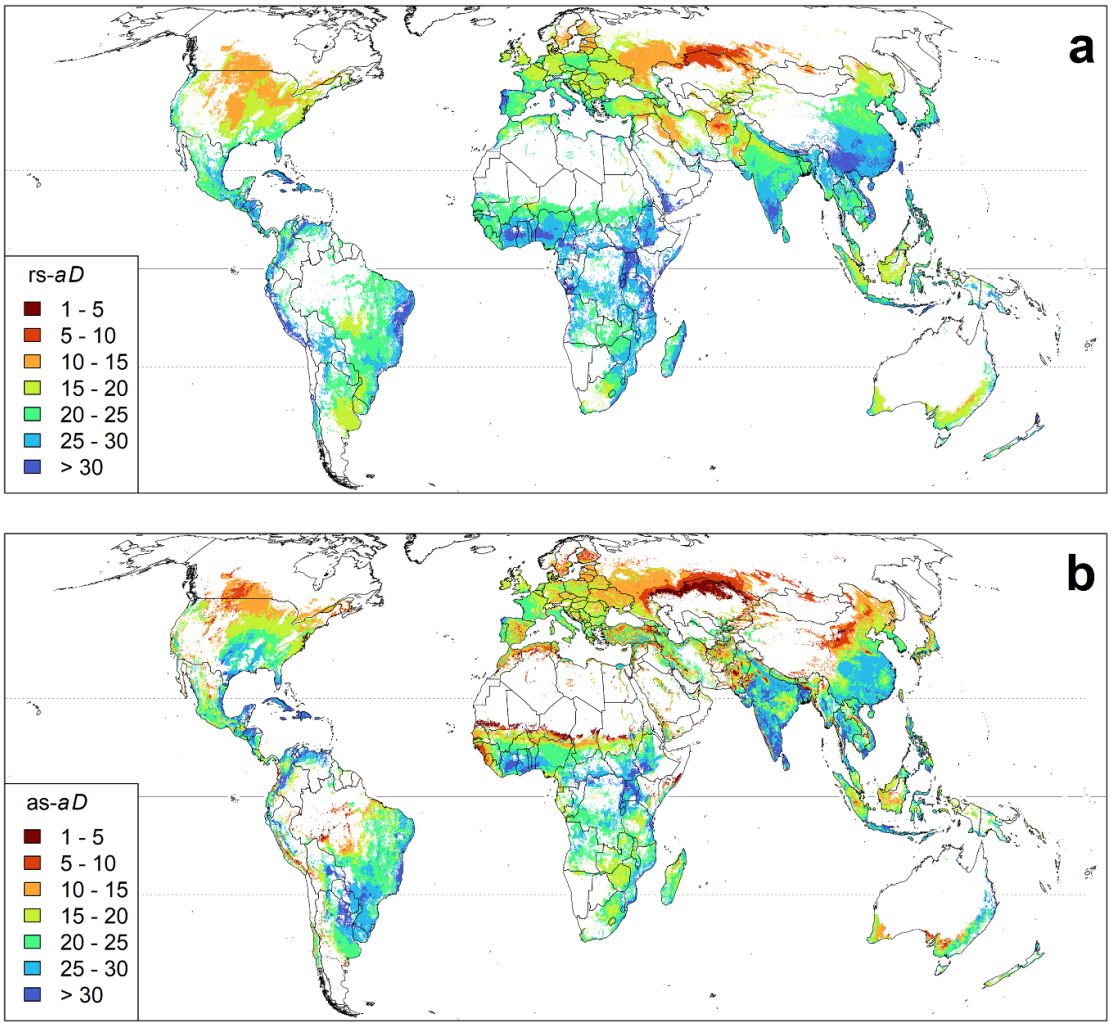

Figure 6. Global patterns of two estimates of attainable crop species diversity: (a) the absolute-suitability-derived attainable diversity (as-*aD*) and (b) the relative-suitability-derived attainable diversity (rs-*aD*), for 86 km^2 cells. White areas have less than 0.5% of cropland coverage.

and supplementary table S1). The annual average temperature strongly affects *aD*: it increases linearly from −10 °C until it reaches a plateau at about 20 °C to 25 °C and slightly decreases at higher temperatures (supplementary figure S10).

4.3. Global diversity gaps

Nearly 84% of the world's cropland has a D_g that is >50%; thus, it has less than half of the crop diversity that it would have if crops were planted to maximize diversity while considering their suitability and current food demand (figure 7). The D_g is especially high in the Americas (82% on average) except for the Andean region, the Caribbean Islands, the US west coast, and Canada. Africa (72%), Asia (71%), and Oceania (76%) also have large D_g values but with much spatial variability. The D_g is relatively small in Europe (56%), especially in the Mediterranean, Eastern Europe, and the Netherlands.

Cropland with a low *cD* tends to have a high D_g (supplementary figure S12). For instance, about 40% of the world's cropland has a *cD* lower than 5. Of this low *cD* cropland, 80% have a $D_g > 75\%$, while less than one percent have a $D_g < 50\%$. In contrast, virtually all the 10% most diverse cropland (*cD* > 12) has a $D_g < 60\%$, and 80% of it has a $D_g < 50\%$.

At the national level, Israel stands out for its low local-average D_g of 14%. Lebanon (30%) and only ten other countries with more than 0.1 Mha of cropland have an average D_g below 50%, while about half have an average D_g higher than 70% (supplementary table S1). However, at the country level, diversity gaps are significantly smaller when considering total diversity ($D\gamma$) instead of local diversity averages ($D\alpha$) (paired *t*-test, $P < 0.01$), illustrating the scale-dependency of diversity.

One strategy to reduce diversity gaps can be to increase the area of the 'most under-utilized' crops, that is, crops showing the highest differences between their attainable and actual crop proportion (figure 3(b) and supplementary figure S13). These crops are primarily major crops because the attainable area (proportion) in a grid cell tends to be greater for these crops than crops with little demand. For

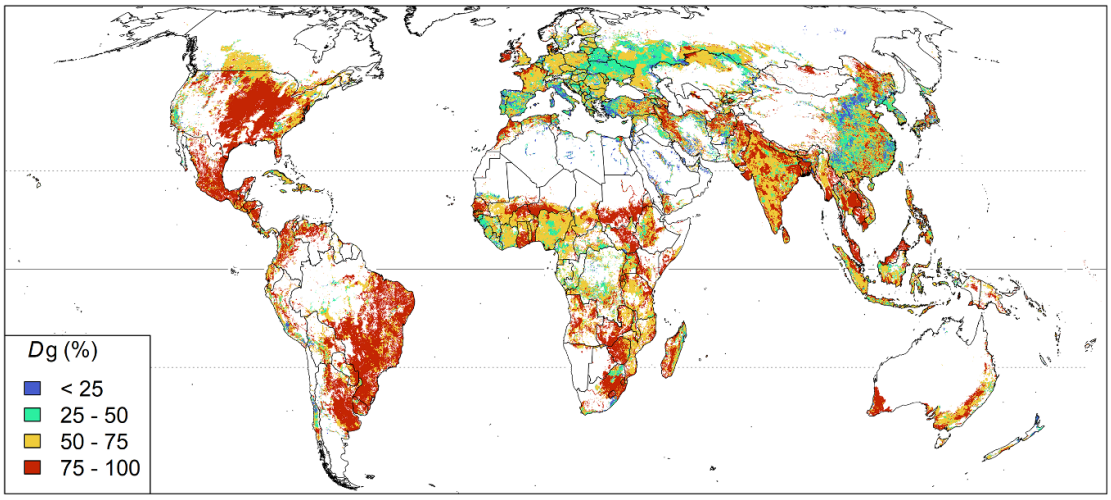

Figure 7. Global crop species diversity gaps. The D_g is the difference between the attainable diversity (average results from the two methods used) and current diversity, relative to the attainable diversity, for 86 km^2 cells. White areas have less than 0.5% of cropland coverage.

instance, the most abundant crop worldwide, wheat, is also one of the most under-utilized, particularly in the US Corn Belt, northeast China, and parts of Europe and Argentina.

5. Discussion

5.1. Specialization and diversification

We assessed global crop diversity gaps considering both ecological and demand constraints to attainable levels of crop diversity. Even when considering the world's heavy reliance on a few major crops for food supply (Cassman and Grassini 2020, Renard and Tilman 2021), our results show vast opportunities for crop diversification: crop species diversity could be doubled on five-sixths of the world's croplands if we only consider environmental constraints and total demand for crops.

There are various reasons why local specialization currently reduces crop diversity this much. While there can be economic benefits to some level of diversification at the farm level, such as risk reduction (Gaudin *et al* 2015), pest and weed pressure mitigation (Davis *et al* 2012), and soil fertility improvement (Tiemann *et al* 2015), these benefits are context-specific and may not be large enough to justify the increase in costs and complexity of managing additional crops (Roesch-McNally *et al* 2018). For example, the benefits of diversification may strongly depend on which crops are added to a cropping system, and further research could investigate opportunities for 'functional diversification'. If increasing farm-level crop diversity is too challenging, it may be possible to increase regional diversity by having different farms specializing in different crops. The effect of this diversification strategy would depend on how farm sizes and configuration shape the landscape (Sirami *et al* 2019), and while this might in some cases reduce transportation costs by decreasing the distance between production and consumption, there could also be a reduced benefit of economies of scale (for example, producing tomatoes near tomato processing plants) and other losses of efficiency associated with regional specialization. At the national level, the opportunity for diversification may be reduced due to policies to assure that a large part of the staple food is produced internally, as imports may be considered less reliable unless there is sufficient land available for new crops (Arsenault *et al* 2015).

When considering opportunities for increasing crop diversity, an important question is which crops should be grown more. For example, while some regions could increase the area with specialty crops, such increases might reduce crop diversity elsewhere if there is no increased demand for those crops. In contrast, given current crop-specific supplies and demands, the most effective strategy to increase crop diversity in a large area might be to reduce the proportion of the most dominant crops and plant more of the most under-utilized crop, which is often a suitable major crop not widely planted in that area. Furthermore, a drastic change in global demand, perhaps through changing diets, could affect attainable and actual diversity, but it is hard to imagine a diet not dominated by starch-producing crops such as wheat, maize, rice, and cassava.

While we focused on crop species diversity, other diversification strategies, such as rotations with cover crops, grassland-cropland integration, and agroforestry, should also be considered when seeking better ecosystem services provision through diversification, particularly in regions where the attainable crop diversity is low (Garrity *et al* 2010, Lemaire *et al* 2015). The best choice will depend on the magnitude of the constraints to diversity and the targeted services to be improved.

5.2. Constraints on crop diversity

Most studies on crop diversity do not consider the environmental constraints that might limit farmers' opportunities for diversification (Kremen and Miles 2012, Renard and Tilman 2019), and very little attention has been given to drivers of crop species diversity (Roesch-McNally *et al* 2018, Goslee 2020). Our analysis of environmental effects on attainable diversity can shed light on some important questions related to crop diversity (Wood 1998), especially the extent to which crop diversity can be increased (Cassman and Grassini 2020). Crop species diversity tends to be greater in tropical than in temperate areas and in coastal than in continental regions, and there is a clear limit to increasing crop diversity in cold environments. Therefore, it is not sensible to expect or call for similar levels of diversity across very different regions, and it cannot be assumed that all countries have the same diversity potential (Jones *et al* 2021), just as is the case with crop yield potential.

Furthermore, there is high spatial variability in current crop diversity that environmental models of attainable diversity cannot explain. This high spatial heterogeneity in diversity and diversity gaps could be related to factors affecting farmers' cropping decisions, such as spatial variation in market access, prices, risk, and policies. Understanding how these factors lead to specialization or limit diversification using spatially explicit models is needed to determine to which extent closing the diversity gap is economically feasible and identify policies that strongly affect diversity, particularly for regions with the highest diversity gaps (Socolar *et al* 2021). Moreover, while increasing diversity may be beneficial in some cases, closing diversity gaps might not always be necessary, such as in regions with high diversity and extremely high potential. Nevertheless, the diversity gap concept is helpful as it allows us to better contrast and compare crop diversity in different regions and investigate what shapes these patterns.

5.3. Diversity gaps

Diversity gaps are smaller in Europe and other areas dominated by relatively small family farms that tend to have higher crop diversity (Ricciardi *et al* 2021). This farm size-crop diversity inverse relationship might be associated with a higher proportion of minor crops in smaller farms (e.g. pulses, roots, tubers, and fruits) (Ricciardi *et al* 2018). Minor crops tend to be planted in more diverse cropping systems and are less likely to take most of the cropland of a region (Aramburu Merlos and Hijmans 2020). This association between minor crops and crop diversity might also explain the small diversity gaps in regions specializing in horticultural crops, such as the US west coast and the Netherlands. In Europe, relatively low D_g may be further supported by agricultural policies that promote diverse landscapes (Stoate *et al* 2009). Gaps are also relatively small in countries that rely less on international markets (Cuba, North Korea) and in places that face high transportation costs (Caribbean islands, desert oases), where most of the production is for local consumption. In contrast, diversity gaps are very high in the sparsely populated plains with a relatively recent agricultural expansion in the Americas (Graesser *et al* 2018). Farms in these regions have larger fields and focus on major crops for export in low diversity cropping systems (Aramburu Merlos and Hijmans 2020).

5.4. Assessing potential and attainable diversity

Diversity gaps can only be calculated after defining appropriate theoretical levels of diversity. Although our approach could be refined, it seems clear that attainable diversity (*aD*) is a much more robust and meaningful diversity benchmark than potential diversity (*pD*). *aD* not only accounts for total demand, making it insensitive to the omission of rare crops, but also it is less sensitive to changes in the suitability estimation method. *pD* estimates depend on how crop-specific suitability indices relate to each other between crops, whereas *aD* estimates only depend on the relative score of the cropland within each crop. However, there might be cases in which it is interesting to assess the potential diversity of a region. In such a case, the suitability estimation method should be carefully selected. Any suitability estimation method that depends on observed data is constrained by the current diversity level of the area of study and data availability for minor crops. Some examples include those methods that rely on crop distribution data (i.e. SDMs) and those that use observed diversity data to fit quantile regressions (Goslee 2020). Quantile regression methods are notably inadequate for potential diversity estimations because farmers generally do not aim at reaching the highest levels of diversity possible. Crop models are more suited for estimating potential diversity because current diversity levels do not affect them.

5.5. Local versus country-level diversity

Crop diversity in space depends on the area of the unit at which diversity estimations are made (Aramburu Merlos and Hijmans 2020). Much analysis of crop diversity and effects relies on national statistics (Khoury *et al* 2014, Mahaut *et al* 2021), in which the country-total diversity ($D\gamma$) is computed. However, most interest in diversification is related to expected effects at farm or landscape levels (Sirami *et al* 2019). Here we provide estimations of diversity and diversity gaps at a 9.26 × 9.26 km resolution (8575 ha), which allows us to compute local-average diversity ($\overline{D\alpha}$) for each country, which is consistently lower than $D\gamma$ and results in larger gaps. $D\alpha$ estimates are more appropriate for studying crop diversity's effects on agroecosystem services and processes, such

as pollination (Aizen *et al* 2019), associated biodiversity (Sirami *et al* 2019), and biological pests control (Tscharntke *et al* 2005). In addition, spatial crop diversity at this resolution is highly correlated with crop rotation diversity because different fields are in different stages of their crop rotation (Aramburu Merlos and Hijmans 2020). While there can be benefits of diversity at the national level (Renard and Tilman 2019), the national to local-average diversity ratio (i.e. $D\beta$), an indicator of regional heterogeneity, should also be considered when assessing diversity effects on the stability of food production (Mahaut *et al* 2021).

6. Conclusions

In this paper, we have contributed to a better understanding of spatial patterns of global crop diversity and opportunities for diversification. By defining theoretical levels of crop diversity, we created a way to compute diversity gaps, the difference between attainable diversity and actual diversity. The (relative) diversity gap is more informative than just the actual diversity because it accounts for environmental variation and limits set by demand. We have shown that even within the limits of the very skewed current levels of production for different crops, crop diversity could increase enormously. However, given the economic benefits of specialization, it remains an important question what the value of diversification could be in different regions and cropping systems, and, where more diversity is desirable, what incentives could be provided to achieve this.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://github.com/aramburumerlos/globcropdiv.

Conflict of interest

The authors declare no competing interest.

ORCID iDs

Fernando Aramburu Merlos https://orcid.org/0000-0003-0957-7482
Robert J Hijmans https://orcid.org/0000-0001-5872-2872

References

Aguiar S, Texeira M, Garibaldi L A and Jobbágy E G 2020 Global changes in crop diversity: trade rather than production enriches supply *Glob. Food Secur.* **26** 100385
Aizen M A *et al* 2019 Global agricultural productivity is threatened by increasing pollinator dependence without a parallel increase in crop diversification *Glob. Change Biol.* **25** 3516–27
Aramburu Merlos F and Hijmans R J 2020 The scale dependency of spatial crop species diversity and its relation to temporal diversity *Proc. Natl Acad. Sci. USA* **117** 26176–82
Arsenault J E, Hijmans R J and Brown K H 2015 Improving nutrition security through agriculture: an analytical framework based on national food balance sheets to estimate nutritional adequacy of food supplies *Food Secur.* **7** 693–707
Cassman K G and Grassini P 2020 A global perspective on sustainable intensification research *Nat. Sustain.* **3** 262–8
Davis A S, Hill J D, Chase C A, Johanns A M and Liebman M 2012 Increasing cropping system diversity balances productivity, profitability and environmental health *PLoS One* **7** e47149
Elith J and Leathwick J R 2009 Species distribution models: ecological explanation and prediction across space and time *Annu. Rev. Ecol. Evol. Syst.* **40** 677–97
FAO 2016 AQUASTAT website (Food and Agriculture Organization of the United Nations (FAO)) (available at: www.fao.org/nr/water/aquastat/irrigationmap/index10.stm)
Fick S E and Hijmans R J 2017 WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas *Int. J. Climatol.* **37** 4302–15
Garrity D P, Akinnifesi F K, Ajayi O C, Weldesemayat S G, Mowo J G, Kalinganire A, Larwanou M and Bayala J 2010 Evergreen agriculture: a robust approach to sustainable food security in Africa *Food Secur.* **2** 197–214
Gaudin A C, Tolhurst T N, Ker A P, Janovicek K, Tortora C, Martin R C and Deen W 2015 Increasing crop diversity mitigates weather variations and improves yield stability *PLoS One* **10** e0113261
Goslee S C 2020 Drivers of agricultural diversity in the contiguous United States *Front. Sustain. Food Syst.* **4** 75
Graesser J, Ramankutty N and Coomes O T 2018 Increasing expansion of large-scale crop production onto deforested land in sub-Andean South America *Environ. Res. Lett.* **13** 084021
Hengl T *et al* 2017 SoilGrids 250m: global gridded soil information based on machine learning *PLoS One* **12** e0169748
Hijmans R J and Graham C H 2006 The ability of climate envelope models to predict the effect of climate change on species distributions *Glob. Change Biol.* **12** 2272–81
Hijmans R 2021 Recocrop: implementation of the ecocrop model for estimating adaptation of plants to environments (available at: https://github.com/cropmodels/Recocrop/)
IFPRI 2019 Global spatially-disaggregated crop production statistics data for 2010 Version 2.0 (available at: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/PRFF8V)
Jackson N D, Konar M, Debaere P and Estes L 2019 Probabilistic global maps of crop-specific areas from 1961 to 2014 *Environ. Res. Lett.* **14** 094023
Jones S K, Estrada-Carmona N, Juventia S D, Dulloo M E, Laporte M-A, Villani C and Remans R 2021 Agrobiodiversity index scores show agrobiodiversity is underutilized in national food systems *Nat. Food* **2** 712–23
Jost L 2006 Entropy and diversity *Oikos* **113** 363e375
Jost L 2007 Partitioning diversity into independent alpha and beta components *Ecology* **88** 2427–39
Khoury C K, Bjorkman A D, Dempewolf H, Ramirez-Villegas J, Guarino L, Jarvis A, Rieseberg L H and Struik P C 2014 Increasing homogeneity in global food supplies and the implications for food security *Proc. Natl Acad. Sci. USA* **111** 4001–6
Kremen C and Miles A 2012 Ecosystem services in biologically diversified versus conventional farming systems: benefits, externalities, and trade-offs *Ecol. Soc.* **17** 40
Lemaire G, Gastal F, Franzluebbers A and Chabbi A 2015 Grassland–cropping rotations: an avenue for agricultural diversification to reconcile high production with environmental quality *Environ. Manage.* **56** 1065–77
Mahaut L, Violle C and Renard D 2021 Complementary mechanisms stabilize national food production *Sci. Rep.* **11** 4922

Monfreda C, Ramankutty N and Foley J A 2008 Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000 *Glob. Biogeochem. Cycles* **22** GB1022

R Core Team 2020 R: A language and environment for statistical computing (R Foundation for Statistical Computing)

Renard D and Tilman D 2019 National food production stabilized by crop diversity *Nature* **571** 257

Renard D and Tilman D 2021 Cultivate biodiversity to harvest food security and sustainability *Curr. Biol.* **31** R1154–8

Ricciardi V, Mehrabi Z, Wittman H, James D and Ramankutty N 2021 Higher yields and more biodiversity on smaller farms *Nat. Sustain.* **4** 651–7

Ricciardi V, Ramankutty N, Mehrabi Z, Jarvis L and Chookolingo B 2018 How much of the world's food do smallholders produce? *Glob. Food Secur.* **17** 64–72

Roesch-McNally G E, Arbuckle J G and Tyndall J C 2018 Barriers to implementing climate resilient agricultural strategies: the case of crop diversification in the US corn belt *Glob. Environ. Change* **48** 206–15

Šavrič B, Patterson T and Jenny B 2019 The equal Earth map projection *Int. J. Geogr. Inf. Sci.* **33** 454–65

Sirami C, Gross N, Baillod A B, Bertrand C, Carrié R, Hass A, Henckel L, Miguet P, Vuillot C and Alignier A 2019 Increasing crop heterogeneity enhances multitrophic diversity across agricultural regions *Proc. Natl Acad. Sci. USA* **116** 16442–7

Socolar Y, Goldstein B R, de Valpine P and Bowles T M 2021 Biophysical and policy factors predict simplified crop rotations in the US Midwest *Environ. Res. Lett.* **16** 054045

Stoate C, Báldi A, Beja P, Boatman N D, Herzon I, van Doorn A, de Snoo G R, Rakosy L and Ramwell C 2009 Ecological impacts of early 21st century agricultural change in Europe—a review *J. Environ. Manage.* **91** 22–46

Tiemann L K, Grandy A S, Atkinson E E, Marin-Spiotta E and McDaniel M D 2015 Crop rotational diversity enhances belowground communities and functions in an agroecosystem *Ecol. Lett.* **18** 761–71

Tilman D and Clark M 2014 Global diets link environmental sustainability and human health *Nature* **515** 518–22

Tscharntke T, Klein A M, Kruess A, Steffan-Dewenter I and Thies C 2005 Landscape perspectives on agricultural intensification and biodiversity–ecosystem service management *Ecol. Lett.* **8** 857–74

Tuomisto H 2010 A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity *Ecography* **33** 2–22

van Ittersum M K, Cassman K G, Grassini P, Wolf J, Tittonell P and Hochman Z 2013 Yield gap analysis with local to global relevance—a review *Field Crops Res.* **143** 4–17

van Ittersum M K and Rabbinge R 1997 Concepts in production ecology for analysis and quantification of agricultural input-output combinations *Field Crops Res.* **52** 197–208

Wood D 1998 Ecological principles in agricultural policy: but which principles? *Food Policy* **23** 371–81

You L, Wood S, Wood-Sichra U and Wu W 2014 Generating global crop distribution maps: from census to grid *Agric. Syst.* **127** 53–60