9-30-2021

Spatial frameworks for robust estimation of yield gaps

Juan I. Rattalino Edreira

José F. Andrade

Kenneth G. Cassman

Martin K. Ittersum

Marloes P. van Loon

See next page for additional authors

Authors

Juan I. Rattalino Edreira, José F. Andrade, Kenneth G. Cassman, Martin K. Ittersum, Marloes P. van Loon, and Patricio Grassini

OPEN

Spatial frameworks for robust estimation of yield gaps

Juan I. Rattalino Edreira[1], José F. Andrade[1,2], Kenneth G. Cassman [1], Martin K. van Ittersum [3], Marloes P. van Loon[3] and Patricio Grassini [1]✉

Food security interventions and policies need reliable estimates of crop production and the scope to enhance production on existing cropland. Here we assess the performance of two widely used 'top-down' gridded frameworks (Global Agro-ecological Zones and Agricultural Model Intercomparison and Improvement Project) versus an alternative 'bottom-up' approach (Global Yield Gap Atlas). The Global Yield Gap Atlas estimates extra production potential locally for a number of sites representing major breadbaskets and then upscales the results to larger spatial scales. We find that estimates from top-down frameworks are alarmingly unlikely, with estimated potential production being lower than current farm production at some locations. The consequences of using these coarse estimates to predict food security are illustrated by an example for sub-Saharan Africa, where using different approaches would lead to different prognoses about future cereal self-sufficiency. Our study shows that foresight about food security and associated agriculture research priority setting based on yield potential and yield gaps derived from top-down approaches are subject to a high degree of uncertainty and would benefit from incorporating estimates from bottom-up approaches.

Meeting food demand on existing cropland, without further encroachment of natural ecosystems such as forests, wetlands and savannahs, is one of the greatest challenges of our time[1]. Orienting investments on agricultural research and development (AR&D) to meet that challenge requires information about where the largest opportunities to increase crop yields exist within the current cultivated area[2–5]. The yield gap, defined as the difference between actual farm yield and the yield potential with good management that minimizes yield losses from biotic and abiotic stresses, is a key biophysical indicator of the available room for crop production increase with current land and water resources[6]. Global assessments of future food security and land-use change published in high-profile journals have followed a 'top-down' approach that relies on crudely calibrated crop models and a gridded spatial framework to organize coarse data on climate, soil, and cropping systems to estimate yield potential and associated yield gaps[7–10] (Fig. 1 and Supplementary Table 1). Recent assessments for specific countries suggest, however, that top-down approaches provide estimates of yield potential and yield gaps that are not useful for effective prioritization of AR&D investments[11,12].

An alternative to the use of top-down spatial frameworks is to follow a 'bottom-up' approach that estimates yield potential and yield gap for a number of sites explicitly chosen to best represent the spatial distribution of crop production area and then upscales the yield potential estimated at those sites to larger spatial scales[13] (Fig. 1 and Supplementary Table 1). While both spatial frameworks (that is, top down and bottom up) can eventually reach complete coverage of the entire cultivated area, they differ in the means to achieve it and the resulting outcomes. Bottom-up approaches favour the use of measured data on weather, soil and cropping systems and the use of crop simulation models calibrated using data from well-managed experiments where yield-limiting and reducing factors were effectively minimized, which, altogether, should lead to more accurate

estimates of yield potential and yield gaps[11,14] (Extended Data Fig. 1). The spatial granularity of the bottom-up approach, in terms of estimating yield potential for a specific combination of climate, soil and cropping systems, allows results to be validated by local experts. Moreover, results for specific locations can be aggregated to regional, national and continental scales by weighting contributions to larger-scale spatial units based on crop production area represented by soil, climate and cropping systems at each location. By contrast, outcomes produced by top-down approaches are difficult to validate because results are necessarily aggregated to the grid level, without differentiating amongst soil types and cropping systems that may exist within the grid. Weather data are also aggregated at the grid scale and may be interpolated from distant weather stations or remotely sensed data.

Yield potential and yield gaps are routinely used as inputs in studies dealing with global food security, biodiversity, land use and climate change[6,15–17]. However, despite the existence of two very different approaches to estimate these two indicators, there has been no explicit attempt to evaluate the performance of top-down versus bottom-up spatial frameworks for estimating yield potential and yield gaps at a local to global scale. We report here a global comparison of the two methods and discuss implications for informing AR&D investments. Our study includes outcomes from two of the most cited studies that utilize top-down approaches: (1) the Global Agro-ecological Zones (GAEZ) model developed by the International Institute for Applied Systems Analysis and the Food and Agricultural Organization of the United Nations (FAO; http://www.fao.org/nr/gaez/en/; refs. [18,19]) and (2) the median of the model ensemble of the Agricultural Model Intercomparison and Improvement Project (AgMIP) (https://agmip.org/; refs. [20,21]). Yield potential, yield gaps and extra production potential reported in these studies are compared against those derived from the bottom-up approach followed by the Global Yield Gap Atlas (GYGA; www.yieldgap.org; refs. [11,13,14]).

[1]Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE, USA. [2]Department of Plant Production, University of Buenos Aires, Buenos Aires, Argentina. [3]Plant Production Systems Group, Wageningen University, Wageningen, The Netherlands. ✉e-mail: pgrassini2@unl.edu
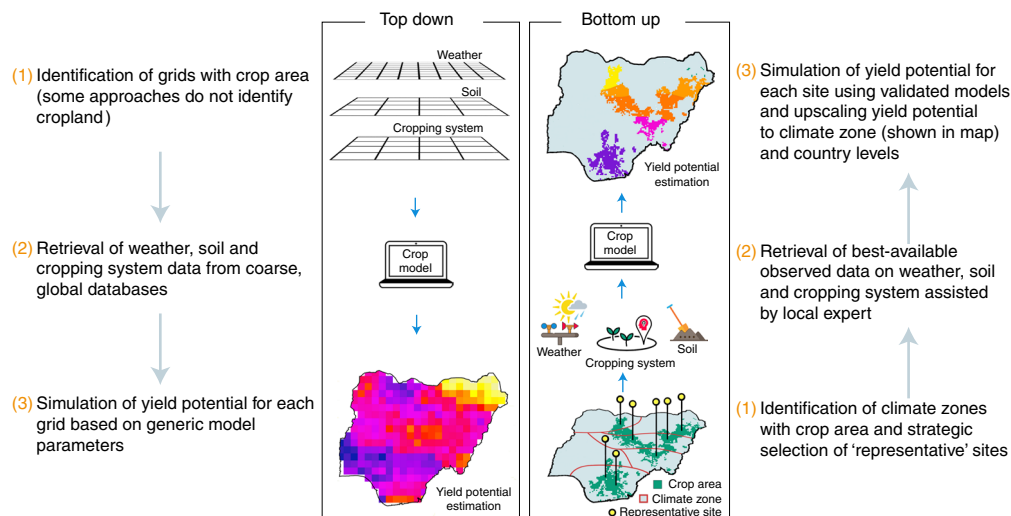
Fig. 1 | Main differences between top-down and bottom-up approaches at estimating crop yield potential. Schematic representation of the steps followed by top-down (left) and bottom-up (right) approaches to estimate yield potential in one country.

Because effective AR&D requires interventions at different spatial scales, we performed a comparison between top-down and bottom-up approaches at three levels, local, subnational ('climate zone') and national or subcontinental, with a respective average size of nearly 9,500, 60,000 and 1,000,000 km². Climate zones are geographic areas with similar temperature and water regimes[21]. We focus on cereal crops, which account for 45% of global calorie intake (https://ourworldindata.org/food-supply). We compare top-down and bottom-up estimates for major cereal crop-producing areas in North and South America, Europe, Asia, Africa and Australia (Extended Data Fig. 2). For simplicity, we show examples on four geographic regions and three crops (maize, rice and wheat). The four regions were selected for being important food exporters and/or importers. As examples of regions with favourable climate and fertile soils (that is, favourable production environments), we include maize in the US Corn Belt, which produces 35% of global maize output, and lowland irrigated and rainfed rice in Asia, which accounts for about 90% of global rice production and about 80% of rice consumption during the 2014–2018 period[22]. As an example of a harsh production environment (less favourable climate and generally infertile soils), we include wheat in Australia, which accounts for 10% of global wheat exports. Maize in sub-Saharan Africa is also included, as this region has rapid population growth rates and domestic cereal demand is projected to increase threefold over the next 30 years[23].

Results

Yield potential and yield gap comparison. Comparison of yield potential derived from top-down (GAEZ and AgMIP) versus bottom-up (GYGA) approaches reveals large discrepancies across all spatial levels. On average, yield potential estimated by AgMIP is 60% lower compared with GYGA across the four case studies (Fig. 2), which is consistent with the findings for other crop-producing regions (Supplementary Tables 2 and 3). As a result, AgMIP gives much more conservative estimates of extra crop production potential on existing cropland compared with GYGA across all spatial scales. Agreement between GYGA and GAEZ was better at national and subcontinental scales, although there were still large discrepancies between the two approaches, ranging from −50% to +30%. These differences were even larger at smaller spatial scales and for specific regions and crops, with GAEZ estimates differing from GYGA by −95% to 480% at local levels (Supplementary Tables 2

and 3). In some cases, yield potential derived from bottom-up and top-down approaches follows the same trend across locations and climate zones, but there remain several substantial disagreements on the absolute level of yield potential. That was the case for maize in the US Corn Belt, where GYGA estimates a yield potential that is 8% and 63% higher than that estimated by GAEZ and AgMIP, respectively (Fig. 2a,f). Similarly, estimated yield potential for rainfed wheat in Australia is 46% higher in GYGA than in AgMIP (Fig. 2g). Besides poor agreement at the national and continental levels in some cases, other cases show a complete lack of association between the yield potential derived from top-down and bottom-up approaches across locations and climate zones, as it is the case for lowland rainfed rice in Asia and maize in sub-Saharan Africa (Fig. 2b,d,f,h). In both regions, the range of yield potential across climate zones is very narrow as estimated following top-down approaches compared with much larger ranges from GYGA. In other words, some of the locations and climate zones reported by GYGA to have the highest yield potential are identified to be among the ones with lowest yield potential by GAEZ and AgMIP and vice versa.

In addition to evaluating the degree of agreement between top-down and bottom-up approaches, we also assess the quality of yield potential estimation per se by comparing the simulated yield potential against the average farm yield currently achieved in farmers' fields (actual yields). By definition, the difference between the two, the so-called yield gap (yield potential minus actual yield), cannot be negative. If an estimated yield potential value is considerably lower than average farm yield, then yield potential is clearly underestimated. We find that the top-down approaches give negative yield gaps for a considerable number of cases worldwide (Fig. 3 and Extended Data Fig. 3). At local levels, yield gaps estimated by GAEZ are negative in 13%, 3% and 3% of the 582, 302 and 478 locations evaluated for maize, rice and wheat, respectively. In the case of AgMIP, yield-gap estimates are negative in 39% (maize), 45% (rice) and 25% (wheat) of the cases. In contrast, no negative yield gaps are estimated by GYGA. Because calculation of yield gaps relies on the same source of average actual yield data for both top-down and bottom-up methods (Methods), the substantial number of cases with negative yield gaps as estimated by top-down approaches can be seen as a strong indication of underestimation of yield potential.

Implications for food self-sufficiency assessments. Although achieving food self-sufficiency is not an essential precondition for
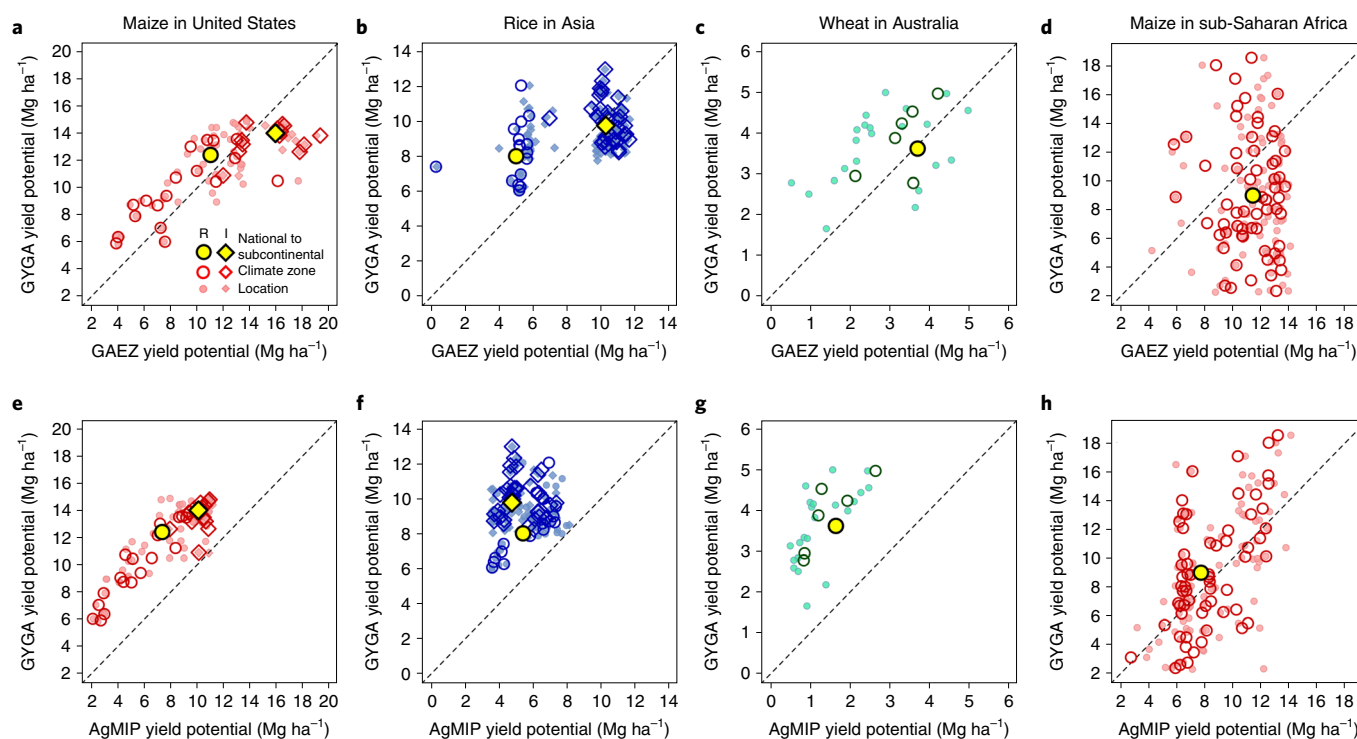
Fig. 2 | Comparison of yield potential derived from top-down and bottom-up approaches. a–h, Yield potential derived from the bottom-up GYGA versus those estimated following the top-down GAEZ (**a–d**) or AgMIP (**e–h**) for rainfed (R) and irrigated (I) maize, rice and wheat in four crop-producing regions (United States (**a,e**), Asia (**b,f**), Australia (**c,g**) and sub-Saharan Africa (**d,h**)) and at three spatial scales: local, regional (climate zone) and national or subcontinental. Each data point represents a long-term average yield potential (10–30 years of data, depending on case study). The dashed line indicates x = y. Comparisons for other cropping systems are shown in Supplementary Tables 2 and 3.

food security, it can be highly relevant for developing countries with limited capacity to purchase food imports and infrastructure to store and distribute it efficiently[24]. A key indicator of food security is the self-sufficiency ratio (SSR), which is the ratio between domestic production and total domestic consumption[25]. Comparison of SSR for different scenarios of yield-gap closure can help assess the degree of food self-sufficiency that a country or region can achieve by increasing productivity on existing cropland[26]. However, as we showed previously, such an assessment will be influenced by the choice of top-down or bottom-up approach in calculating yield potential, yield gaps and associated extra production potential. For example, self-sufficiency estimates for major cereal crops (maize, sorghum, millet, rice and wheat) vary widely across SSA countries, assuming a production scenario in which average cereal crop yields reach 80% of the yield potential by year 2050 without changes in cropland area (Fig. 4). GAEZ forecasts that the region could become self-sufficient for cereal grain by 2050 by an ample margin via narrowing current yield gaps, with the potential production exceeding expected demand by 36% (that is, SSR = 1.36). In contrast, GYGA also estimates that the region could be self-sufficient in cereals if yield gaps are closed, but with production levels very close to the expected demand by 2050 (SSR = 1.03). In the case of AgMIP, estimates of crop production potential fall short of sufficiency, indicating that cropland expansion and/or increase in food imports will be needed to meet projected cereal demand by 2050 (SSR = 0.96). Discrepancies among approaches become larger when zooming in on specific countries or regions. For example, while GAEZ predicts that yield-gap closure would result in cereal surplus in seven of the ten countries, outputs from GYGA and AgMIP suggest that most of the countries could not meet cereal demand by 2050. While SSR estimates at the subcontinental scale are similar for GYGA and AgMIP, there are large differences in estimated SSR at the national

scale, with AgMIP estimations differing from GYGA between −24% and 39%.

Discussion

A key question for AR&D is where to invest in terms of crops and regions to maximize the return on investment. While yield gap alone is not sufficient to answer this question, together with other biophysical and socioeconomic factors that influence technology adoption, it is an important parameter to guide public and private investments in agriculture, because it specifies where and how much crop production can be increased. Here we show that the choice of spatial framework to make such assessments (that is, top down or bottom up) has important implications for projecting the return on investment. For example, different approaches lead to contrasting answers about the prognosis of a given country to reach a desired level of food self-sufficiency. Even in those cases in which both approaches give similar yield-gap estimates at the subcontinental level, there are large discrepancies when looking at specific countries or regions within each country. The considerable number of locations with a negative yield gap estimated by top-down approaches raises important questions about the accuracy of these approaches in estimating yield potential and gives caution to their use for effective prioritization of AR&D investments. While we focused on extra production potential and food availability, the uncertainty associated with top-down analysis will also apply to other studies focusing on land use, climate change and biodiversity that follow a similar approach to estimate crop production potential[7,8].

Causes of inaccurate estimation of yield gaps following top-down approaches have been investigated elsewhere[26,27]; here we point out some of them. Top-down approaches are based on secondary (unmeasured) gridded data and coarse global soil maps and cropping systems data (Table 1 and Supplementary Data Table 1), which
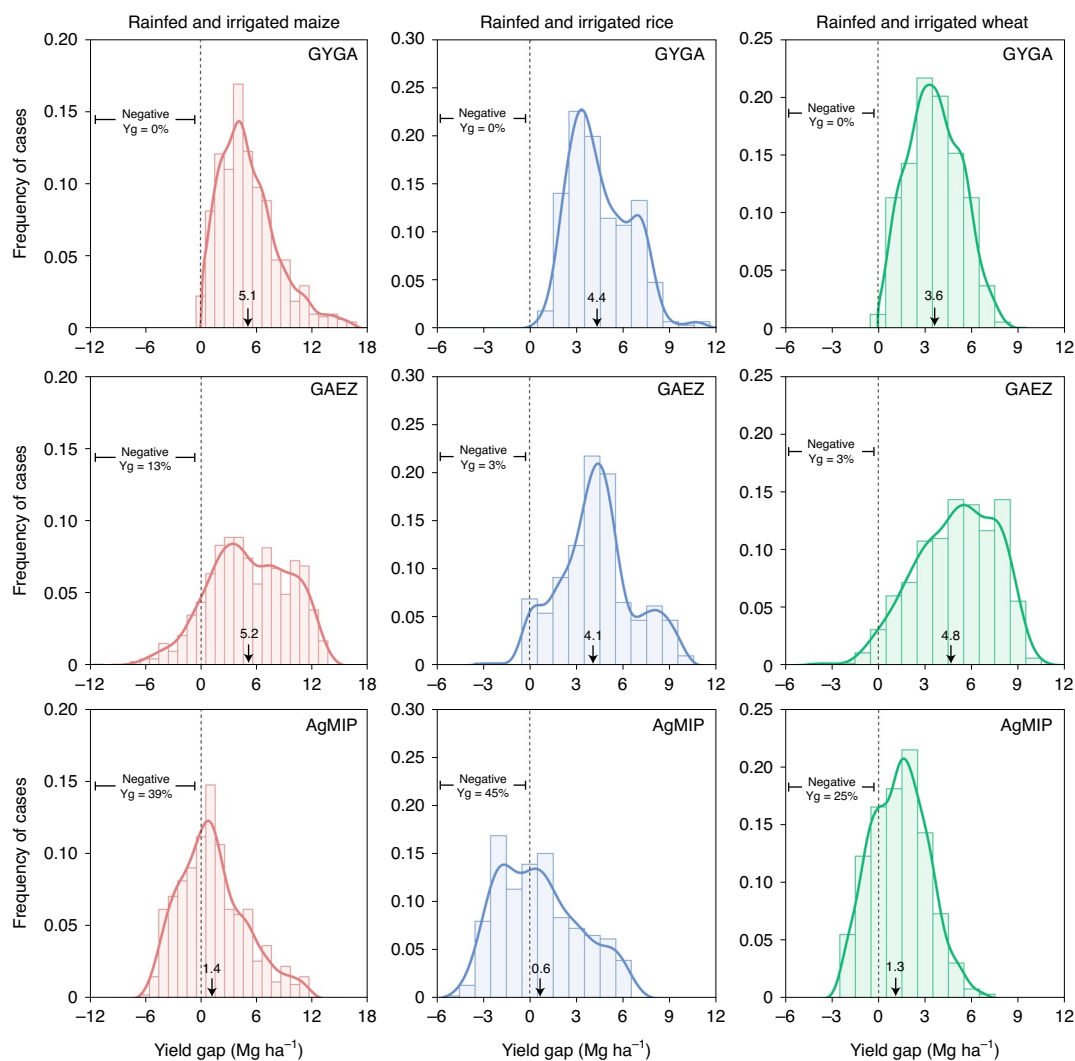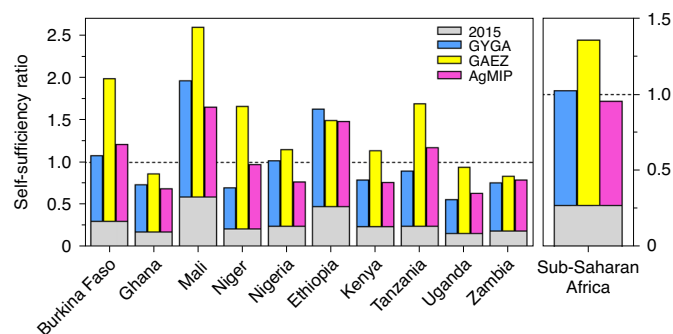
Fig. 3 | Yield gaps as estimated using bottom-up and top-down approaches. Histogram of yield gap (Yg) estimation for rainfed and irrigated maize, rice and wheat at the local scale using the GYGA, GAEZ and AgMIP frameworks. Arrows indicate the mean yield gap estimated across 582 locations for maize, 302 for rice and 478 for wheat. The vertical dotted line indicates zero yield gap, that is, no difference between yield potential and actual yield.

give a false sense of data quality and availability in spatial grids that are typically 0.5–2.0° (ca. 3,000–50,000 km² at the Equator). Indeed, previous studies have shown important biases when simulating yield potential using coarse gridded weather data compared with simulations based on measured data[27–29] or without proper selection of the dominant soil types within an agricultural area[30,31]. Similarly, the cropping-system context, including cropping intensity (that is, number of crops per year), crop calendar (sowing window and crop cycle duration) and water regime (irrigated or rainfed), is critical for the estimation of yield potential. While GYGA works with local experts to obtain reliable information about the cropping system context, the two top-down approaches rely on an in silico optimization of the cropping system (GAEZ) or coarse global crop calendars (AgMIP), predicting in many cases crop systems that do not match the dominant existing systems or even systems that simply do not exist (Table 1). For example, in the US Corn Belt, the global dataset MIRCA 2000[32] employed by AgMIP sets a maize-sowing window between April and October, but, in reality, producers typically do not sow beyond June to prevent crop loss due to fall frost[33]. Likewise, top-down approaches generally use generic crop model coefficients that do not account for the specificity of crop cultivars in terms of responses to temperature and photoperiod[21,34,35]; these models are

also rarely validated for their ability to estimate yield potential based on data collected from well-managed crops where yield-limiting and yield-reducing factors have been effectively controlled. In summary, estimates of yield potential and yield gaps derived from top-down approaches are subject to a high degree of uncertainty considering the errors associated with the underpinning data.

The accuracy of the spatial sampling framework of the GYGA bottom-up approach has been validated for regions where high-quality and spatially detailed data are available. Hochman et al.[36] conducted a study on yield gaps of rainfed wheat in Australia following two approaches: (1) the bottom-up approach of GYGA and (2) a data-rich method using high-density data available in the Australian grain zone (both relying on measured weather data). These researchers reported that the two approaches gave similar estimates of yield potential and yield gaps at climate zone and national levels. Similarly, Aramburu Merlos et al.[37] and Morell et al.[38] show that national average actual yield estimates for Argentina and the United States, calculated using a limited number of selected locations following the GYGA protocols, were remarkably similar to the reported national average yield based on data from hundreds of subnational-level administrative units covering the entire crop production area. Finally, Van Wart et al.[39] and van Bussel et al.[13]

Fig. 4 | Cereal SSR for a scenario of yield-gap closure in sub-Saharan Africa by 2050 as estimated using bottom-up and top-down approaches. The SSR was calculated as the quotient between production and demand of the most important cereals. The grey portion of the bars represents the SSR by 2050, assuming production level as for 2015; the coloured portion represents the SSR if farmers close the exploitable yield gap (that is, reach 80% of the yield potential). Separate bars are shown for bottom-up (GYGA) and top-down approaches (GAEZ and AgMIP).

Table 1 | Summary of key advantages and disadvantages associated with different spatial frameworks to estimate yield potential and yield gaps at local to global scale

Attribute	Top-down	Bottom-up
Required effort to achieve full cropland coverage	Low	High
Data input requirements	High	High
Required time and data-collection efforts	Low	High
Consistency in databases and models across countries	High	Low
Accuracy of weather, soil and cropping system data	Low	High
Effort on calibrating crop models	Low	High
Accuracy on yield potential and yield gaps	Low	High
Local relevance	Low	High

showed that variability in weather and simulated yield potential was relatively low for sites located within the same climate zones, which provides further support for a stratified (instead of random) selection of sites and use of the climate zone framework as a basis for upscaling results from location to region and country. Altogether, these studies provide strong evidence of robust estimates of yield potential and yield gaps following the bottom-up approach of the GYGA.

Given the 'global public goods' nature of food self-sufficiency estimates across local to global scales, we see an urgent need for robust estimates of yield gaps for major cropping systems worldwide as input to a national and international dialogue on future global food security under climate change. Such a project can be accomplished with a modest investment on a bottom-up approach that gives priority to use primary measured weather data, finer-scale soil maps and accurate cropping systems data for a given location[31]. We still see a number of areas for complementarity between top-down and bottom-up approaches. For example, the bottom-up approach requires more granular and detailed data on climate, soils and cropping systems than the top-down approaches evaluated in this study, which makes its application difficult in regions where these data are scarce or simply do not exist. In these regions, it may be necessary to rely on top-down approaches that obtain the required data by interpolation and informed guesses using data from coarser spatial scales. Similarly, top-down modellers may benefit from using bottom-up estimates of yield potential, and underpinning weather and soil databases, to evaluate their model outcomes. We conclude that foresight studies of food security, land use and climate change and associated priority setting on AR&D based on yield potential and yield gaps would benefit from using a bottom-up spatial framework and good-quality data to reduce uncertainties in previously reported estimates of food production potential under current and future climates.

Methods

Yield definitions. Yield potential (Yp; megagrams per harvested hectare) is defined as the yield of a cultivar in an environment to which it is adapted, when grown with sufficient water and nutrients in the absence of abiotic and biotic stress[40]. In irrigated fields, Yp is determined by solar radiation, temperature, atmospheric CO_2 concentration and management practices that influence crop cycle duration and light interception, such as sowing date, cultivar maturity and plant density. In rainfed systems where water supply from stored soil water at sowing and in-season precipitation is not enough to meet crop water requirements, water-limited Yp

(Yw) is determined by water supply amount and its distribution during the growing season, as well as by soil properties influencing the crop–water balance, such as the rootable soil depth, texture and terrain slope. Actual yield is defined as the average grain yield (megagrams per harvested hectare) obtained by farmers for a given crop with a given water regime. The difference between Yp (or Yw) and farmer actual yield is known as the yield gap[11]. In the case of irrigated crops, Yp is the proper benchmark to estimate yield gaps, while Yw is the meaningful benchmark for rainfed crops. With good, cost-effective crop management, reaching 70–80% of Yp (or Yw) is a reasonable target for farmers with good access to markets, inputs and extension services, which is usually referred to as 'attainable yield'[41,42]. Beyond this yield level, the small return to extra input requirement and labour does not justify the associated financial and environmental costs and level of sophistication in crop and soil management practices.

Sources of Yp data derived from top-down and bottom-up approaches. We retrieved data generated from two initiatives following a top-down approach: (1) the GAEZ (http://www.fao.org/nr/gaez/en/; refs. [18,19]) and (2) the AgMIP (https://agmip.org/data-and-tools-updated/; refs. [20,21]). As the bottom-up approach, we used results from the GYGA (www.yieldgap.org; refs. [11,31,43]). The main features of these databases are summarized elsewhere (Supplementary Table 1 and Supplementary Section 1). In the process of selecting the specific dataset, we explicitly attempted to reduce biases in the comparisons to the extent this was possible. For example, in all cases, we used simulations that meet the yield definitions provided in the previous section. We also tried to be consistent in terms of the time period for which Yp (or Yw) was simulated; however, this was not always possible, because while GAEZ and AgMIP use weather datasets that cover the time period between 1961 and 1990 and between 1980 and 2010, respectively, GYGA uses more recent weather data (Supplementary Table 1). Similarly, comparisons between databases were limited to those regions for which there were estimates of Yp (or Yw) for each of the top-down and bottom-up approaches. More detailed information about the three approaches can be found in Supplementary Section 1. We acknowledge that, when assessing different approaches, it is conceivable that there would be an inherent bias depending on who performs it and his/her preference. Although the authors of this current study have all contributed to the development of GYGA, we have maintained neutrality when conducting the analysis and made inferences solely based on the results shown here, avoiding any inherent bias. Additionally, methods and data sources are fully documented and publicly accessible for other researchers who may be interested in replicating our comparison.

Comparison of bottom-up and top-down approaches at different spatial levels. Comparison of the three databases needs to account for the different spatial resolution at which the data are reported (grid in GAEZ and AgMIP versus buffer in GYGA). In the present study, we compared Yp (or Yw) among the three databases at three spatial levels: local (also referred to as buffer), climate zone (CZ) and country (or subcontinent). An example of the three spatial levels evaluated in this study as well as the Yw estimated by each of the three databases for rainfed maize is shown in Extended Data Fig. 4. We note that buffer is the lowest spatial level at which Yp and Yw are reported in GYGA. For a country such as the United States, where maize production is concentrated on flat geographic areas, the average size of buffers and CZs selected by GYGA is 17,000 and 60,000 km², respectively; the size is smaller for countries with greater terrain and climate heterogeneity, such as Ethiopia, where the average size of buffers and CZs selected for maize by GYGA is a respective 4,000 and 21,000 km², or for smaller countries, such as in Europe.

The GYGA already provides estimates of Yp (or Yw) and yield gaps at those three spatial levels. Following a bottom-up approach, GYGA estimates the Yp

(or Yw) at the buffer level based on the Yp (or Yw) simulated for each crop cycle and soil type (within a given buffer) and their associated harvested area (within that same buffer) using a weighted average. Subsequently, Yp (or Yw) at buffer levels are upscaled to CZ, national or subcontinental levels using a weighted average based on harvested area retrieved from the Spatial Production Allocation Model (SPAM) 2010[44]. Details on the GYGA upscaling method can be found in van Bussel et al.[13] In the case of top-down approaches, for comparison purposes, it was necessary to aggregate Yp (or Yw) reported for each individual grid into buffers, CZs and countries in order to make them comparable to those reported by GYGA. To do so, Yp (or Yw) from GAEZ and AgMIP was scaled up to buffer, climate zone and country (or subnational levels) considering the crop-specific area within each pixel, as reported by SPAM 2010[44]. For example, for a given buffer, the average Yp (or Yw) was estimated using a weighted average, in which the value of Yp (or Yw) reported for each of the GAEZ or AgMIP grids located within the GYGA buffer was 'weighted' according to the SPAM crop-specific area within that grid. The same approach was used to estimate average Yp (or Yw) at the CZ and country (or subcontinental) levels for GAEZ and AgMIP.

For a given buffer, CZ or country (or subcontinent), the yield gap was calculated as the difference between Yp (or Yw) and the average farmer yield (actual yield, Ya). The Yp and Yw were taken as the appropriate benchmarks to estimate yield gaps for irrigated and rainfed crops, respectively. To avoid biases due to the source of average actual yield in the estimation of yield gap, we used the average actual yield dataset from GYGA, because it provides estimates of average actual yield disaggregated by water regime for the most recent time period. Actual yield data from GYGA were retrieved from official statistics available at subnational administrative units such as municipalities, counties, departments and subdistrict. The exact number of years of data to calculate average yield is determined by GYGA on a case-by-case basis, following the principle of including as many recent years of data as possible to account for weather variability while avoiding the bias due to a technological time trend and long-term climate change[31]. Using the GYGA database on average actual yield for estimation of yield gaps does not bias the results from our study, as GYGA favours the use of official sources of average yields at the finer available spatial resolution, which is the same source of actual yield data used by other databases such as FAO and SPAM[22,44]. In this study, we opted not to use actual yield data from GAEZ, because they derived from FAOSTAT statistics of the years 2000 and 2005, and thus, they could lead to an overestimation of the yield gap in those regions where actual yields have increased over the past two decades[19]. Finally, extra production potential was calculated based on the yield gap estimated by each approach and the SPAM crop-specific harvested area reported for each buffer, CZ and country (or subcontinent). The top-down and bottom-up approaches were compared in a total of 67 countries, which together account for 74%, 67% and 43% of global maize, rice and wheat harvested areas, respectively (Extended Data Fig. 2). Overall, our comparison included a total of 1,362 buffers located within 870 CZs, with 422 buffers (within 249 CZs) for rainfed maize, 160 buffers (116 CZs) for irrigated maize, 93 buffers (66 CZs) for rainfed rice, 209 buffers (114 CZs) for irrigated rice, 400 buffers (274 CZs) for rainfed wheat and 78 buffers (49 CZs) for irrigated wheat. In all cases, Yp (or Yw), yield gaps and extra production potential were expressed at standard commercial moisture content (that is, 15.5% for maize, 14% for rice and 13.5% for wheat).

We assessed the agreement in Yp (or Yw), yield gap, and extra production potential between GYGA and the two databases that follow a top-down approach (GAEZ and AgMIP) separately for each of the spatial levels (buffer, CZ, country or subcontinent) by calculating root-mean-square error (RMSE) and absolute mean error (ME):

$$\text{RMSE} = \sqrt{\frac{\sum (Y_{TD} - Y_{BU})^2}{n}} \qquad (1)$$

$$\text{ME} = \frac{\sum (Y_{TD} - Y_{BU})}{n} \qquad (2)$$

where Y_{TD} and Y_{BU} are the estimated Yp (or Yw), yield gap, or extra production potential for database i following a top-down approach and for GYGA, respectively, and n is the number of paired Y_{TD} versus Y_{BU} comparisons at a given spatial scale for a given crop in a given country. Separate comparisons were performed for irrigated and rainfed crops.

Impact of Yp estimates on food self-sufficiency analysis. We assessed the impact of discrepancies in Yp (or Yw) between top-down and bottom-up approaches on the SSR, which is an important indicator for food security. To do so, we focused on cereal crops in sub-Saharan Africa, and we calculated the SSR for the five main cereal crops in this region (that is, maize, millet, rice, sorghum and wheat) following van Ittersum et al.[23]. Millet and sorghum were included in the analysis of SSR in sub-Saharan Africa, because together they account for ca. 25% of the total cereal production and ca. 40% of the total cereal harvested area in this region (average over the 2015–2019 period)[22]. Briefly, we computed current national demand (assumed equal to the 2015 consumption) and the 2015 production of the five cereals to estimate the baseline SSR (that is, in 2015) in ten countries

for which Yw (or Yp) data were available in GYGA. Current total cereal demand per country were calculated as the product of current population size derived from United Nations population prospects and cereal demand per capita based on the International Model for Policy Analysis of Agricultural Commodities and Trade (IMPACT)[35,45]. The annual per-capita demand for the five cereals was expressed in maize yield equivalents by using the crop-specific grain caloric contents, with caloric contents based on FAO food balances[46]. Current domestic grain production per cereal crop per country (approximately 2015) was calculated as mean actual crop yield (2003–2012) as estimated in GYGA times the 2015 harvested area per crop by FAO[22]. Total future annual cereal demand per capita (2050), for each of the five cereals and each country, was retrieved from IMPACT modelling results[35] using the shared socioeconomic pathway (SSP2, no climate change) from the Intergovernmental Panel on Climate Change fifth assessment[47]. Total cereal demand per country in 2050 was calculated based on projected 2050 population (medium-fertility variant of United Nations population prospects; https://population.un.org/wpp/) multiplied by the per-capita cereal demand in 2050 from the SSP2 scenario. In our study, we assumed an attainable yield of 80% of Yw for rainfed crops, which is consistent with the original approach followed by van Ittersum et al.[23], but, in our study, we also used 80% of Yp for irrigated crops as an estimate of the attainable yield, instead of 85% as in van Ittersum et al.[23], to be slightly more conservative. Because the goal was to understand the level of SSR on existing cropland, we assumed no expansion of rainfed or irrigated cropland and no change in net planted area for each of the cereal crops. Our calculations for sub-Saharan Africa may be too pessimistic if genetic progress to increase Yp is achieved. Historically, genetic progress in Yp has contributed to progress in farm yields, although the magnitude of Yp increase is debatable. Progress in elevating Yp of the major cereals would imply, however, that even larger yield gaps need to be overcome than the already large gaps reported herein. Hence, we did not account for changes in genetic Yp in our calculation of SSR by 2050, also because climate change is likely to have a negative effect on Yp and Yw in sub-Saharan Africa.

Data availability

Data on yield potential and actual yield from GYGA are available at www.yieldgap.org. Data on yield potential from AgMIP and GAEZ can be downloaded from www.fao.org/nr/gaez/en and https://agmip.org/data-and-tools-updated/, respectively. Source data are provided with this paper.

References

1. Cassman, K. G. & Grassini, P. A global perspective on sustainable intensification research. *Nat. Sustain.* **3**, 262–268 (2020).
2. Pardey, P. G., Chan-Kang, C., Dehmer, S. P. & Beddow, J. M. Agricultural R&D is on the move. *Nature* **537**, 301–303 (2016).
3. Wood, S. & Pardey, P. G. Agroecological aspects of evaluating agricultural R&D. *Agric. Syst.* **57**, 13–41 (1998).
4. Alston, J. M., Norton, G. W. & Pardey, P. G. *Science under Scarcity: Principles and Practice for Agricultural Research Evaluation and Priority Setting* (Cornell Univ. Press, 1995).
5. Folberth, C. et al. The global cropland-sparing potential of high-yield farming. *Nat. Sustain.* **3**, 281–289 (2020).
6. Godfray, H. C. J. et al. Food security: the challenge of feeding 9 billion people. *Science* **327**, 812–818 (2010).
7. Springmann, M. et al. Options for keeping the food system within environmental limits. *Nature* **562**, 519–525 (2018).
8. Erb, K.-H. et al. Exploring the biophysical option space for feeding the world without deforestation. *Nat. Commun.* **7**, 11382 (2016).
9. Koh, L. P. & Ghazoul, J. Spatially explicit scenario analysis for reconciling agricultural expansion, forest protection, and carbon conservation in Indonesia. *Proc. Natl Acad. Sci. USA* **107**, 11140–11144 (2010).
10. van Vliet, J. Direct and indirect loss of natural area from urban expansion. *Nat. Sustain.* **2**, 755–763 (2019).
11. van Ittersum, M. K. et al. Yield gap analysis with local to global relevance—a review. *Field Crops Res.* **143**, 4–17 (2013).
12. Deng, N. et al. Closing yield gaps for rice self-sufficiency in China. *Nat. Commun.* **10**, 1725 (2019).
13. van Bussel, L. G. J. et al. From field to atlas: upscaling of location-specific yield gap estimates. *Field Crops Res.* **177**, 98–108 (2015).
14. Grassini, P. et al. Robust spatial frameworks for leveraging research on sustainable crop intensification. *Glob. Food Sec.* **14**, 18–22 (2017).
15. Foley, J. A. et al. Solutions for a cultivated planet. *Nature* **478**, 337–342 (2011).
16. Bruinsma, J. *The Resource Outlook to 2050: By How Much Do Land, Water and Crop Yields Need to Increase by 2050?* (FAO, 2009).

17. Suh, S. et al. Closing yield gap is crucial to avoid potential surge in global carbon emissions. *Glob. Environ. Change* **63**, 102100 (2020).
18. Fischer, G., Shah, M., van Velthuizen, H. & Nachtergaele, F. O. *Global Agro-ecological Assessment for Agriculture in the 21st Century* (IIASA, 2001).
19. *Global Agro-ecological Zones (GAEZ v3.0)* (IIASA and FAO, 2012); http://www.fao.org/nr/gaez/en/
20. Rosenzweig, C. et al. The Agricultural Model Intercomparison and Improvement Project (AgMIP): protocols and pilot studies. *Agric. For. Meteorol.* **170**, 166–182 (2013).
21. Elliott, J. et al. The global gridded crop model intercomparison: data and modeling protocols for phase 1 (v1.0). *Geosci. Model Dev.* **8**, 261–277 (2015).
22. *FAOSTAT* (FAO, 2021); www.faostat.fao.org
23. van Ittersum, M. K. et al. Can sub-Saharan Africa feed itself? *Proc. Natl Acad. Sci. USA* **113**, 14964–14969 (2016).
24. Clapp, J. Food self-sufficiency: making sense of it, and when it makes sense. *Food Policy* **66**, 88–96 (2017).
25. Alexandratos, N. & Bruinsma, J. *World Agriculture towards 2030/2050: The 2012 Revision* (Agricultural Development Economics Division, FAO, 2012).
26. Blankenau, P. A., Kilic, A. & Allen, R. An evaluation of gridded weather data sets for the purpose of estimating reference evapotranspiration in the United States. *Agric. Water Manag.* **242**, 106376 (2020).
27. Mourtzinis, S., Rattalino Edreira, J. I., Conley, S. P. & Grassini, P. From grid to field: assessing quality of gridded weather data for agricultural applications. *Eur. J. Agron.* **82A**, 163–172 (2017).
28. van Wart, J., Grassini, P. & Cassman, K. G. Impact of derived global weather data on simulated crop yields. *Glob. Change Biol.* **19**, 3822–3834 (2013).
29. Ramirez-Villegas, J. & Challinor, A. Assessing relevant climate data for agricultural applications. *Agric. For. Meteorol.* **161**, 26–45 (2012).
30. Hendriks, C. M. J., Stoorvogel, J. J. & Claessens, L. Exploring the challenges with soil data in regional land use analysis. *Agric. Syst.* **144**, 9–21 (2016).
31. Grassini, P. et al. How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis. *Field Crops Res.* **177**, 49–63 (2015).
32. Portmann, F. T., Siebert, S. & Döll, P. MIRCA2000—global monthly irrigated and rainfed crop areas around the year 2000: a new high-resolution data set for agricultural and hydrological modeling. *Glob. Biogeochem. Cycles* https://doi.org/10.1029/2008gb003435 (2010).
33. Grassini, P., Specht, J. E., Tollenaar, M., Ciampitti, I. A. & Cassman, K. G. in *Crop Physiology: Applications for Genetic Improvement and Agronomy* (eds Sadras, V. O & Calderini, D. F.) 15–42 (Academic Press, 2015).
34. Müller, C. et al. The Global Gridded Crop Model Intercomparison phase 1 simulation dataset. *Sci. Data* **6**, 50 (2019).
35. Robinson, S. et al. *The International Model for Policy Analysis of Agricultural Commodities and Trade (IMPACT): Model Description for Version 3* (International Food Policy Research Institute, 2015).
36. Hochman, Z., Gobbett, D., Horan, H. & Navarro Garcia, J. Data rich yield gap analysis of wheat in Australia. *Field Crops Res.* **197**, 97–106 (2016).
37. Aramburu Merlos, F. et al. Potential for crop production increase in Argentina through closure of existing yield gaps. *Field Crops Res.* **184**, 145–154 (2015).
38. Morell, F. J. et al. Can crop simulation models be used to predict local to regional maize yields and total production in the U.S. Corn Belt? *Field Crops Res.* **192**, 1–12 (2016).
39. van Wart, J. et al. Use of agro-climatic zones to upscale simulated crop yield potential. *Field Crops Res.* **143**, 44–55 (2013).
40. Evans, L. T. *Crop Evolution, Adaptation, and Yield* (Cambridge Univ. Press, 1993).
41. Cassman, K. G. Ecological intensification of cereal production systems: Yield potential, soil quality, and precision agriculture. *Proc. Natl Acad. Sci. USA* **96**, 5952–5959 (1999).
42. Lobell, D. B., Cassman, K. G. & Field, C. B. Crop yield gaps: their importance, magnitudes, and causes. *Annu. Rev. Environ. Resour.* **34**, 179–204 (2009).
43. van Wart, J., Kersebaum, K. C., Peng, S., Milner, M. & Cassman, K. G. Estimating crop yield potential at regional to national scales. *Field Crops Res.* **143**, 34–43 (2013).
44. International Food Policy Research Institute Global spatially-disaggregated crop production statistics data for 2010 version 2.0 *Harvard Dataverse* https://doi.org/10.7910/DVN/PRFF8V (2019).
45. *World Population Prospects: The 2015 Revision* (United Nations, 2015).
46. *Food Balance Sheets: A Handbook* (FAO, 2001).
47. O'Neill, B. C. et al. A new scenario framework for climate change research: the concept of shared socioeconomic pathways. *Climatic Change* **122**, 387–400 (2014).
48. Lollato, R., Edwards, J. & Ochsner, T. Meteorological limits to winter wheat productivity in the U.S. southern Great Plains. *Field Crops Res.* **203**, 212–226 (2017).
49. Yang, H., Grassini, P., Cassman, K. G., Aiken, R. M. & Coyne, P. I. Improvements to the hybrid-maize model for simulating maize yields in harsh rainfed environments. *Field Crops Res.* **204**, 180–190 (2017).
50. Li, T. et al. From ORYZA2000 to ORYZA (v3): an improved simulation model for rice in drought and nitrogen-deficient environments. *Agric. For. Meteorol.* **237–238**, 246–256 (2017).
51. Hochman, Z., Holzworth, D. & Hunt, J. Potential to improve on-farm wheat yield and WUE in Australia. *Crop Pasture Sci.* https://doi.org/10.1071/CP09064 (2009).
52. Agustiani, N. et al. Simulating rice and maize yield potential in the humid tropical environment of Indonesia. *Eur. J. Agron.* **101**, 10–19 (2018).

Author contributions
Research was conceived by J.I.R.E., P.G., K.G.C. and M.K.v.I. Data acquisition, data processing, and statistical analysis were performed by J.I.R.E., J.F.A. and M.P.v.L. The manuscript was written by J.I.R.E. and P.G. with input from all authors.

Competing interests
The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s43016-021-00365-y.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43016-021-00365-y.

Correspondence and requests for materials should be addressed to Patricio Grassini.

Peer review information *Nature Food* thanks Tony Fischer and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.
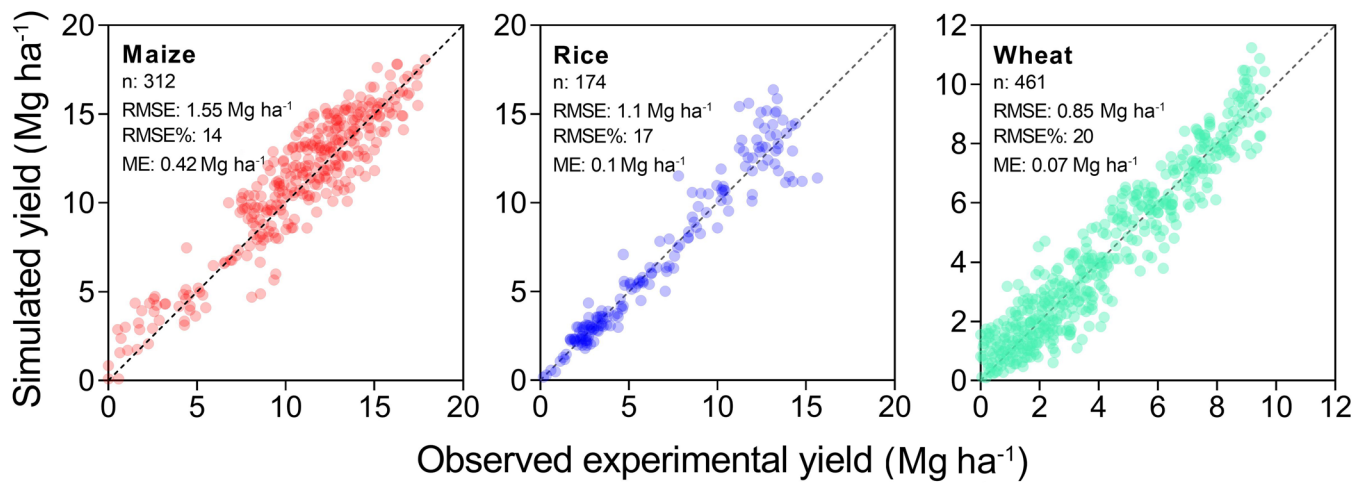
Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended Data Fig. 1 | Evaluation of crop simulation models used in GYGA. Simulated yield potential plotted against measured yields in well-managed experiments conducted across a wide range of rainfed and irrigated environments. Details are provided elsewhere[12,37,48–52]. Number of data points (n), root mean square error (RMSE), RMSE as percentage of the average measured yield (RMSE%), and mean error (ME) are indicated. Diagonal dashed line indicates y = x.
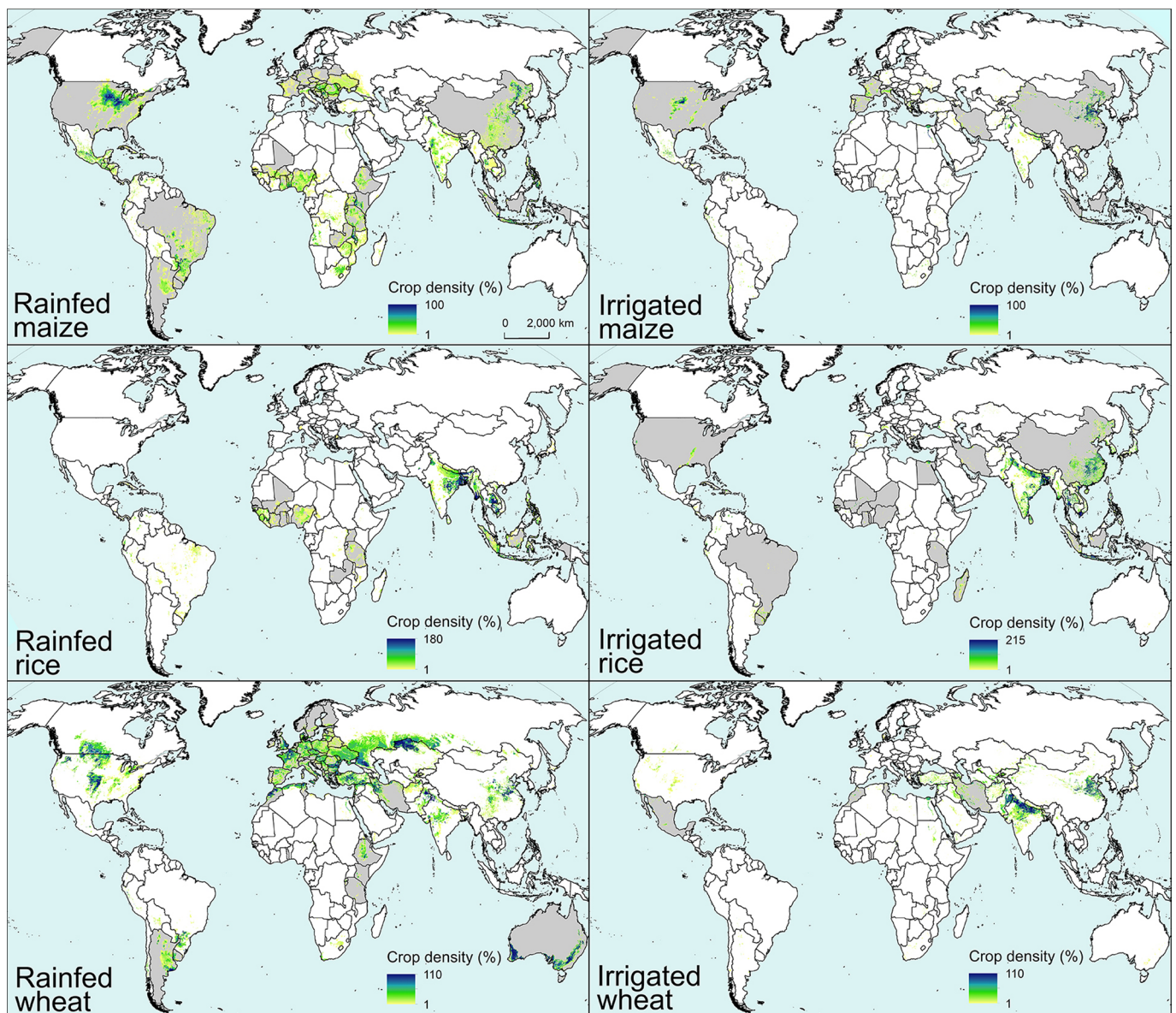
Extended Data Fig. 2 | Target crops and countries. Crop density of rainfed and irrigated maize, rice, and wheat, and countries included in the comparison shown in the present study (dark grey). Crop harvested area density was retrieved from SPAM 2010[44].
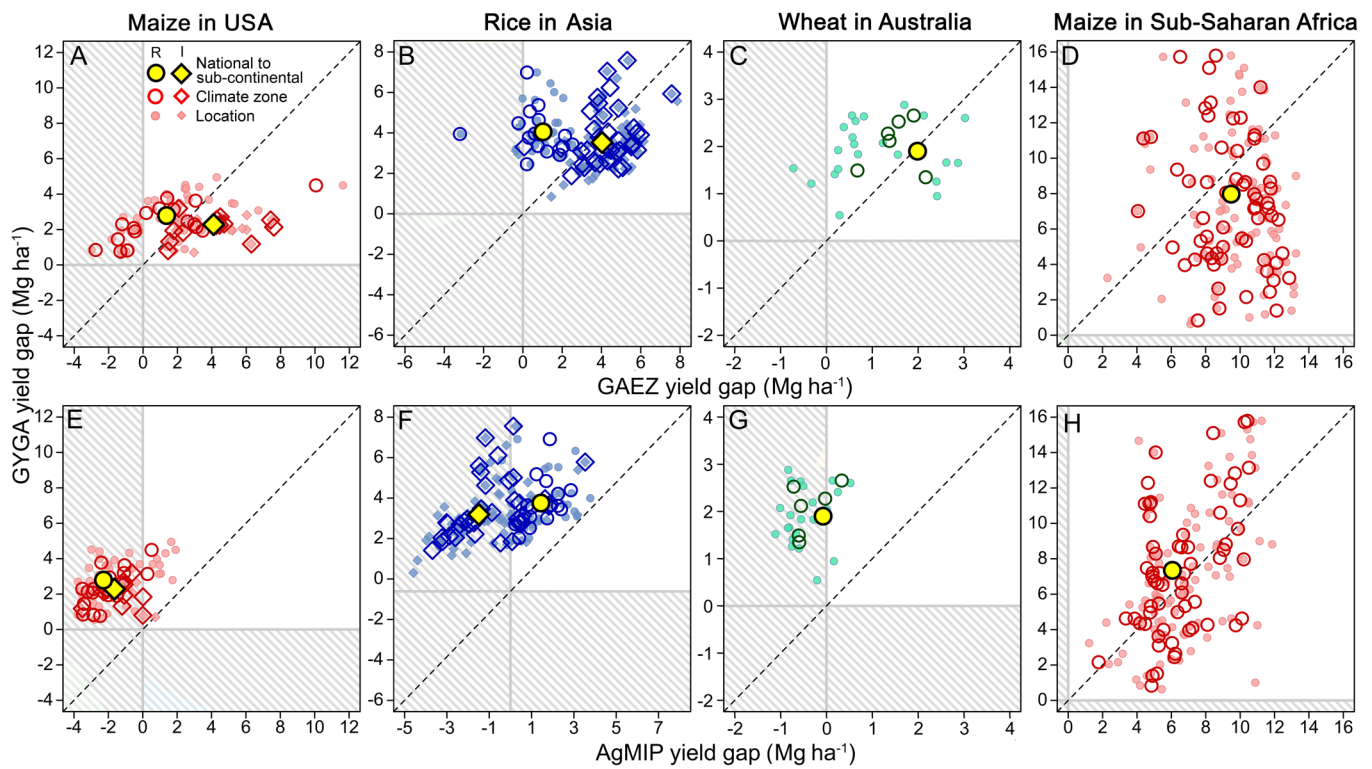
Extended Data Fig. 3 | Comparison of yield gap derived from top-down and bottom-up approaches. Yield gap estimated following a bottom-up approach (GYGA) versus those derived from top-down GAEZ (upper panels) and AgMIP (bottom panels) for rainfed (R) and irrigated (I) maize, rice, and wheat in important producing regions and at three spatial scales: local, regional (climate zone), and national or subcontinental. Inside each panel, dashed areas indicate negative yield gaps. Each data point represents the average yield gap. The dashed line corresponds to x = y. Comparisons for other cropping systems are shown in the Supplementary Materials.
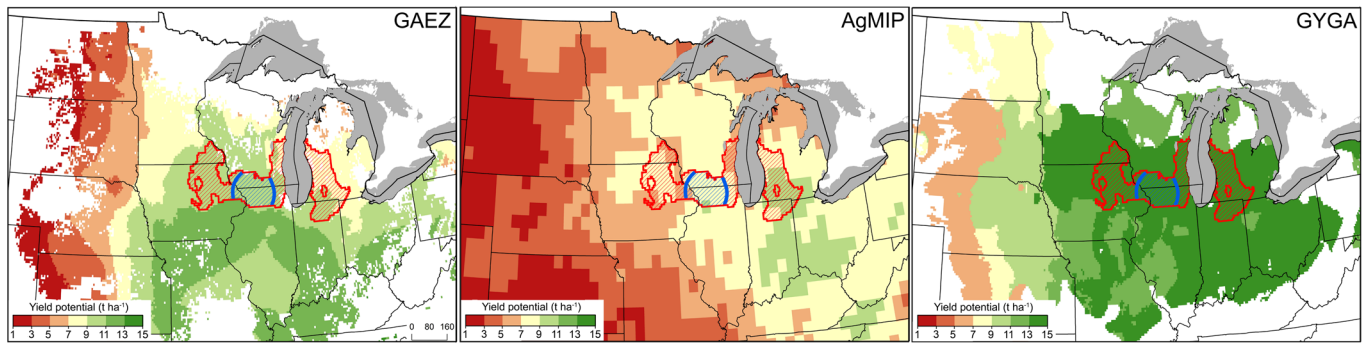
Extended Data Fig. 4 | Yield potential derived from top-down and bottom-up approaches. Comparison of water-limited yield potential estimated for rainfed maize in the US Corn Belt following top-down (GAEZ and AgMIP) and bottom-up (GAEZ) approaches. Figures show three spatial resolutions: local (buffer with blue borders), subnational (climate zone with red borders), and national. For simplicity, only one buffer and one climate zone are shown.

nature research

Corresponding author(s): Patricio Grassini

Last updated by author(s): 07-29-2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Data on yield were downloaded from the Global Yield Gap Atlas, GAEZ, and AgMIP websites. Data on crop distribution were downloaded from SPAM MAP, and data on crop area were download from FAOSTAT. All these sources are publicly available
Data analysis	Data were analyzed using R and mapped using ArcGIS.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data on yield potential and actual yield from GYGA are available at www.yieldgap.org. Data on yield potential from AgMIP and GAEZ can be download from www.fao.org/nr/gaez/en and www.agmip.org, respectively.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We assess the performance of two widely used "top-down" gridded frameworks (GAEZ and AgMIP) versus an alternative "bottom-up" approach that estimates extra production potential locally, for a number of representative sites, and then upscales the results to larger spatial scales (GYGA).
Research sample	Our comparison included a total of 1362 buffers located within 870 climate zones, with 422 buffers (within 249 CZs) for rainfed maize, 160 buffers (116 CZs) for irrigated maize, 93 buffers (66 CZs) for rainfed rice, 209 buffers (114 CZs) for irrigated rice, 400 buffers (274 CZs) for rainfed wheat, and 78 buffers (49 CZs) for irrigated wheat.
Sampling strategy	The top-down and bottom-up approaches were compared in a total of 67 countries, which together account for 74%, 67%, and 43% of global maize, rice, and wheat harvested areas, respectively.
Data collection	Data were downloaded from the GAEZ, AgMIP, and GYGA websites.
Timing and spatial scale	Comparison of the three databases was performed at three spatial levels: local (also referred as "buffer"), climate zone, and country. We also try to be consistent in terms of the time period for yield potential was simulated; however, this was not always possible because while GAEZ and AgMIP use weather data sets that cover the time period between 1961-1990 and 1980-2010, respectively, GYGA uses more recent weather data. This would, however, not biased the conclusions from the study as we are comparing long-term averages in all cases rather than specific years.
Data exclusions	No data were excluded from the analysis.
Reproducibility	Raw data inputs used for model simulations and other analyses are available upon request.
Randomization	Not applicable as we did not conduct field experiments.
Blinding	Not applicable.

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒ ☐	Antibodies
☒ ☐	Eukaryotic cell lines
☒ ☐	Palaeontology and archaeology
☒ ☐	Animals and other organisms
☒ ☐	Human research participants
☒ ☐	Clinical data
☒ ☐	Dual use research of concern

Methods

n/a	Involved in the study
☒ ☐	ChIP-seq
☒ ☐	Flow cytometry
☒ ☐	MRI-based neuroimaging