Matches and mismatches between the global distribution of major food crops and climate suitability.

Lucie Mahaut, Samuel Pironon, Jean-Yves Barnagaud, François Bretagnolle, Colin K. Khoury, Zia Mehrabi, Ruben Milla, Charlotte Phillips, Loren H. Rieseberg, Cyrille Violle, et al.

Matches and mismatches between the global distribution of major food crops and climate suitability

Lucie Mahaut[1], Samuel Pironon[2,3], Jean-Yves Barnagaud[1], François Bretagnolle[4], Colin K. Khoury[4,6], Zia Mehrabi[7], Ruben Milla[8], Charlotte Phillips[2], Loren H. Rieseberg[9], Cyrille Violle[1*], Delphine Renard[1*]

[1]CEFE, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France.

[2]Royal Botanic Gardens, Kew, Richmond, UK.

[3] UN Environment Programme World Conservation Monitoring Center (UNEP-WCMC), Cambridge, UK.

[4]Université Bourgogne Franche Comte, Biogeosciences, UMR 6282, Centre National de la Recherche Scientifique (CNRS), Dijon, France.

[5] International Center for Tropical Agriculture (CIAT), Km 17, Recta Cali-Palmira, Apartado Aéreo 6713, 763537 Cali, Colombia

[6]San Diego Botanic Garden, 230 Quail Gardens Dr, Encinitas, CA 92024 USA

[7] Institute for Resources Environment and Sustainability, School of Public Policy and Global Affairs, University of British Columbia, Canada.

[8] Universidad Rey Juan Carlos, Escuela Superior de Ciencias Experimentales y Tecnología, Mostoles, Spain.

[9] Department of Botany and Biodiversity Research Centre, University of British Columbia, Vancouver, BC, V6R 2A5, Canada.

*****Both considered as co-last authors

Abstract

Over the course of history, humans have moved crops from their regions of origin to new locations across the world. The social, cultural, and economic drivers of these movements have generated differences not only between current distributions of crops and their climatic origins, but also between crop distributions and climate suitability for their production. Although these mismatches are particularly important to inform agricultural strategies on climate change adaptation, they have, to date, not been quantified consistently at the global level. Here, we show that the relationships between the distributions of twelve major food crops and climate suitability for their yields display strong variation globally. After investigating the role of biophysical, socio-economic and historical factors, we report that high-income world regions display a better match between crop distribution and climate suitability. In addition, although crops are farmed predominantly in the same climatic range as their wild progenitors, climate suitability is not necessarily higher there, a pattern that reflects the legacy of domestication history on current crop distribution. Our results reveal how far the global distribution of major crops diverges from their climatic optima and call for greater consideration of the multiple dimensions of the crop socio-ecological niche in climate change adaptive strategies.

Introduction

Ensuring food security while adapting to climate change and preserving the environment is a central challenge for humanity [1,2]. Adaptive strategies in agriculture notably include the migration of cropland to areas that will be more suitable for agricultural production according to different scenarios of climate change (e.g. [3–6] and the selection of new species and varieties better adapted to future climatic regimes (e.g., [7–9]). These strategies should lead to a better adjustment of the distribution and identity of crops to climatic conditions which could in turn reduce the need for agricultural inputs [10] and improve the resilience of agricultural systems to climate change [11]. However, apart from limited evidence of crop presence in areas of low climatic suitability (e.g., [12]), the relationships between the current distribution of crops and the suitability of climate for their growth have been largely overlooked, leaving unknown the potential for improving matches between crop distribution and climate.

Recent shifts in where crops are grown at both global and regional scales suggest that human societies modulate the distribution of cropland areas to buffer the adverse effects of climate change on crop yields [13,14]. Conversely, environmental and agricultural factors that directly affect crop yields, such as soil characteristics (e.g., [3]), crop pests (e.g. [15]) or the use of agricultural inputs (e.g. irrigation and fertilisation, [16,17]) can decouple the distribution of crops from the suitability of climate for their growth. The fact that farmers adapt the crops they grow according to complex socioeconomic factors can also lead to strong inconsistencies in the distribution of crops with respect to climate [18]. Deciphering the role of multiple biophysical, agricultural and socio-economic factors in shaping the relationships between crop distribution and climate suitability is therefore a crucial step in assessing the potential levers that human societies have to reduce crop distribution-climate suitability mismatches.

In addition, crops have undergone substantial geographical expansion over millennia, from their respective centres of origin to the rest of the world [19]. Such historical changes in the geographical distributions of crops are well documented globally (e.g. [20]). However, the extent to which climatic conditions in areas where crops have been introduced (i.e. the introduced range) differs from that in the crop's centres of origin (i.e. the native range) remains largely unknown. This is particularly problematic given that sites of crop origin host much of the associated agrobiodiversity, including numerous crop wild progenitors and relatives as well as traditional farmer varieties [21,22]. As such, these centres are perceived as reservoirs from which to draw new genetic resources to improve crop resilience and adaptation to future climatic conditions [23,24]. However, if historical globalization, together with innovations in agricultural practices, have shifted the climatic niche of crops (i.e. the range of climates where a crop is grown) from the one of their wild progenitors, domesticated plants may no longer share the same climatic optima as their wild progenitors.

In this study, we use a global dataset of the geographic distribution of crop areas and yields [25] to characterise the current climatic niche of twelve major food crops and investigate how the distribution of their harvested areas co-vary with climate suitability for their yields (Figure 1). Using multiple regressions, we quantify the marginal effects of climate on crop yields while controlling for the role of multiple soil, topography, agricultural inputs and socio-economic factors to determine climate suitability for the yield of each crop. Then, we build a mismatch index to analyse the relationship between harvested areas and climate suitability and to examine the role of non-climatic factors in producing these (mis)matches. Finally, we quantify shifts between the native and the introduced climatic range of crops and test whether harvested areas, climate suitability and crop areas-climate suitability mismatch differ across the two ranges.

Material and Methods

Crop data set

We tested the relationship between the distribution of harvested areas and climate suitability for barley, cassava, groundnut, maize, potato, rapeseed, rice, sorghum, soybean, sugarbeet, sunflower and wheat (Supplementary Table S1). The 12 crops were chosen because they are widely cultivated worldwide, providing more than 70% of calorie intake globally [16], and because agronomic and wild progenitor data are available. For each of these crops, the global distribution of harvested area and yield was extracted from public data sources that have been created by combining national-, state-, and county-level census statistics with a global data set of croplands for the 1997-2003 period at a 5 arc-minute (~10km) resolution [25]. We controlled for the influence of the availability of agricultural land [26] by computing the fraction of available cropland allocated to each crop's growth in each grid cell. For each crop, harvested areas is therefore expressed in terms of percentage of available cropland.

Crop progenitors data set

We used native occurrences of each crop´s wild progenitors to delineate the climatic range of the centre of origin of each crop (i.e. native climatic range). We identified 23 progenitors of the 12 crops using published literature (Supplementary Table S1; [27]). Wheat includes *Triticum aestivum* and *Triticum durum*, which are the two main wheat species cultivated worldwide. Therefore, we considered the progenitors of both wheats. There are also two species of cultivated rice, *Oryza sativa* (Asian rice) and *Oryza glaberrima* (African rice). However, because *O. sativa* is by far the main cultivated species of rice, we only considered the wild progenitor of the Asian rice in our analysis. We extracted progenitor occurrence records from five global and two regional databases. We used the BIEN and rgbif packages in R to download global occurrence records from the Botanical Information and Ecology

Network [28] and the Global Biodiversity Information Facility databases, respectively. In addition, we extracted global occurrences of crop wild progenitors from "A global database for the distributions of crop wild relatives", the BioTIME database [29] and GENESYS (https://www.genesys-pgr.org/). We downloaded further regional occurrences using the RAINBIO database, which contains records for Sub-Saharan African vascular plants [30] and speciesLink, a national database for plant and animal occurrences in Brazil (https://splink.cria.org.br/). Because occurrence data are often inaccurate [31], we removed records that were documented before 1950. Further, we used the *CoordinateCleaner* package in R [32] to remove occurrence records found within a 1 km radius of country and capital centroids, with equal longitude and latitude coordinates, or assigned to institutional locations such as botanical gardens, herbaria or the GBIF Headquarters. We also removed records with high coordinate uncertainty (over 10 km), cultivated records (e.g. breeding/research material, advanced/improved cultivar, GMO and those found in markets or shops, institutes/research stations and genebanks and from seed companies), and records located in the sea/oceans. For progenitors with the same species name as their crop, we only extracted records confirmed as wild and ignored any record with an unknown cultivation status. We also removed records outside of the wild species native range to ensure no introduced or cultivated records were included. Native ranges were identified at administrative levels according to the USDA Agricultural Research Service, Germplasm Resources Information Network. After filtering, 8715 occurrence points remained for 23 wild progenitor species of the 12 crops (Supplementary Table S1).

Climatic, soil, topographic, agronomic and socio-economic data set

We collected eight bioclimatic variables from the CHELSA database [33] at a 30 arc-seconds (~1 kilometre) resolution that corresponded to the mean of annual data for the years 1979–2013 (Supplementary Table S2). This includes mean annual temperature (BIO1), temperature seasonality (BIO4), maximum temperature of the warmest month (BIO5), minimum temperature of the coldest month (BIO6), annual sum of precipitation (BIO12), precipitation of the wettest (BIO13) and driest (BIO14) months, and precipitation seasonality (BIO15). These eight variables were selected to capture mean, seasonal and extreme (potentially limiting) temperature and precipitation regimes.

We estimated the role of agricultural factors by gathering the amount of nutrients applied on individual crops (hereafter fertilisation in ton/ha) and the percentage of land area equipped for irrigation (hereafter irrigation). We used the Nutrient-Application-For-Major-Crops maps [17] that compiled national and subnational data of mineral fertilizers, manure, and atmospheric deposition with detailed maps of crop areas to produce crop-specific nutrient application rate (tons per ha) at a 5 arc-minute resolution (Supplementary Table S2). These data represent mean annual values for the 1997-2003 period. We downloaded the Global Map of Irrigation Areas from the FAO AquaStat portal that represents the percentage of area equipped for irrigation circa 2005 as a 5 arc-minute raster (Supplementary Table S2). Soil pH, organic content and maximum water holding capacity were used to define soil conditions. These data were collected from the ISRIC World Soil information portal at a 250m resolution (Supplementary Table S2). Topography was characterized by terrain slope from the EarthEnv portal at a 1km resolution (Supplementary Table S2). Finally, socio-economic factors were approximated by the Human Development Index (HDI) and Gross Domestic Product (GDP). HDI is a composite index of average achievement in three key dimensions of human development (a long and healthy life, being knowledgeable and having a decent standard of living [34]). We downloaded these data from the Gridded global datasets for Gross Domestic Product and Human Development Index over 1990-2015 at a 5 arc-min resolution (Supplementary Table S2). For each index, we computed the mean of 1990-2015 annual data. Climate, soil and topography layers were all upscaled to a 5 arc-minutes resolution using the *aggregate* function from the *raster* package [35]in the R software to match the resolution of the crop dataset.

Data analyses

We used a PCA to characterize the main axes of variation of the climatic conditions of the terrestrial surfaces occupied by cropland globally (step 1 in Figure 1). We used the global distribution of cropland areas at a 5 arc-minute resolution [26] and selected pixels with cropland cover >0. We extracted the values of the bioclimatic variables previously described for each of these pixels and summarized their variation according to the two first principal component axes (Dim. 1 and Dim. 2, respectively). Dim.1 explained 51.1% of the variance and principally reflected the variation in temperature with increasing values corresponding to increasing mean annual temperatures and decreasing seasonality in temperature (Supplementary Figure 1). Dim.2 explained 26.8% of the variance. It was related to precipitation, with increasing values corresponding to increasing seasonality in precipitation and decreasing total annual precipitation (Supplementary Figure S1). The PCA was performed using the *factoMineR* R package (Husson et al., 2020). Hereafter we use "climatic niche" to refer to the portion of the climatic space made by Dim.1 and Dim.2 that is occupied by a crop (i.e. the range of climatic conditions where a crop is found). Dim.1 and Dim.2 are identical among crops and thus comparable.

We then projected harvested areas (in % of available cropland) and yields within the climatic niche of each crop (step 2 in Figure 1). We used the function *rasterize* from the *raster* R library [35] to build a 150 x 150 pixels grid defined by Dim.1 and Dim.2. For each pixel of this climatic grid, we computed the average values of the fraction of cropland allocated to each crop and average crop yield. This grid size was chosen so that each of these 150x150 pixels corresponded on average to 25 pixels of the geographical distribution of crops (and thus to 25 measures of yields and harvested areas). We smoothed harvested areas values using a loess regression to identify hot- and cold-spots of cultivation (Supplementary Figure S2). Furthermore, we assessed climate suitability by computing the marginal effects of climate (i.e. Dim.1 and Dim.2) on the yields of each crop while accounting for soil (i.e. soil organic content, soil pH, soil water capacity), topography (i.e. slope), agricultural inputs (i.e. fertilisation rate and irrigation) and socio-economics (i.e. gross domestic product and human development index) factors. For each 150 x 150 pixel of the climatic grid and each crop, we calculated the average values of these non-climatic factors. We used linear models and the function *predict* from the *stats* R library to quantify the marginal effects of Dim. 1 and Dim. 2 (as second order polynomials) while controlling for the effects of non-climatic factors. Climate suitability thus corresponds to the expected yields given the climate, all other things being equal.

Third, we quantified the extent to which the distribution of harvested areas matched climate suitability for the yields of each crop (step 3 in Figure 1). We calculated Spearman correlation between smoothed values of harvested areas and climate suitability. For each crop, we randomly sampled 1,000 pixels of the 150 x 150 pixels climatic grid, computed Spearman correlation index and repeated the process 1,000 times to calculate 95% confidence intervals of the correlation index.

Then, we analysed variation in the relationships between harvested areas and climate suitability within the climatic niche of each crop by computing a local mismatch index (step 4 in Figure 1). We computed for each crop and for each pixel of the climatic niche the centiles of the smoothed values of harvested areas and the centiles of climate suitability. We calculated the log-ratio between climate suitability centiles and crop areas centiles. The resulting mismatch index varied between log(0.01) (climate suitability value belonging to the 1^{st} centile/harvested areas value belonging to the 100^{th} centile) and log(100) (climate suitability value belonging to the 100th centile/ harvested areas value belonging to the 1st centile). When equal to zero, this index indicates a balanced situation (i.e. absence of mismatch) between harvested areas and climate suitability. Positive values describe situations where climate suitability is disproportionately large with respect to the portion of cropland allocated to a crop while negative values indicate situations where crop harvested areas are disproportionately large with respect to climate suitability.

We examined the role of agricultural inputs, soil conditions, topography and socio-economic factors in generating matches and mismatches between harvested areas and climate suitability. We used linear models to quantify the effects of each of these factors on the absolute values of the mismatch index of each crop. Using the absolute value of the crop area-climate suitability mismatch index allowed differentiating factors that are associated with a low mismatch from the ones associated with a high mismatch, whether the mismatch was in the direction of harvested areas or climate suitability. All covariates were scaled before analysis so that estimated coefficients are comparable.

Finally, we compared the climatic conditions in areas where crops have been introduced (i.e. introduced range) to those where the wild progenitors live (i.e. native range) to investigate how agricultural expansion could affect the relationships between current crop areas and climate suitability. We defined the native climatic range of each crop by drawing convex polygons around occurrence points of its wild progenitors (see step 2 in Figure 1; [21]). Because species occurrence data is prone to geo-referencing errors and because the use of convex polygons is sensitive to outliers, we removed crop wild relative occurrence points with Mahalanobis distance to the centroid of the native climatic range ≥ 10. The introduced climatic range of a crop corresponded to the portion of its climatic niche occupied by the crop but outside its native range. For each crop, we compared mismatch index, crop distribution and climate suitability between the two climatic ranges through an analysis of variance. All analyses were conducted using R version 4.0.3.

Results

Relationships between harvested areas and climate suitability

The models used to quantify climate suitability explained between 35% (groundnut) and 73% (maize) of the variance of crop yields (Supplementary Table S3). We found significant positive correlations between the distribution of harvested areas and climate suitability for eight of the 12 major crops under study (barley, groundnut, maize, rapeseed, soybean, sugarbeet, sunflower and wheat) (Figure 2). The other four crops - cassava, potato, rice and sorghum - displayed significant, negative harvested areas-climate suitability correlations (Figure 2). The magnitude of positive correlations varied between 0.2 (maize) and 0.9 (barley) while negative correlations varied between -0.25 (sorghum) and -0.5 (cassava) (Figure 2).

Local (mis)matches between harvested areas and climate suitability

The analysis of the mismatch index revealed strong variation in the relationships between harvested areas and climate suitability for crop yields (Figure 3). This index allowed the identification of climatic zones where crop areas (in % of available cropland) were disproportionately large with respect to climate suitability (blue zones in Figure 3), zones where climate suitability was disproportionately large with respect to harvested areas (red zones in Figure 3) and zones with balanced harvested areas and climate suitability (beige zones in Figure 3). Imbalanced zones were particularly marked for cassava, potato, rice and sorghum (i.e. crops displaying negative harvested areas - climate suitability correlations). The harvested areas of cassava, rice and sorghum was disproportionally large with respect to climate suitability in the warmest climates with strong precipitation seasonality while zones with lower harvested areas with respect to climate suitability were concentrated in colder climates with strong temperature seasonality (Figure 3). The mismatch index of barley (i.e. the crop showing the strongest positive correlation between harvested areas and climate suitability) showed the lowest variation, situations with balanced harvested areas and climate suitability predominating (Figure 3).

Drivers of harvested areas - climate suitability (mis)matches

Overall, soil conditions, topography, agricultural inputs and socio-economic factors together explained between 16% (groundnut) and 43% (wheat) of the variation in the absolute value of the mismatch index (Figure 4). For all crops except groundnut, cells with higher Human Development Index (HDI) and/or Gross Domestic Product (GDP) showed a higher match between the amount of cropland area devoted to a crop and climate suitability for its yield (i.e. negative effect on the mismatch index; Figure 4 & Supplementary Table S3). Soil conditions and topography were other important drivers of the mismatches but their effects varied among crops (Figure 4 & Supplementary Table S3). For all crops except cassava and sugarbeet, cells with higher soil organic content displayed greater mismatches between harvested areas and climate suitability (Figure 4 & Supplementary Table S3). Higher soil pH was associated to higher mismatches for barley, groundnut, maize, potato, rape, rice, sorghum, soybean and sugarbeet and wheat and to lower mismatches for cassava and sunflower (Figure 4 & Supplementary Table S3). Compared to these biophysical factors, agricultural inputs exerted lower influences on harvested areas - climate suitability mismatches (Figure 4 & Supplementary Table S3). Increasing fertilisation had negative effects on the mismatch index for barley, groundnut, maize, rapeseed, sorghum and wheat (Figure 4 & Supplementary Table S3). Similarly, increasing irrigation had negative effects on the mismatch index for barley, cassava, groundnut, rape, rice, soybean and wheat (Figure 4 & Supplementary Table S3).

Native vs introduced climatic ranges

Native and introduced climatic ranges always fully overlapped, with native ranges corresponding approximately to the centre of the climatic niche of each crop (Figure 3). The average mismatch between harvested areas and climate suitability in the native climatic ranges was negative for all crops except barley, indicating relatively large harvested areas with respect to climate suitability for crop yields (Figure 5). By contrast, the average mismatches in the introduced climatic ranges was null or slightly positive, with the highest values observed for rice and sorghum (Figure 5). In addition, the harvested areas of all crops were larger in native climatic ranges (Figure 5 & Supplementary Figure 2) while the average climate suitability for crop yields was worse in introduced ranges (Figure 5 & Supplementary Figure 3).

Discussion

Here, we report for the first time the relationship between the global distribution of twelve major food crops and climate suitability for their yields. We find that the amount of available cropland allocated to a crop and climate suitability for its yields are positively correlated for eight of these crops, confirming that climate is a determinant element in the choice of cultivation areas [13,14]. However, we report negative correlations between harvested areas and climate suitability for cassava, potato, rice and sorghum. We also reveal strong variation in the relationships between harvested areas and climate suitability for all crops, even for those crops that display positive harvested areas-climate suitability correlations. Together, these results clearly indicate that the global distribution of crop areas is not optimised for climate and that several other factors limit this adjustment.

Our results demonstrate the importance of socio-economic factors in regulating the relationships between the distribution of crop harvested areas and climate suitability for yields. Globally, regions with high gross domestic product and high human development index display better adjustments of crop distributions to climate. The richest regions of the world also display the lowest differences between observed yields and those potentially attainable (so-called “yield-gap”, [17,36]). These regions might therefore be best able to optimize the distribution of crop areas in relation to climate and environment, potentially through strong investments in agricultural policies, centralized planning and greater regionalization of production areas in these countries [37]. In addition, we report negative correlations between harvested areas and climate suitability for the yields of three crops that are all commonly found across smallholder and subsistence food systems, namely cassava, sorghum and rice. Cassava is notably a key crop for food security across the African continent and represents a lifeline when other crops fail due to droughts or pest outbreaks, especially during famine period [38]. Sorghum is also staple crops of smallholder farmers in sub-Saharan Africa [12,39], and almost 30% of rice cultivation in South Asia takes place in low-input farming systems [40]. These results suggest that poorer regions may grow staple crops that still produce food under harsh climatic conditions that deviate from crop climatic optima, so that harvested areas are disproportionately large with respect to climate suitability.

Soil conditions and farming practices are other important drivers of the relationships between harvested areas and climate suitability, which can either enhance or limit the adjustment of crop distribution to climate. Soil conditions are indeed important drivers of crop yield and modulate crop responses to extreme climatic events such as drought [41]. However, a panel of farming practices can modify edaphic conditions, including liming, mulching and the addition of organic matter [42]. These practices can therefore enhance the potential of crop cultivation in zones with suitable climate but unsuitable soil characteristics. Conversely, irrigation and fertilisation can be used to overcome low climate suitability by allowing crops to be cultivated beyond their biophysical limits. Accordingly, the use of agricultural inputs can generate mismatches between harvested areas and climate suitability. This is notably the case for potato (i.e. the fourth crop that displays negative harvested areas - climate suitability correlation) for which we found positive effects of both irrigation and fertilisation on the mismatch index (Figure 4). However, we also report numerous cases where agricultural inputs have negative effects on the mismatch index, indicating higher rates of fertilization and irrigation where climate suitability and harvested areas are in balance. This result can reflect the fact that agricultural inputs are mostly used to increase the productive capacity of areas that are already (climatically) suitable for crop yield [16,17].

Other factors, not directly considered in this study, can further explain the decoupling between the distribution of harvested areas and climate suitability for crop yields. The fact that some crops share the same climatic optima (e.g. wheat and barley, Supplementary Figure 3), leading to competition for space, can for example increase the mismatches between harvested areas and climate suitability. These mismatches may also arise because local food systems, and especially smallholder ones, typically cultivate a diversity of crops to serve multiple purposes (livelihood, subsistence, pest control [43]). Management factors such as multiple cropping systems [44] and year-to-year crop rotations [45] can further create patterns of harvested areas that deviate from climate suitability. Interspecific interactions, and notably the presence of natural enemies (pests and pathogens), may be another important driver of these mismatches since farmers may choose to reduce the area allocated to a crop that is exposed too intensely to pests. Although there is no globally available dataset for pests across many crops, we were able to test this hypothesis for sunflower using published occurrence records of *Sclerotinia sclerotiorum* [46], a major sunflower pathogenic fungus worldwide (see Supplementary Figure S4 for details). We found that harvested areas were disproportionately large compared climate suitability in presence of the pest (Supplementary Figure S4). This result shows that the presence of *S. sclerotiorum* and the associated yield damage can hardly be separated from the effects of climate at this scale, probably because the pest share the same climatic niche as its host [46]. Nevertheless, these results, if true for other crops and pests, would suggest a strong ecological anchoring of the harvested areas - climate suitability mismatches.

Finally, our results bring novel insights to understand how the geographical expansion of crops from their origin centres to the rest of the world has affected their climatic niche. Despite the global homogenization of cropping patterns [20,47], we show that major crops are still farmed predominantly in their native climatic range. Such differences in the distribution of harvested areas between native and introduced climatic ranges can highlight the biological constraints of crop species to wide climate adaptation. However, we find that native climatic ranges do not necessarily represent the highest climate suitability for crop yields. This might reflect the fact that farming practices and plant breeding has modified the climatic requirements of crops with regard to their wild progenitors [48], so that crops and their wild progenitors no longer share the same climatic optimums. These findings have major implication for agricultural science and breeding programs that increasingly look at crop wild relative as promising reservoir of genetic diversity to enhance crop adaptation to heat or drought (e.g. [21,24,49,50]. Accounting for differences in climate suitability between crops and their wild progenitors will be a critical issue for the development of novel crop varieties from these wild species.

Before concluding, it is important to mention some limitations of our study, which are mainly related to the nature of the datasets available for this work. First, we cannot exclude that our approach might overestimate the mismatches observed between crop’s harvested areas and climate suitability. This is notably because we cannot account for intra-specific (e.g., commercial varieties, landraces) variations in climatic ranges within each crop species. However, intra-specific variations are expected to play a lesser role at global scale, compared to a finer spatial scale[51]. Second, conclusions about the role of irrigation might be limited since the nature of the data (i.e. % of land equipped for irrigation) does not depict the actual irrigation of each crop. Finally, the agricultural data we used may appear a bit dated, as it presents the average, global distribution of crop areas and yields for the 2000s [25]. It also neglects year-to-year variability in the global distribution of crop areas and yields. Although temporal changes in the global distribution of crop harvested areas can mitigate the negative effects of climate on agricultural production [13], the extent to which these dynamics do affect the magnitude of the harvested areas-climate suitability mismatch remains to be quantified. Another promising extension of this work will be to consider the temporal variability of crop yields, yield stability being key for the adaptation of global agriculture to climate change [52].

To conclude, many strategies for coping with climate change in agriculture implicitly aim to improve the match between the distribution of crop areas and climate. Accordingly, studies have focused on the climatic niche of crops to predict future patterns of crop distribution and to identify crops and wild relatives best adapted to future conditions. However, the fact that multiple ecological, agricultural, socio-economic and historical factors govern the choice of farmers to cultivate a particular species strongly limits the global adjustment of crop distribution to climate, today. A main challenge for the coming years will therefore be to comprehensively integrate these dimensions to define the socio-ecological niche of crops [53] so as to design efficient strategies that can sustainably deal with climate change in agriculture. Extrapolating the socio-ecological niche of crops to future climate conditions will also provide crucial information for estimating the socio-economic costs that will need to be invested in order to improve the match between crops and future climate.

References

1. Godfray HCJ *et al.* 2010 Food Security: The Challenge of Feeding 9 Billion People. *Science* **327**, 812–818. (doi:10.1126/science.1185383)

2. Tilman D *et al.* 2001 Forecasting agriculturally driven global environmental change. *Science* **292**, 281–284. (doi:10.1126/science.1057544)

3. Akpoti K, Kabo-bah AT, Dossou-Yovo ER, Groen TA, Zwart SJ. 2020 Mapping suitability for rice production in inland valley landscapes in Benin and Togo using environmental niche modeling. *Sci. Total Environ.* **709**, 136165. (doi:10.1016/j.scitotenv.2019.136165)

4. Arenas-Castro S, Goncalves JF, Moreno M, Villar R. 2020 Projected climate changes are expected to decrease the suitability and production of olive varieties in southern Spain. *Sci. Total Environ.* **709**, 136161. (doi:10.1016/j.scitotenv.2019.136161)

5. Feng L, Wang H, Ma X, Peng H, Shan J. 2021 Modeling the current land suitability and future dynamics of global soybean cultivation under climate change scenarios. *Field Crops Res.* **263**, 108069. (doi:10.1016/j.fcr.2021.108069)

6. Shi X, Wang C, Zhao J, Wang K, Chen F, Chu Q. 2021 Increasing inconsistency between climate suitability and production of cotton (Gossypium hirsutum L.) in China. *Ind. Crops Prod.* **171**, 113959. (doi:10.1016/j.indcrop.2021.113959)

7. Campos H, Cooper A, Habben JE, Edmeades GO, Schussler JR. 2004 Improving drought tolerance in maize: a view from industry. *Field Crops Res.* **90**, 19–34. (doi:10.1016/j.fcr.2004.07.003)

8. Araus JL, Slafer GA, Royo C, Dolores Serret M. 2008 Breeding for Yield Potential and Stress Adaptation in Cereals. *Crit. Rev. Plant Sci.* **27**, 377–412. (doi:10.1080/07352680802467736)

9. Parent B, Leclere M, Lacube S, Semenov MA, Welcker C, Martre P, Tardieu F. 2018 Maize yields over Europe may increase in spite of climate change, with an appropriate use of the genetic variability of flowering time. *Proc. Natl. Acad. Sci.* **115**, 10642–10647. (doi:10.1073/pnas.1720716115)

10. Davis KF, Rulli MC, Seveso A, D'Odorico P. 2017 Increased food production and reduced water use through optimized crop distribution. *Nat. Geosci.* **10**, 919-+. (doi:10.1038/s41561-017-0004-5)

11. Acevedo M *et al.* 2020 A scoping review of adoption of climate-resilient crops by small-scale producers in low- and middle-income countries. *Nat. Plants* **6**, 1231-+. (doi:10.1038/s41477-020-00783-z)

12. Snapp SS, Cox CM, Peter BG. 2019 Multipurpose legumes for smallholders in sub-Saharan Africa: Identification of promising 'scale out' options. *Glob. Food Secur.* **23**, 22–32. (doi:10.1016/j.gfs.2019.03.002)

13. Sloat LL, Davis SJ, Gerber JS, Moore FC, Ray DK, West PC, Mueller ND. 2020 Climate adaptation by crop migration. *Nat. Commun.* **11**, 1243. (doi:10.1038/s41467-020-15076-4)

14. Wang H, Hijmans RJ. 2019 Climate change and geographic shifts in rice production in China. *Environ. Res. Commun.* **1**, 011008. (doi:10.1088/2515-7620/ab0856)

15. Chaloner TM, Gurr SJ, Bebber DP. 2021 Plant pathogen infection risk tracks global crop yields under climate change. *Nat. Clim. Change* **11**, 710–715. (doi:10.1038/s41558-021-01104-8)

16. Tilman D, Balzer C, Hill J, Befort BL. 2011 Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 20260–20264. (doi:10.1073/pnas.1116437108)

17. Mueller ND, Gerber JS, Johnston M, Ray DK, Ramankutty N, Foley JA. 2012 Closing yield gaps through nutrient and water management. *Nature* **490**, 254–257. (doi:10.1038/nature11420)

18. Labeyrie V *et al.* 2021 The role of crop diversity in climate change adaptation: insights from local observations to inform decision making in agriculture. *Curr. Opin. Environ. Sustain.* **51**, 15–23. (doi:10.1016/j.cosust.2021.01.006)

19. Diamond J. 2002 Evolution, consequences and future of plant and animal domestication. *Nature* **418**, 700–707. (doi:10.1038/nature01019)

20. Khoury CK *et al.* 2016 Origins of food crops connect countries worldwide. *Proc. R. Soc. B-Biol. Sci.* **283**, 20160792. (doi:10.1098/rspb.2016.0792)

21. Pironon S, Etherington TR, Borrell JS, Kühn N, Macias-Fauria M, Ondo I, Tovar C, Wilkin P, Willis KJ. 2019 Potential adaptive strategies for 29 sub-Saharan crops under future climate change. *Nat. Clim. Change* **9**, 758–763. (doi:10.1038/s41558-019-0585-7)

22. Ramirez-Villegas J, Khoury CK, Achicanoy HA, Mendez AC, Diaz MV, Sosa CC, Debouck DG, Kehel Z, Guarino L. 2020 A gap analysis modelling framework to prioritize collecting for ex situ conservation of crop landraces. *Divers. Distrib.* **26**, 730–742. (doi:10.1111/ddi.13046)

23. Pironon S *et al.* 2020 Toward Unifying Global Hotspots of Wild and Domesticated Biodiversity. *Plants* **9**, 1128. (doi:10.3390/plants9091128)

24. Castañeda-Álvarez NP *et al.* 2016 Global conservation priorities for crop wild relatives. *Nat. Plants* **2**, 1–6. (doi:10.1038/nplants.2016.22)

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. *Glob. Biogeochem. Cycles* **22**. (doi:10.1029/2007GB002947)

26. Ramankutty N, Foley JA, Norman J, McSweeney K. 2002 The global distribution of cultivable lands: current patterns and sensitivity to possible climate change. *Glob. Ecol. Biogeogr.* **11**, 377–392. (doi:10.1046/j.1466-822x.2002.00294.x)

27. Milla R. 2020 Crop Origins and Phylo Food: A database and a phylogenetic tree to stimulate comparative analyses on the origins of food crops. *Glob. Ecol. Biogeogr.* **29**, 606–614. (doi:10.1111/geb.13057)

28. Maitner BS *et al.* 2018 The bien r package: A tool to access the Botanical Information and Ecology Network (BIEN) database. *Methods Ecol. Evol.* **9**, 373–379. (doi:https://doi.org/10.1111/2041-210X.12861)

29. Dornelas M *et al.* 2018 BioTIME: A database of biodiversity time series for the Anthropocene. *Glob. Ecol. Biogeogr.* **27**, 760–786. (doi:https://doi.org/10.1111/geb.12729)

30. Dauby G *et al.* 2016 RAINBIO: a mega-database of tropical African vascular plants distributions. *PhytoKeys* **74**, 1–18. (doi:10.3897/phytokeys.74.9723)

31. Meyer C, Weigelt P, Kreft H. 2016 Multidimensional biases, gaps and uncertainties in global plant occurrence information. *Ecol. Lett.* **19**, 992–1006. (doi:https://doi.org/10.1111/ele.12624)

32. Zizka A *et al.* 2019 CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods Ecol. Evol.* **10**, 744–751. (doi:https://doi.org/10.1111/2041-210X.13152)

33. Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder HP, Kessler M. 2017 Climatologies at high resolution for the earth's land surface areas. *Sci. Data* **4**, 170122. (doi:10.1038/sdata.2017.122)

34. Hopkins M. 1991 Human-Development Revisited - a New Undp Report. *World Dev.* **19**, 1469–1473. (doi:10.1016/0305-750X(91)90089-Z)

35. Hijmans RJ *et al.* 2021 raster: Geographic Data Analysis and Modeling.

36. Licker R, Johnston M, Foley JA, Barford C, Kucharik CJ, Monfreda C, Ramankutty N. 2010 Mind the gap: how do climate and agricultural management explain the 'yield gap' of croplands around the world? *Glob. Ecol. Biogeogr.* **19**, 769–782. (doi:10.1111/j.1466-8238.2010.00563.x)

37. OECD. 2021 *Agricultural Policy Monitoring and Evaluation 2021: Addressing the Challenges Facing Food Systems*. Paris: Organisation for Economic Co-operation and Development. See https://www.oecd-ilibrary.org/agriculture-and-food/agricultural-policy-monitoring-and-evaluation-2021_2d810e01-en.

38. Tittonell P, Giller KE. 2013 When yield gaps are poverty traps: The paradigm of ecological intensification in African smallholder agriculture. *Field Crops Res.* **143**, 76–90. (doi:10.1016/j.fcr.2012.10.007)

39. Hoffman AL, Kemanian AR, Forest CE. 2018 Analysis of climate signals in the crop yield record of sub-Saharan Africa. *Glob. Change Biol.* **24**, 143–157. (doi:10.1111/gcb.13901)

40. Yu Q *et al.* 2020 A cultivated planet in 2010 – Part 2: The global gridded agricultural-production maps. *Earth Syst. Sci. Data* **12**, 3545–3572. (doi:10.5194/essd-12-3545-2020)

41. Daryanto S, Wang L, Jacinthe P-A. 2017 Global synthesis of drought effects on cereal, legume, tuber and root crops production: A review. *Agric. Water Manag.* **179**, 18–33. (doi:10.1016/j.agwat.2016.04.022)

42. Shang Q, Ling N, Feng X, Yang X, Wu P, Zou J, Shen Q, Guo S. 2014 Soil fertility and its significance to crop productivity and sustainability in typical agroecosystem: a summary of long-term fertilizer experiments in China. *Plant Soil* **381**, 13–23. (doi:10.1007/s11104-014-2089-6)

43. Ricciardi V, Ramankutty N, Mehrabi Z, Jarvis L, Chookolingo B. 2018 How much of the world's food do smallholders produce? *Glob. Food Secur.* **17**, 64–72. (doi:10.1016/j.gfs.2018.05.002)

44. Waha K, Dietrich JP, Portmann FT, Siebert S, Thornton PK, Bondeau A, Herrero M. 2020 Multiple cropping systems of the world and the potential for increasing cropping intensity. *Glob. Environ. Change* **64**, 102131. (doi:10.1016/j.gloenvcha.2020.102131)

45. Karlen DL, Varvel GE, Bullock D, Cruse RM. 1994 Crop Rotations for the 21st Century. *Adv. Agron.* **53**, 1–45. (doi:10.1016/S0065-2113(08)60611-2)

46. Mehrabi Z, Pironon S, Kantar M, Ramankutty N, Rieseberg L. 2019 Shifts in the abiotic and biotic environment of cultivated sunflower under future climate change. *OCL* **26**, 9. (doi:10.1051/ocl/2019003)

47. Khoury CK, Bjorkman AD, Dempewolf H, Ramirez-Villegas J, Guarino L, Jarvis A, Rieseberg LH, Struik PC. 2014 Increasing homogeneity in global food supplies and the implications for food security. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 4001–4006. (doi:10.1073/pnas.1313490111)

48. Vikram P *et al.* 2015 Drought susceptibility of modern rice varieties: an effect of linkage of drought tolerance with undesirable traits. *Sci. Rep.* **5**, 14799. (doi:10.1038/srep14799)

49. Fumia N, Pironon S, Rubinoff D, Khoury CK, Gore MA, Kantar MB. In press. Wild relatives of potato may bolster its adaptation to new niches under future climate scenarios. *Food Energy Secur.* **n/a**, e360. (doi:10.1002/fes3.360)

50. Dempewolf H, Bordoni P, Rieseberg LH, Engels JMM. 2010 Food Security: Crop Species Diversity. *Science* **328**, 169–170. (doi:10.1126/science.328.5975.169-e)

51. Siefert A, Fridley JD, Ritchie ME. 2014 Community Functional Responses to Soil and Climate at Multiple Spatial Scales: When Does Intraspecific Variation Matter? *PLOS ONE* **9**, e111189. (doi:10.1371/journal.pone.0111189)

52. Mahaut L, Violle C, Renard D. 2021 Complementary mechanisms stabilize national food production. *Sci. Rep.* **11**, 4922. (doi:10.1038/s41598-021-84272-z)

53. Descheemaeker K, Ronner E, Ollenburger M, Franke AC, Klapwijk CJ, Falconnier GN, Wichern J, Giller KE. 2019 WHICH OPTIONS FIT BEST? OPERATIONALIZING THE SOCIO-ECOLOGICAL NICHE CONCEPT. *Exp. Agric.* **55**, 169–190. (doi:10.1017/S001447971600048X)

Data availability

The sources of all data used in this study are referenced in the Methods and all raw data are freely accessible at the URLs provided in Extended Data Table 2. The dataset used for the analyses is available from the corresponding author upon request (lucie.mahaut@cefe.cnrs.fr).

Acknowledgements

LM's work was funded by the French National Research Agency under the Programme “Investissements d’Avenir” under the reference ANR 17 MPGA 0004. LHR's research on sunflower climate adaptation is funded by Genome Canada, Genome BC, and The Natural Sciences and Engineering Research Council of Canada. The Ministry of Economy and Competitivy of Spain (Grants CGL2014-56567-R and CGL2017-83855-R; Ministerio de Economía y Competitividad, Spain) fund RM research on crop’s wild progenitors. SP thanks the Bentham-Moxon Trust for funding a short stay at the University of British Columbia

(BMT35-2017). C.K.K. was supported by grant no. 2019-67012-29733/project accession no. 1019405 from the USDA National Institute of Food and Agriculture. C.V. was supported by the European Research Council (ERC) Starting Grant Project "Ecophysiological and biophysical constraints on domestication in crop plants" (Grant ERC-StG-2014-639706-CONSTRAINTS). We thank Ian Ondo for his help with soil data, Emily Warschefsky and Navin Ramankutty for feedback on the manuscript. We thank Gilles Dauby and Thomas Couvreur for providing unpublished data from the Rainbio database. We also thank Nora Castañeda-Álvarez and Matija Obreza for extracting occurrence records from GENESYS and Dora A. L. Canhos and Sidnei De Souza for obtaining records from SpeciesLink database.

Author contributions

L.M. led the data analysis and writing. L.M., S.P., D.R. and C.V. designed the experiment. F.B., C.K.K., R.M and C.P. contributed substantially to the crop wild progenitors' analysis and writing. J.Y.B., S.P., D.R. and C.V. assisted with data analysis and writing. Z.M. and L.H.R. assisted with study design and writing.

Competing interests

The authors declare no competing financial interests.

Figures

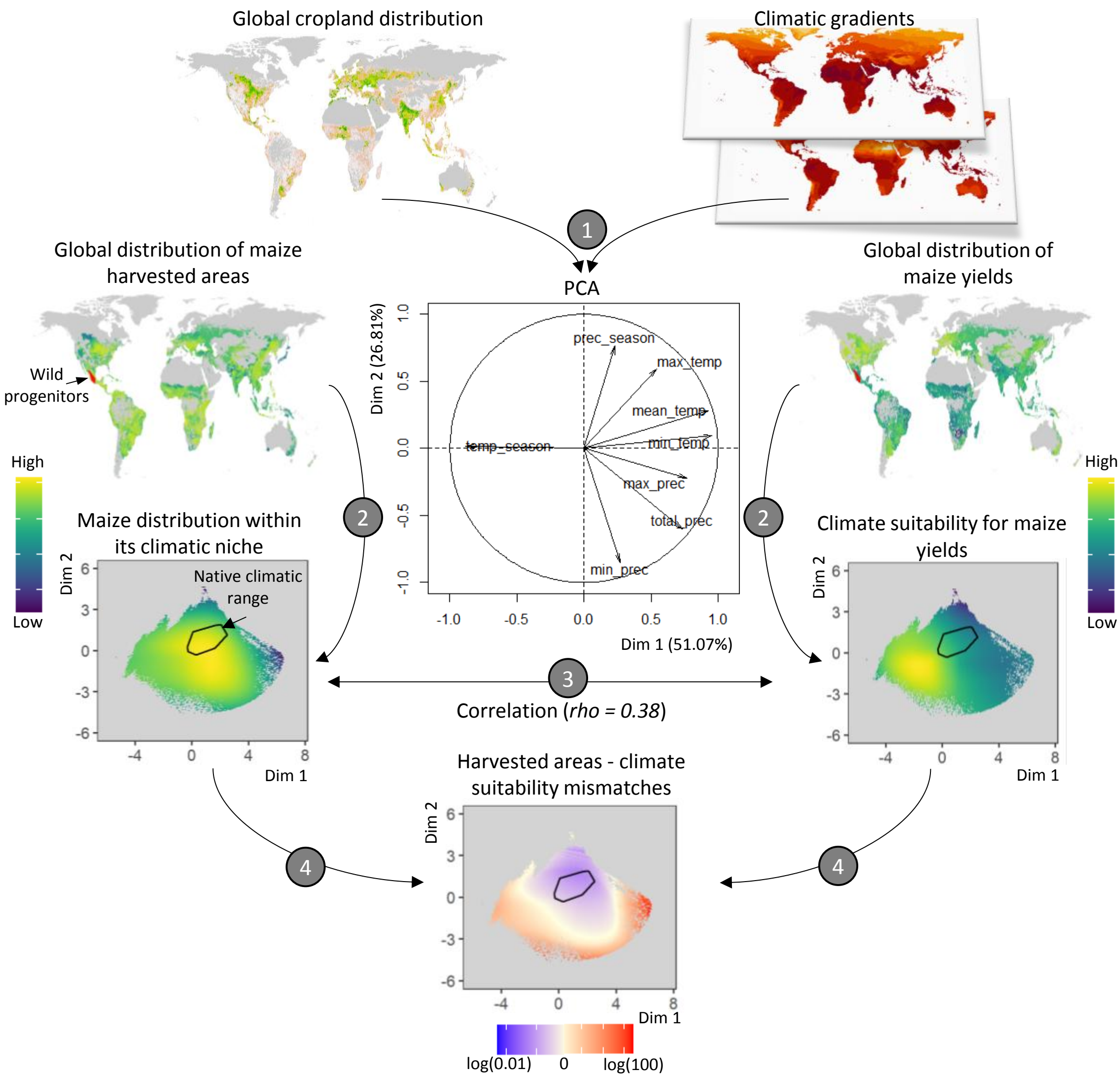

Figure 1: General framework for estimating (mis)matches between harvested areas and climate suitability for crop yields. First, we compute a unique 2D climatic space based on the global distribution of cropland and the two first dimensions of a PCA (Dim.1 and Dim. 2) that summarizes variations in eight bioclimatic variables related to mean, seasonality and extremes values of temperature (temp) and precipitation (prec). The portion of this climatic space that a crop occupies defines its climatic niche. Second, we project the global distributions of harvested areas and yields of twelve major food crops (here maize) as well as the distribution of their wild progenitors into the climatic niche of each crop (see Methods for details). The portion of the climatic niche that is occupied by wild progenitors (black polygons) represents the native range of the crop. We use multiple regression to model climate suitability as the marginal effects of Dim.1 and Dim.2 on crop yields while controlling for numerous non-climatic factors. Third, we quantify the relationship between harvested areas and climate suitability by computing Spearman's correlation index. Fourth, we analyse variation in the harvested areas-climate suitability relationship by computing a local mismatch index (see Methods for details). Positive values (red) indicate climatic zones with higher climate suitability with respect to harvested areas. Negative values (blue) indicate climatic zones with higher harvested areas with respect to climate suitability. Null values (beige) represent balanced harvested areas and climate suitability.

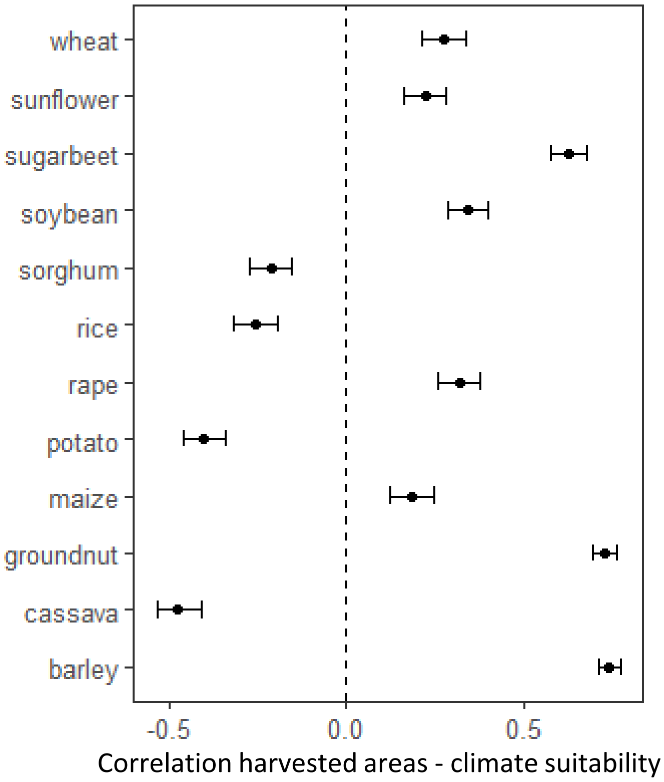

Figure 2: Correlation between the distribution of harvested areas and climate suitability for the yield of twelve major food crops. For each crop, we randomly sampled 1,000 pixels of the climate grid, computed Spearman correlation index between harvested area and climate suitability and repeated the process 1,000 times. Black dots indicate mean spearman correlation indices and horizontal bars indicate 95% confidence intervals.

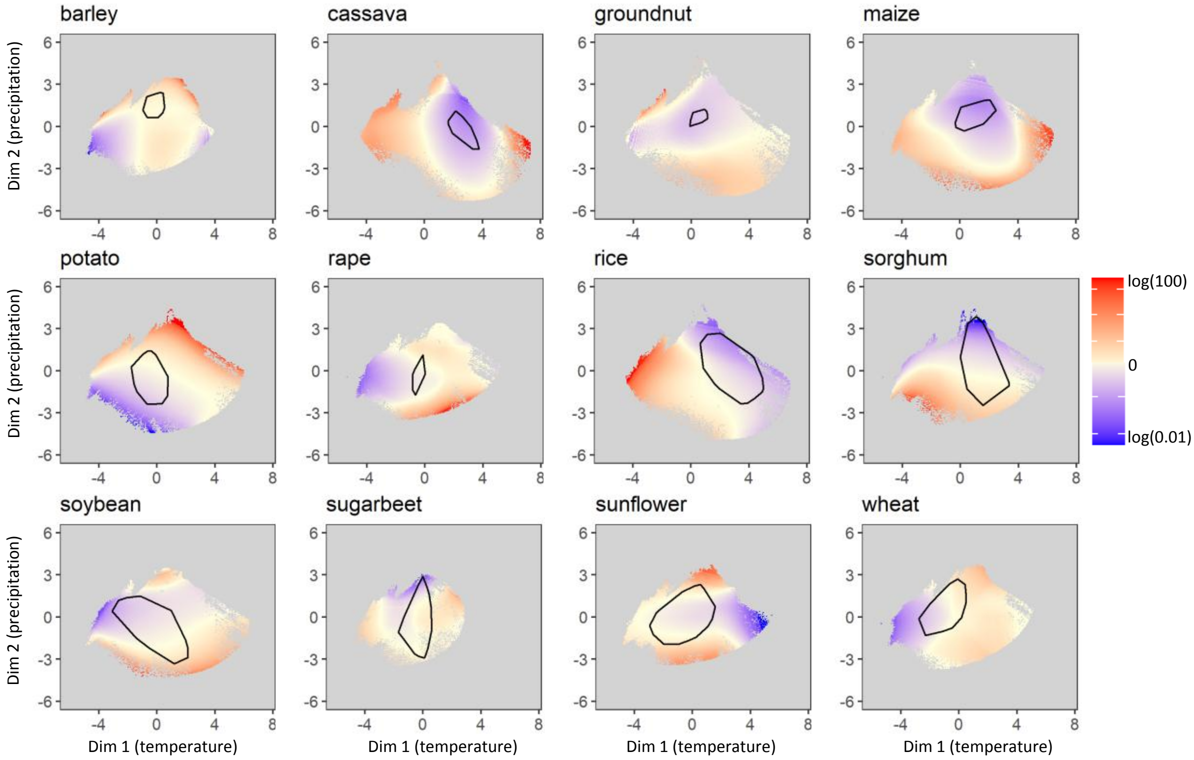

Figure 3: Variation in the harvested areas-climate suitability mismatch index within the climatic niche of twelve major food crops. Dim. 1 summarizes the variation in mean, seasonal and extreme temperature. Dim. 2 summarizes the variation in mean, seasonal and extreme precipitation. Beige colours indicate balanced harvested areas and climate suitability. Red represents zones where climate suitability is disproportionately high with respect to harvested areas. Conversely, blue indicates regions where harvested areas are disproportionately large with respect to climate suitability. Black polygons delineate the climatic space occupied by crop wild progenitors (i.e. native climatic range).

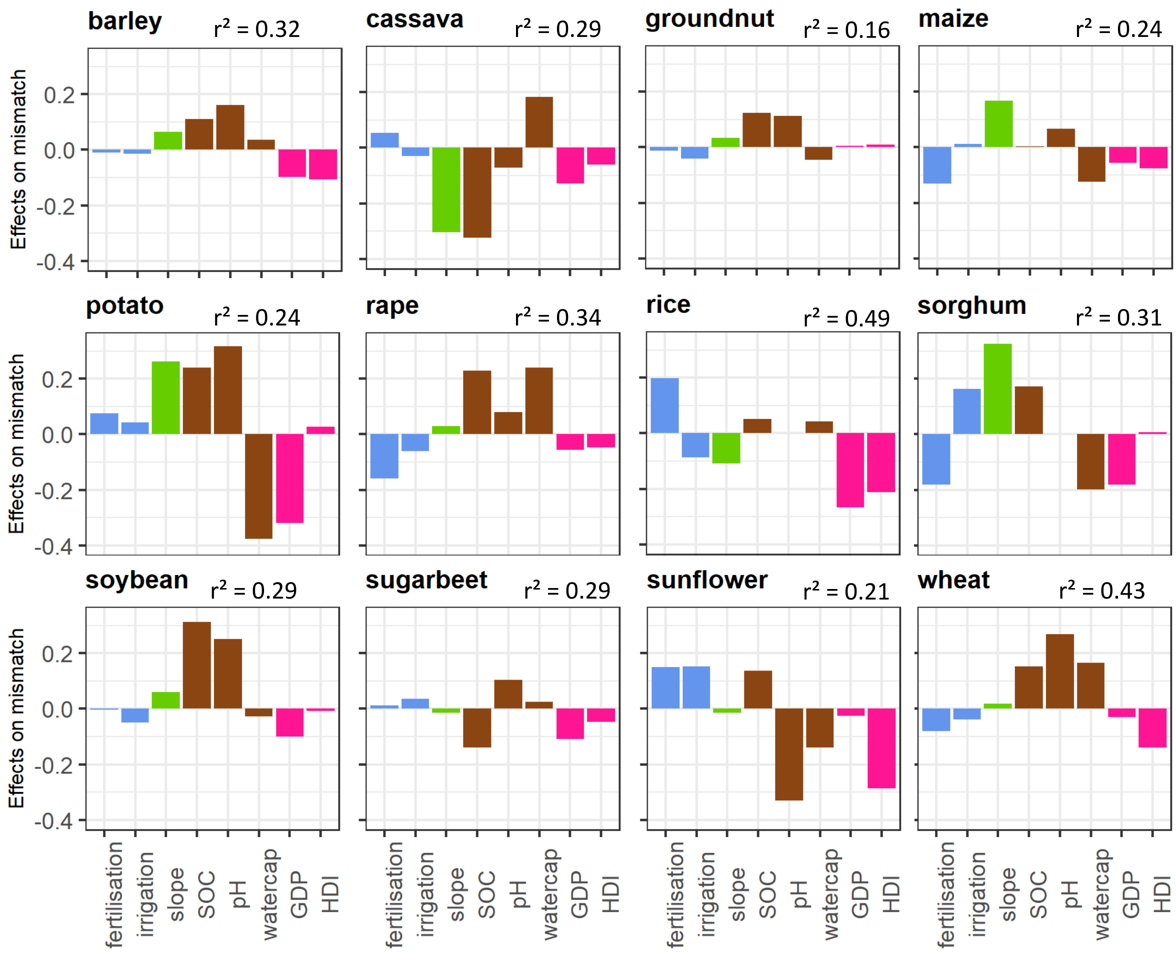

Figure 4: Estimated coefficients of the mismatch models for twelve major food crops. The effects of agricultural inputs (Fertilisation and Irrigation; blue), topography (slope; green), soil (SOC, pH, water capacity; brown) and socio-economic (GDP and HDI; pink) factors are modelled as linear terms. Positive effects on the absolute values of crop distribution – climate suitability mismatch indicate an increase in the decoupling between the distribution of harvested areas and climate suitability for crop yields while negative effects indicate an increasing match between both parameters. Coefficients of determination are shown (r^2).

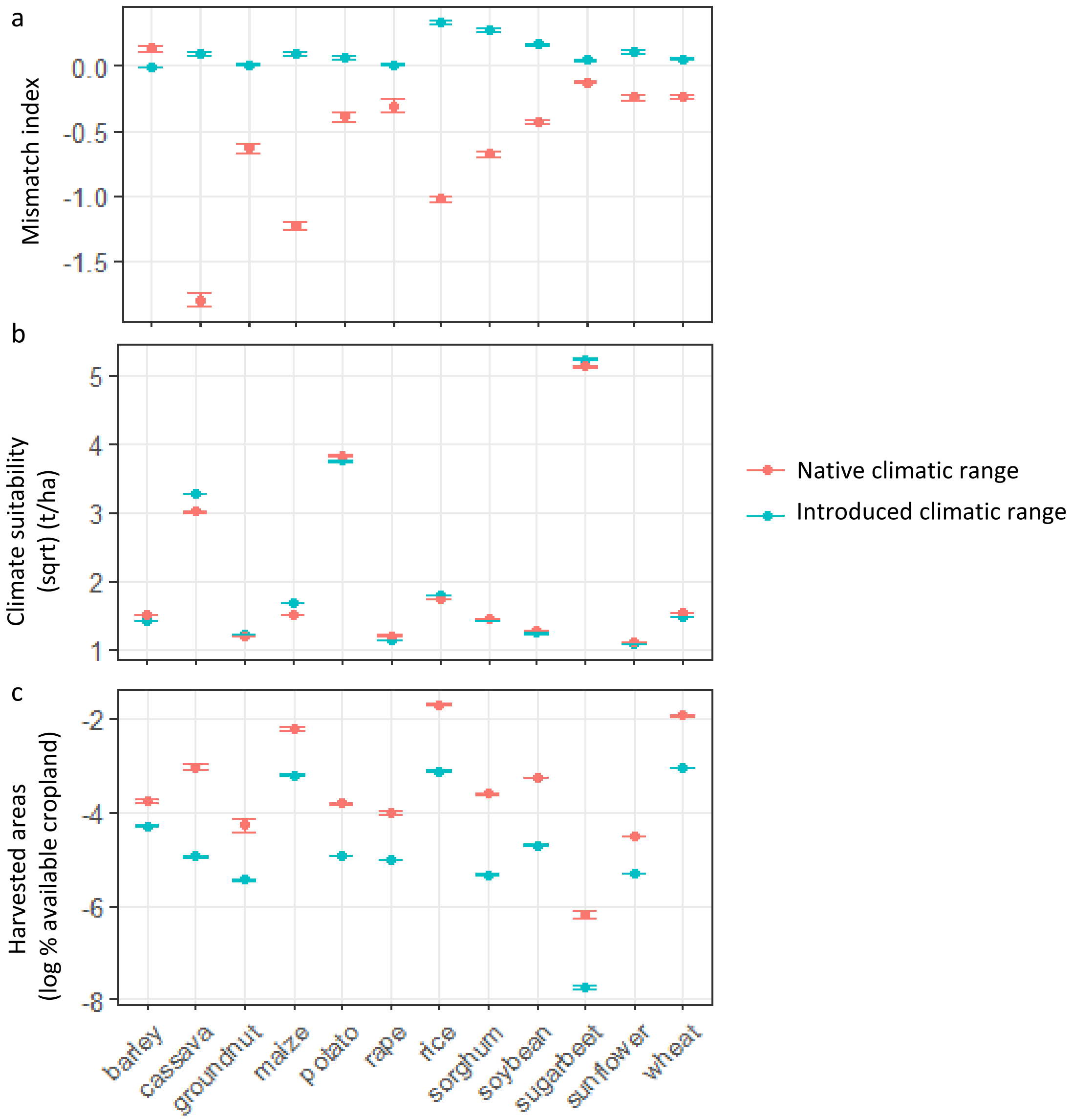

Figure 5: Average mismatch index (a), climate suitability (b) and harvested areas (c) in the native (red) and introduced (blue) climatic ranges of twelve major food crops. Dots indicate average values +/- 95% confidence intervals.