Global Ecology and Biogeography, (Global Ecol. Biogeogr.) (2014)

Global Ecology and Biogeography
A Journal of Macroecology

RESEARCH PAPER

An analysis of methodological and spatial differences in global cropping systems models and maps

Weston Anderson[1]*, Liangzhi You[1,2]*, Stanley Wood[3], Ulrike Wood-Sichra[1] and Wenbin Wu[2]

[1]*International Food Policy Research Institute, 2033 K Street, NW, Washington, DC 20006, USA,* [2]*Key Laboratory of Agri-informatics, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China,* [3]*Bill and Melinda Gates Foundation, 500 Fifth Avenue North, Seattle, WA 98109, USA*

*Correspondence: Weston Anderson, Liangzhi You, International Food Policy Research Institute, 2033 K Street, NW, Washington, DC 20006, USA
E-mail: weston.b.anderson@gmail.com; l.you@cgiar.org

ABSTRACT

Aim Agricultural practices have dramatically altered the land cover of the earth, but the spatial extent and intensity of these practices is often difficult to catalogue. Information on the distribution and performance of specific crops is often only available through national or subnational statistics. Recently, however, there have been multiple independent efforts to incorporate the detailed information available from statistical surveys with supplemental spatial information to produce a spatially explicit global dataset specific to individual crops. While these datasets provide decision makers with improved information on global cropping systems, the final global cropping maps differ substantially from one another. This study aims to explore and quantify systematic similarities and differences between four major global cropping systems products and the subsequent implications for analyses dependent on those models.

Location This study was conducted at a global scale.

Methods Each global cropping systems model was assessed by latitude as a measure of biophysical plausibility and each pair of models was compared using a Gaussian filter to remove trivial spatial discrepancies. Model disagreement was explored in relation to the interdependent input data of each model pair with a particular focus on cropland extent. The influence of the observed model differences on subsequent analyses was demonstrated using model-dependent estimates of the yield gap as an example.

Results The results of our analysis indicate that the choice of cropping systems model is non-trivial: considerable differences exist between model-specific estimates of the yield gap across nearly all climate zones and the average model difference exceeds the average estimated yield gap in certain regions. The differences in crop-specific harvested area and yield products of each model are significant, and mostly result from differences in the input datasets and downscaling methodologies. In particular, the choice of dataset on cropland extent proved to be influential regardless of the downscaling process employed.

Main conclusions Discrepancies in the final products of cropping systems models are currently poorly understood, but have implications for basic policy decisions relating to agricultural production and food security. The considerable disagreement between models serves as a reminder of the ongoing challenges to the creation of spatially explicit estimates of harvested area and yield based on crop statistics. Our analysis helps shed light on the importance of model choice by demonstrating the implications for further analyses that depend on cropping systems models, and works to overcome these challenges by characterizing model-dependent differences in harvested area and yield.

Keywords

Cropping systems, GAEZ, global cropland, harvested area, M3, MIRCA, SPAM, yield, yield gap.

DOI: 10.1111/geb.12243
http://wileyonlinelibrary.com/journal/geb

INTRODUCTION

Cropland accounts for nearly 15×10^6 km^2 of the earth's land cover – amounting to 12% of the earth's ice-free land surface – yet information on the distribution and performance of specific crops is often only available through national or subnational statistics (Foley *et al.*, 2005; Ramankutty *et al.*, 2008). Cataloguing the increasing extent and productivity of cropland has implications for analyses of food security, studies of land degradation and resource management. While subnational statistics provide information on raw quantities, they provide only limited information that is useful for spatially explicit applications. Detailed mapping of the impacts of agriculture is vital to understanding water usage (Rosegrant *et al.*, 2002), nutrient cycling (Bondeau *et al.*, 2007; Liu *et al.*, 2010), soil erosion (Yang *et al.*, 2003), loss of biodiversity (Foley *et al.*, 2011) and the impact on regional climate (Ramankutty *et al.*, 2006; Voldoire *et al.*, 2007).

Remote-sensing products offer spatially disaggregated information, but those currently available on a global scale are ill-suited to many applications due to the limited separation of crop types within the area classified as cropland. Recently, however, there have been multiple independent efforts to incorporate the detailed information available from statistical surveys with supplementary spatial information to produce a spatially explicit global dataset specific to individual crops for the year 2000. These studies have generated increasingly sophisticated results portraying the downscaling of crop production statistics at a global extent at moderately high spatial resolution. Global studies have been reported by Leff *et al.* (2004), Monfreda *et al.* (2008), Portmann *et al.* (2010) and most recently by Fischer *et al.* (2013) and You *et al.* (2014). While these datasets provide analysts and decision makers with improved information on global cropping systems, the final global cropping systems maps differ from one another substantially. This spatial uncertainty is not unique to cropping systems models and is, in fact, similar in nature to the uncertainties arising from the multiple available land-cover datasets. However, while there have been efforts to characterize and quantify uncertainties in global land-cover datasets (Giri *et al.*, 2005; Fritz & See, 2008; Herold *et al.*, 2008; Fritz *et al.*, 2011) there has to date been no such analysis for global cropping systems models. To more completely characterize the uncertainties in global agricultural impacts it is necessary to analyse each of the global cropping systems models, as has been done with land-cover datasets.

This study aims to explore and quantify systematic similarities and differences between four major global cropping systems products and the subsequent implications for analyses dependent on those models. The models used in this analysis include the monthly irrigated and rainfed crop areas around the year 2000 (MIRCA2000; Portmann *et al.*, 2010), the spatial production allocation model (SPAM2000; You *et al.*, 2014), the global agro-ecological zone (GAEZ) dataset (Fischer *et al.*, 2013) and the M3 dataset developed by Monfreda *et al.* (2008). We begin with an overview of the methods and outputs of each dataset before comparing the downscaling methodologies used by each model. We further compare the input datasets used and evaluate how interdependences between models may propagate through the downscaling methodology. Following the qualitative analysis, we quantitatively evaluate discrepancies between the four models, focusing on wheat (analyses of rice and maize are available in Appendix S5 in Supporting Information). Finally, using the global yield gap as an example, we demonstrate the influence of observed model differences on subsequent analyses. We conclude with a summary and some recommendations for users of these products.

OVERVIEW OF GLOBAL CROPPING SYSTEMS MODELS

Research on cropping system models has been reported at the global scale by Leff *et al.* (2004), You *et al.* (2014), Monfreda *et al.* (2008) and Portmann *et al.* (2010), while regional applications have been reported for Latin America and the Caribbean (You & Wood, 2006) and sub-Saharan Africa (You *et al.*, 2009). This paper focuses on four global cropping system models: M3, MIRCA, GAEZ and SPAM. Although differences in methodology and final products can at times make comparisons between the four models difficult, Table 1 provides an overview of each model that may be used to infer basic similarities and differences between products. A brief description of each model is included in the following sections, with complete details available in Sections S1.1–S1.4 of Appendix S1.

M3 cropping system model

Out of the four global cropping systems models considered, the M3 approach (Monfreda *et al.*, 2008) attempts spatial downscaling of the most complete coverage of crops (175 crops, including tree and forage crops and managed grasslands) for both harvested area and yield. M3 is described in the companion paper to Ramankutty *et al.* (2008), which uses remote-sensing products to construct a new dataset for croplands and pasture around the year 2000 at a 5-arcmin resolution. The M3 dataset applies minimal modelling to distribute subnational statistics of yield and harvested area, opting for ease of interpretation and a limit to requisite assumptions over complexity.

Monthly irrigated and rainfed crop areas around the year 2000 (MIRCA)

MIRCA downscales 26 crops including two aggregate categories of 'other annual' and 'other perennial' crops, all of which are divided into rainfed and irrigated production areas. But unique amongst the four approaches, MIRCA also performs a temporal downscaling so as to provide estimates of rainfed and irrigated area disaggregated by month (Portmann *et al.*, 2010). MIRCA uses M3 downscaled crop data results as its starting point for allocating the total harvested area for each crop into rainfed and irrigated areas, and apportions the M3 crop area allocations into 402 spatial 'calendar units' globally for which the MIRCA team

Table 1 Summary of the four global cropping system products.

	M3	MIRCA	SPAM	GAEZ
Crops:				
Crop classes (including 'other')	175	26 irrigated, 26 rainfed	21	23
Includes forages	Yes	Yes	No	Yes
Other category	n.a.	Other annual crops other perennial crops	Other crops	Other cereals, other crops
Grassland/pasture	Ramankutty pastureland*	managed grassland/pasture	n.a.	Bioenergy feedstock
Crop system disaggregation	None	Irrigated, rainfed	Irrigated, rainfed (commercial), rainfed (non-commercial), rainfed (subsistence)	Irrigated, rainfed
Production indicators	Harvested area, yield	Harvested area	Harvested area, physical area, yield and production	Harvested area, production value, yield
Seasonality	Annual	Monthly	Annual	Annual
Data portal	http://www.geog.mcgill.ca/landuse/pub/Data/175crops2000	http://www2.uni-frankfurt.de/45218023/MIRCA	http://mapspam.info/data/	http://www.gaez.iiasa.ac.at/ http://gaez.fao.org/

n.a., not applicable.
*See Ramankutty *et al.* (2008), the companion paper to Monfreda *et al.* (2008), for details.

has been able to compile unique sets of ancillary information on irrigation, crop-specific irrigated water use, crop calendars and cropping intensities.

Spatial production allocation model (SPAM)

SPAM has the most limited crop coverage, just 20 crops, but downscales the area and yield for each crop into three different production systems: (high-input) irrigated, high-input rainfed and low-input rainfed. The low-input rainfed category is itself further subdivided into low input and subsistence (You *et al.*, 2014). SPAM relies on a separate collection of subnational statistical data from that of MIRCA, focusing on increased coverage in developing countries. The SPAM approach relies on constructing prior probabilities through expert elicitation and using ancillary information – including crop prices, population density (CIESIN, IFPRI and WRI, 2000) and crop-specific biophysical suitability (Fischer *et al.*, 2013) – to distribute subnational statistics within the cropland extent (Ramankutty *et al.*, 2008) based on a method known as cross entropy (Golan *et al.*, 1996; Lence & Miller, 1998; Zhang & Fan, 2001).

Global agro-ecological zones cropping system model (GAEZ)

GAEZ downscales 23 crops including forages and other cereals in either irrigated or rainfed production systems for both harvested area and yield (Fischer *et al.*, 2013). GAEZ develops a cropland extent independent from SPAM, MIRCA and M3 – which all use the cropland extent from Ramankutty *et al.* (2008) – but does so using many of the same source datasets as Ramankutty *et al.* (2008). In the downscaling procedure, GAEZ develops and relies on an extensive analysis of crop-specific agro-climatic and edaphic suitability criteria. The model uses a methodology mathematically comparable to the cross-entropy framework of SPAM to incorporate ancillary information such as population density, crop price and market access as a means of distributing subnational statistics.

MODEL INPUT DATA AND INTERDEPENDENCES

The major determinants of the potential reliability of downscaling efforts are: (1) the quality of the cropland extent dataset indicating the physical extent and area intensity of cropland (e.g. share of cropland area in each 5-arcmin grid cell), and (2) the resolution and reliability of the subnational crop statistics. Each model builds on a common set of available data as well as previous work in cropping systems modelling. Table 2 illustrates both the broad linkages and increasing sets of input data and assumptions that each of the M3, MIRCA, GAEZ and SPAM models relies upon. These linkages are briefly outlined in the following paragraphs, and are described in detail in Appendix S2.

Differences between M3, MIRCA, SPAM and GAEZ persist all the way back to basic data collection. As a matter of calibration, all four datasets draw on FAOSTAT national data to provide control totals for cropland area, the harvested area and yields of specific crops. However, each product also expends considerable efforts to collect subnational crop statistics to allow as detailed a disaggregation of national totals within subnational administrative boundaries as possible. This gives rise to region-specific differences in the subnational data employed.

Table 2 Input data layers of the four models.

Category	Dataset		M3	MIRCA	SPAM	GAEZ
National/subnational statistics	FAOSTAT – national land use and crop production statistics		X	X	X	X
	FAO AGRO-MAPS – subnational crop statistics (SAGE and IFPRI collaboration)*		X	Via M3 data	X	X
	Additional crop distribution data		X	X	X	No documentation available
	Expert elicitation				X	
Cropland extent and cropping systems	FAO AQUASTAT – national irrigated crop statistics			X	X	X
	GMIA – global irrigated land			X	X	X
	GLC2000		X	M3 cropland extent modified	M3 cropland extent modified	X
	Boston University MODIS-derived land cover		X	M3 cropland extent modified	M3 cropland extent modified	
	GAEZ suitability index				X	X
Ancillary data	Population density	FAO-SDRN (derived from LANDSCAN 2003)				X
		GRUMP**			X	
	FAO ruminant livestock density					X
	Distance to market					X
	Crop prices				X	

*MIRCA2000 relies on M3 as does GAEZ 'for selected crops in countries where more than 50% was covered by sub-national statistics'.
**CIESIN, IFPRI and CIAT.

As a first step towards delineating crop-specific harvested area and yield, each cropping system model defines a spatially explicit layer of cropland extent, representing the proportion of cropland in each 5-min pixel globally. Because each subsequent step in the modelling process relies on the definition of cropland extent, the degree to which each pair of cropland extent products agree represents an upper bound of intermodel agreement on the spatial distribution of physical crop areas. Ramankutty *et al.* (2008) provides the base dataset for cropland extent in M3, MIRCA and SPAM, although MIRCA and SPAM both modify the dataset. GAEZ does not use the Ramankutty *et al.* (2008) data, but instead runs a cross-sectional regression on a number of land-cover distributions to derive weights, which are then applied in an iterative adjustment procedure to match estimated reference values such that the geographic and statistical data are consistent.

Beyond the cropland area, MIRCA, SPAM and GAEZ distinguish between rainfed and irrigated cultivations. All three models use GMIA version 4.0, released in 2007 (Siebert *et al.*, 2010), to identify the location and area intensity of irrigated production. Only SPAM and GAEZ, however, constrain potential crop distribution using biophysical and socio-economic suitability prior to allocating the harvested area of each crop. These two models further incorporate ancillary data such as road infrastructure, livestock density, population density and distance to market as a means of differentiating between production levels within cropping systems.

METHODS OF ANALYSIS

Each model produces spatially explicit cropping system maps on the same 5-arcmin grid; however, comparing these grids directly (pixel-wise comparison) may produce artificially inflated disagreement between products. While it is important to provide some indication of pixel-specific performance, as this is the format in which these models are often used, a pixel-specific approach implicitly assumes each pixel to be a result independent of any neighbouring pixels.

Each product was therefore assessed using methods that incorporate the spatial dimension of the data in biophysically and mathematically meaningful ways in addition to a pixel-specific comparison. To depict overall model agreement in raw output, pixel-wise maps of model consensus were produced for each crop at multiple harvested area thresholds. The sum of crops or cropland for each product was evaluated by latitude to account for the biophysical evolution of crops and cropland by growing regions while still allowing for methodological differences in crop distribution. As a means of focusing our analysis on regions of major production only, we masked areas with fewer than 50 hectares of harvested area (in a 5-min pixel whose size varies from about 5000 to 9000 hectares) prior to analysing the final products of each model in a spatially explicit manner. To compare the final distributions of each product with one another, a Gaussian filter with a kernel density of three standard deviations was applied to the results of each product prior to a pair-wise comparison. Although the raw model outputs, without filters or masks, and associated pairwise comparisons are available in Appendix S7 for reference, only the results of the Gaussian filtered analysis will be discussed here. Pre-processing the data using a Gaussian filter expands the analysis to incorporate neighbouring pixels, which allows for a more general analysis of spatial trends in each model. The kernel density for the Gaussian filters, i.e. the number of neighbouring pixels to consider, was chosen following a sensitivity analysis using a kernel density of one, two, three and four standard deviations. The results of the sensitivity analysis and full documentation of the implementation of the Gaussian filter are given in Appendix S2.

Both the by-latitude and the pixel-wise Gaussian filtered analyses were used to assess the cropland extent, the harvested area and the yield for each product. SPAM, M3 and MIRCA all rely on the Ramankutty cropland data, while GAEZ has developed its own cropland extent. The cropland extent delineates the domain to which each product's downscaling approaches are applied, and therefore represents an important intermediate step in the modelling process. Each model produced maps for harvested area, but only M3, GAEZ and SPAM produced maps of yield. Pairwise comparisons of each product were made accordingly.

Previous analyses have made use of spatially distributed estimates of harvested area or crop yield, but often use the chosen cropping systems model as a reference dataset without adequate attention to uncertainties in the model (Licker *et al.*, 2010; Neumann *et al.*, 2010; Deryng *et al.*, 2011; Foley *et al.*, 2011; Mueller *et al.*, 2012). To date, no analyses have explored the implications of model choice on final results. To illustrate the importance of continuing work aimed at characterizing the spatial distribution of crops and cropland, we calculated the yield gap using each model following the methods of Licker *et al.* (2010). This method entails using a set of climate zones to sample existing crop yields to derive an area-weighted distribution. The 90th percentile of this distribution is considered the potential yield for all areas within the same climate zone. The yield gap is then calculated as the difference between the potential and the actual yield. Although the potential yield could also be derived using crop model simulations, the method of Licker *et al.* (2010) provides two primary benefits: (1) it does not rely on a definition of cropland extent, which is a requisite condition for our analysis, and (2) as a relatively simple method, it may be applied at the global scale without extensive region-specific calibration data, as is required for many crop model simulations. For a relevant summary of existing methods used to calculate the yield gap, as well as the strengths and weaknesses of each, see van Ittersum *et al.* (2013).

Following the calculation of the yield gap using each cropping systems model, we computed the average yield gap $\left(\overline{YG}\right)$, the average difference in model-estimated yield gap (ΔYG) and the ratio of average model difference to average model estimate (the yield gap uncertainty ratio):

$$\overline{YG} = \frac{(\mathrm{M3_{yg}} + \mathrm{SPAM_{yg}} + \mathrm{GAEZ_{yg}})}{3}$$

$$\Delta YG = \frac{(|\mathrm{M3_{yg}} - \mathrm{SPAM_{yg}}| + |\mathrm{GAEZ_{yg}} - \mathrm{SPAM_{yg}}| + |\mathrm{M3_{yg}} - \mathrm{GAEZ_{yg}}|)}{3},$$

$$\text{uncertainty ratio} = \frac{\Delta YG}{\overline{YG}},$$

where the subscript yg denotes the model estimated yield gap calculated using each model. The quantity $\Delta YG/\overline{YG}$ would be of interest for anyone seeking to intensify food production. As the value of $\Delta YG/\overline{YG}$ approaches 1, the disagreement between models approaches the estimated yield gap; values greater than 1 indicate that the average model disagreement is greater than the average estimated yield gap.

As a means of differentiating between areas that have already achieved large yields and those yet to realize a significant proportion of their potential yield, we explore the relation between ΔYG and existing yields using the ratio $\Delta YG/\overline{Y}$, where $\overline{Y}$ is the average yield across all cropping systems models. This measure of ΔYG relative to achieved yields will complement both the existing absolute measure of ΔYG and the uncertainty ratio $\Delta YG/\overline{YG}$. Taken together, the three indices provide both absolute and relative information on the subsequent importance of model choice for calculating the yield gap. Of particular importance are areas that display high values over all indices, indicating that the estimated potential increase in yields depends heavily on the model chosen and contains a large degree of uncertainty both in an absolute sense and relative to existing production. These are areas where not only are the results of the yield gap analysis highly uncertain, but the implications for global food security are also relatively large.

The results of each analysis are described in following sections for wheat. The results for rice and maize are given in Appendix S5.

RESULTS

Cropland extent

Both Ramankutty *et al.* (2008) and GAEZ use the GLC2000 land-cover dataset as one input to the definition of cropland extent: Ramankutty *et al.* (2008) blends GLC2000 with remote-sensing observations from the BU-MODIS dataset while GAEZ supplements the GLC2000 data with independent information on the extent of protected areas, forests and agricultural extent (see Table 2) (FAO, 2001; Friedl *et al.*, 2002; Ramankutty *et al.*, 2008; WDPA, 2009). The methodological differences in delineating cropland extent result in distributions of cropland that broadly resemble one another but which differ significantly over certain regions. The cropland extent maps largely agree in Europe, southern Africa, East Africa and through much of China. But the products significantly disagree in the Great Plains of North America, West Africa, south-east Australia, India and south-east South America (see Fig. 1). These differences

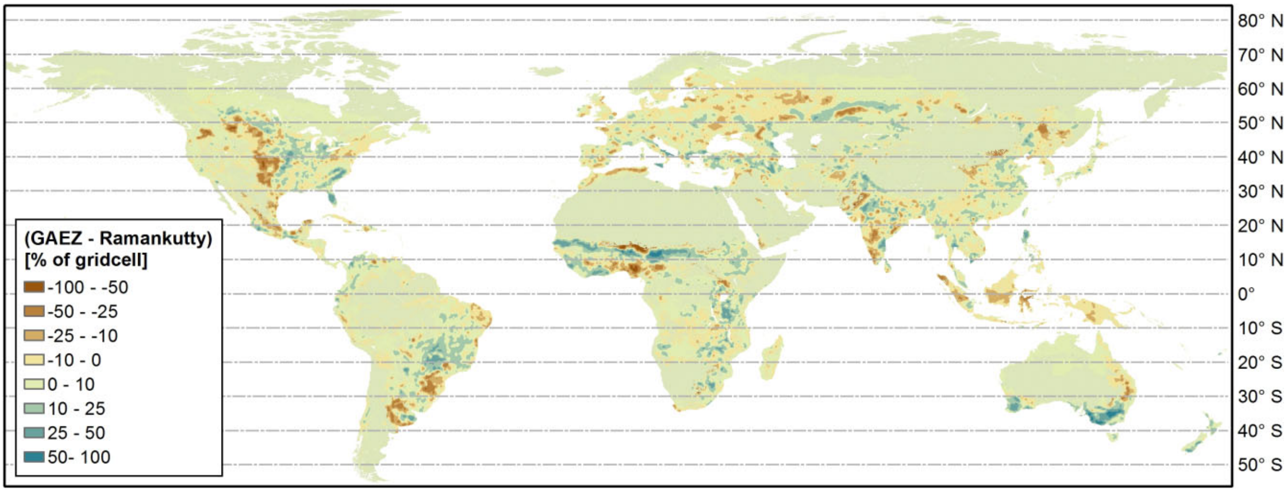

Figure 1 Differences between GAEZ and Ramankutty cropland extent after applying a Gaussian filter with a kernel density of three sigma.

propagate through each model to underpin the difference between models that use the Ramankutty *et al.* (2008) dataset and those using the GAEZ cropland.

Wheat: harvested area and yield

The four cropping system models displayed considerable dissimilarities in the spatial distribution of wheat across all thresholds of harvested area (see Fig. S1 in Appendix S5). The greatest extent of disagreement relates to areas with only minimal harvested area and is largely due to differences in model methodology. M3 and MIRCA, which tend to spread harvested area equally across all plausible croplands, account for the majority of disagreement in these low-intensity harvested areas (see Fig. S1 in Appendix S6). Agreement between models at low to intermediate thresholds is mixed but does not clearly reflect differences in cropland extent, nor are these inconsistencies dominated by any single model. When considering the threshold of highest intensity, SPAM differs significantly from the other three models over western Russia. This difference is probably due to both the downscaling methodology of SPAM and the collection of additional subnational statistics.

Despite the dissimilarities explained by divergence of individual model methodologies, the differences in the harvested area of wheat between GAEZ and the products that use the Ramankutty cropland extent (see Fig. 2) largely mirror the differences in cropland extent depicted in Fig. 1. Dissimilarities in the harvested area of wheat for Australia, the United States and Russia are nearly identical to those same differences in cropland extent. The distribution of differences in harvested area between any two models appears to be log-normally distributed, with the possible exception of the MIRCA–M3 comparison, which indicates that models are generally not systematically skewing or biasing crop statistics in their spatial disaggregation procedures for wheat (see inset histograms in Fig. 2). This is not always the case for maize and rice (see Appendix S6). Figure 2 reiterates the fact that discrepancies between the M3 and MIRCA products are minor, which is to be expected given that MIRCA uses M3 output directly as model input (see Table 2). The relatively larger inconsistencies between SPAM/M3 and SPAM/MIRCA – particularly in Europe, north India and coastal China – arise from differences in downscaling methodology and subnational data collection.

The estimated yields from M3, GAEZ and SPAM (MIRCA does not produce estimates of yields) matched less well than did the harvested areas from each product. GAEZ predicted higher yields over the majority of the domain (see Fig. 3), but particularly in the eastern United States, Europe, central/eastern China and Australia (see Fig. 4). While M3 and SPAM agreed to a greater extent, differences remained large over much of Europe and Asia. The disagreement between all models was by far the greatest in areas with minimal harvested area (see Fig. S2 in Appendix S7), although these areas may be considered less important to global food security analyses. Overall, the model discrepancies for wheat yield were comparable with those for the harvested area of wheat. However, the models showed less agreement on the yield analyses for both rice and maize when compared with the respective analyses of harvested area (see Appendices S6 & S7). This is evidenced in both the by-latitude comparisons and the histograms depicting the distribution of differences (see Figs S3, S4, S8 & S9, and associated histogram insets in Appendix S6). The observed differences between models on even the fundamental spatial patterns of harvested areas and yields serve as a reminder of the ongoing challenges to creating spatially explicit estimates of cropping system.

Discrepancies in the final products of cropping system models have implications for basic policy decisions relating to agricultural production and food security. Spatially explicit models influence our current estimates of yield, our estimate of potential yields and by extension our estimation of the yield gap. Figure 5 illustrates the model-dependent differences, and resulting uncertainty, in calculating the yield gap (panel a) using both an absolute measure (panel c) and with regard to existing yields

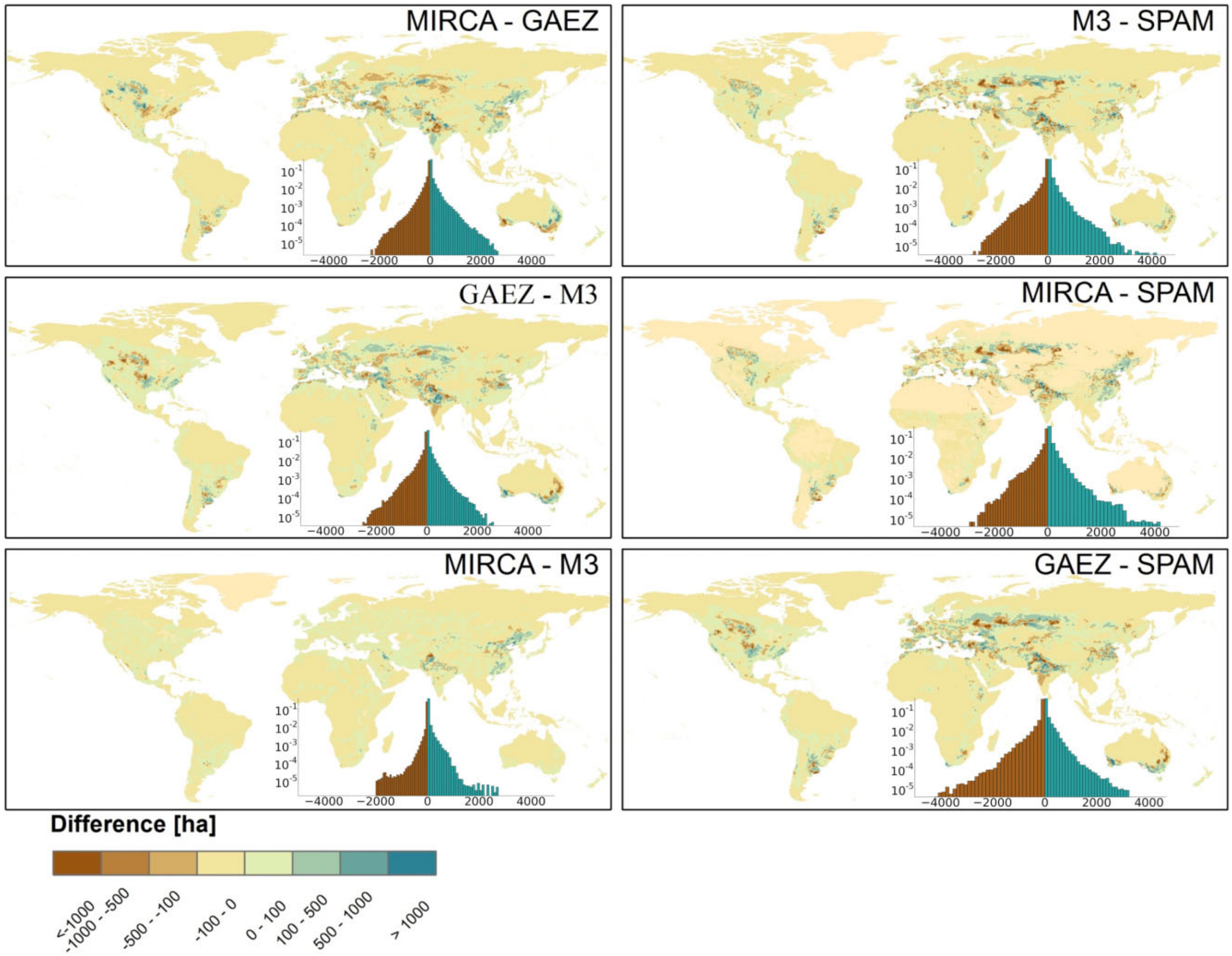

Figure 2 Comparison of wheat harvested area by model following a Gaussian filter of three-sigma kernel density. Histograms in each panel display the normalized percentage of pixels as a function of harvested area, y-axis (log scale) limits [0, 50%], x-axis limits [–5000, 500] ha.

(panel b). Areas in which the yield gap uncertainty ratio approaches 1 are areas in which uncertainty dominates the estimate of the yield gap. However, it is equally important to contextualize these uncertainties relative to existing yields and using an absolute measure of model difference. Areas displaying large values in all three panels indicate places in which the model-estimated yield gaps disagree (panel c), where this disagreement is a significant proportion of the estimated yield gap (panel a) and in which the differences are important in the context of existing food production systems (panel b).

The model-dependent uncertainty exceeds the estimated yield gap in significant parts of every continent, indicating that model choice is an important aspect of calculating the yield gap. While some of the large values in panel A of Fig. 5 arise owing to relatively high realized yields (and therefore small yield gaps), this is not universally the case, most notably not in Sudan, Nigeria, South Africa, Mexico, south-west India and northern China. In parts of Russia, Kazakhstan and Belarus the model-dependent uncertainty is high in an absolute sense and with regard to existing production, but is a less significant proportion of the overall yield gap (Fig. 5).

DISCUSSION

Calculation of the yield gap is only one example demonstrating the importance of improving spatially explicit estimations of harvested areas and yields. Other lines of research that will benefit from improved spatially explicit estimates include, but are not limited to, estimating the impact of climate change or extreme weather on agriculture (Nelson *et al.*, 2008), characterizing past and future anthropogenic components of land-cover change (Foley *et al.*, 2005; Pongratz *et al.*, 2008) and better understanding the historical evolution of global farming systems (Iizumi *et al.*, 2014).

This study represents a first step towards reconciling global, spatially explicit estimates of harvested areas and yields in that we quantify and demonstrate the subsequent importance of discrepancies in existing products. We first quantified model

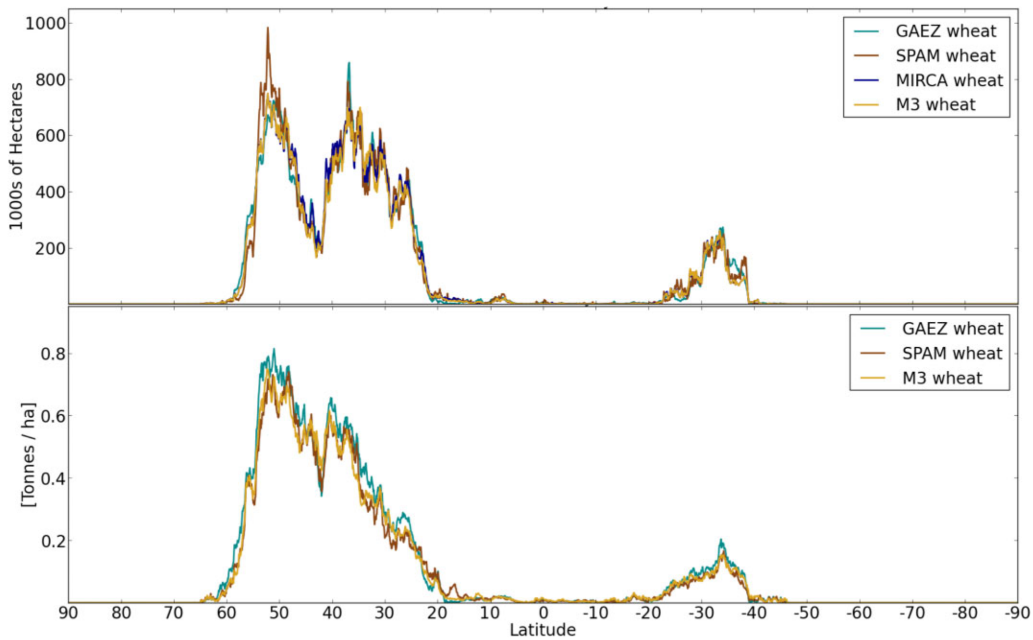

Figure 3 Sum of wheat harvested area and mean wheat yield by latitude.

consensus, then analysed each product by latitude and using a pair-wise Gaussian-filtered comparison and finally demonstrated the impact that model choice has on subsequent food security analyses. The results of our analysis indicate that the selection of a cropping systems model is non-trivial: differences between model-specific estimates of the yield gap are significant across nearly all climate zones and the average model difference exceeds the average estimated yield gap in select regions. While aspects of the disagreement may be attributed to model methodology, the choice of the dataset for cropland extent proved to be influential regardless of the downscaling process employed.

Despite interdependences between models, the input data used in each model vary substantially. These differences include fundamentally different subnational statistics used as a benchmark for each model, differences in the dataset on cropland extent and differences in the ancillary datasets used to provide supplementary information. The differences in input data are one reason why the four models produce substantially different results in many regions, and it is therefore imperative that potential users understand the extent and quality of input data used in each model.

Even in regions for which the four models rely on comparable input data, differences in downscaling methodologies resulted at times in large differences between products. Methodologies ranged from a philosophy of using minimal modelling to distribute national and subnational statistics – opting for ease of interpretation and a limit to requisite assumptions over modelling complexity – to one of including all available input datasets in a cross-entropy framework to account for both crop-specific biophysical and economic suitability. Our past experiences (e.g. You & Wood, 2006) demonstrate that while more input data and increasingly complex modelling don't necessarily lead to better or more accurate results, relying on only one input layer alone – either cropland, crop suitability or irrigated area – may not always be sufficient.

Because true crop distribution is unknown we are unable to evaluate the performance of any single model. Our analysis instead characterizes discrepancies among the four models and provides information for researchers and analysts to make knowledgeable decisions. Potential users of these cropping systems products need to understand the differences in input data and modelling prior to choosing the model that is best fitted to their purpose. If the study requires coverage of a large number of crops, potential uses may have to use M3 as it covers over 175 crops while the other three products include fewer than 30 major crops. On the other hand, if distinguishing the production system is critical for the study, MIRCA, SPAM and GAEZ may be used as M3 doesn't separate irrigated and rainfed productions. For the greatest detail on production systems, SPAM provides data on crop harvested area and yields disaggregated by system (irrigated, high-input rainfed, low-input rainfed and subsistence). Although GAEZ models detail production systems internally, they provide data to users at a scale similar to MIRCA: disaggregated into irrigated and rainfed only. In the studies which applied these modelling results, M3 and MIRCA have been used (at times together) for food security, climate change, water resource management and yield gap analyses (Licker *et al*., 2010; Siebert *et al*., 2010; Deryng *et al*., 2011; Foley *et al*., 2011; Mueller *et al*., 2012; Iizumi *et al*., 2014; Ponti *et al*., 2014). SPAM has been used for irrigation investment, nutrient management and climate change studies (Nelson *et al*., 2008; Liu *et al*., 2010; You *et al*., 2011). More recently, GAEZ has been used for analyses relating

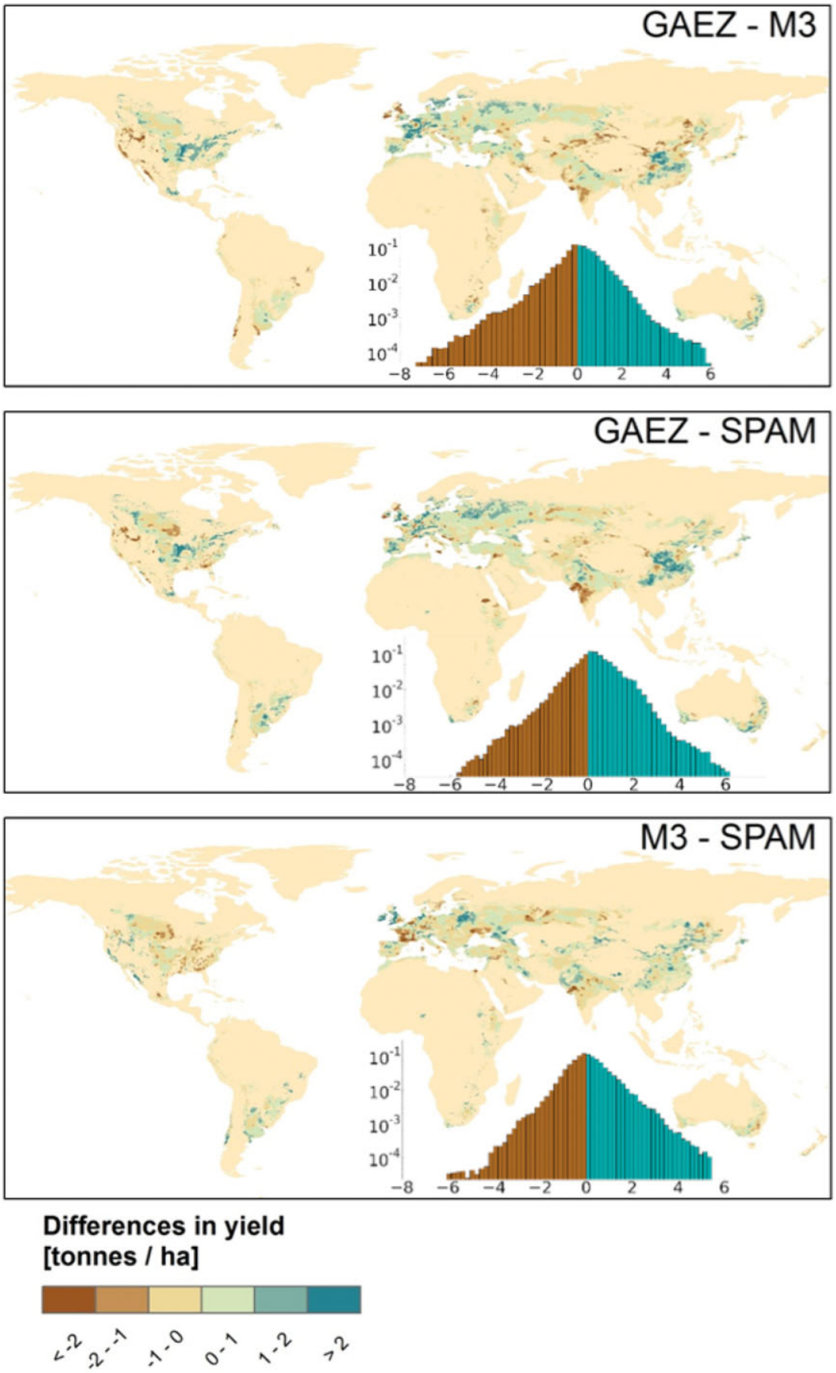

Figure 4 Comparison of wheat yield by model following a Gaussian filter of three-sigma kernel density. Histograms in each panel display the normalized percent of pixels as a function of yield, *y*-axis (log scale) limits [0, 35%], *x*-axis limits [–8, 6] tonnes ha^{-1}.

to climate change, ecology and food security (Teixeira *et al.*, 2011, 2013; Müller & Robertson, 2014; Ponti *et al.*, 2014). Looking to the future, there is certainly room for the four modelling teams to develop some community of practice. This has started in a recent workshop convened by International Food Policy Research Institute (IFPRI), but more needs to be done. Until we improve our fundamental understanding of where agricultural cropping systems exist, it will remain difficult to accurately characterize global food security.

ACKNOWLEDGEMENTS

The authors would like to thank Dr Van Velthuizen and Dr Fischer for their helpful correspondence regarding the GAEZ model; Dr Portmann and Dr Siebert for providing additional data from the MIRCA model; and Dr Licker for providing the climate zone dataset used to calculate the yield gap. This paper would not have been possible without their generous contributions. The authors would also like to thank two anonymous referees and Nathan Mueller for their detailed

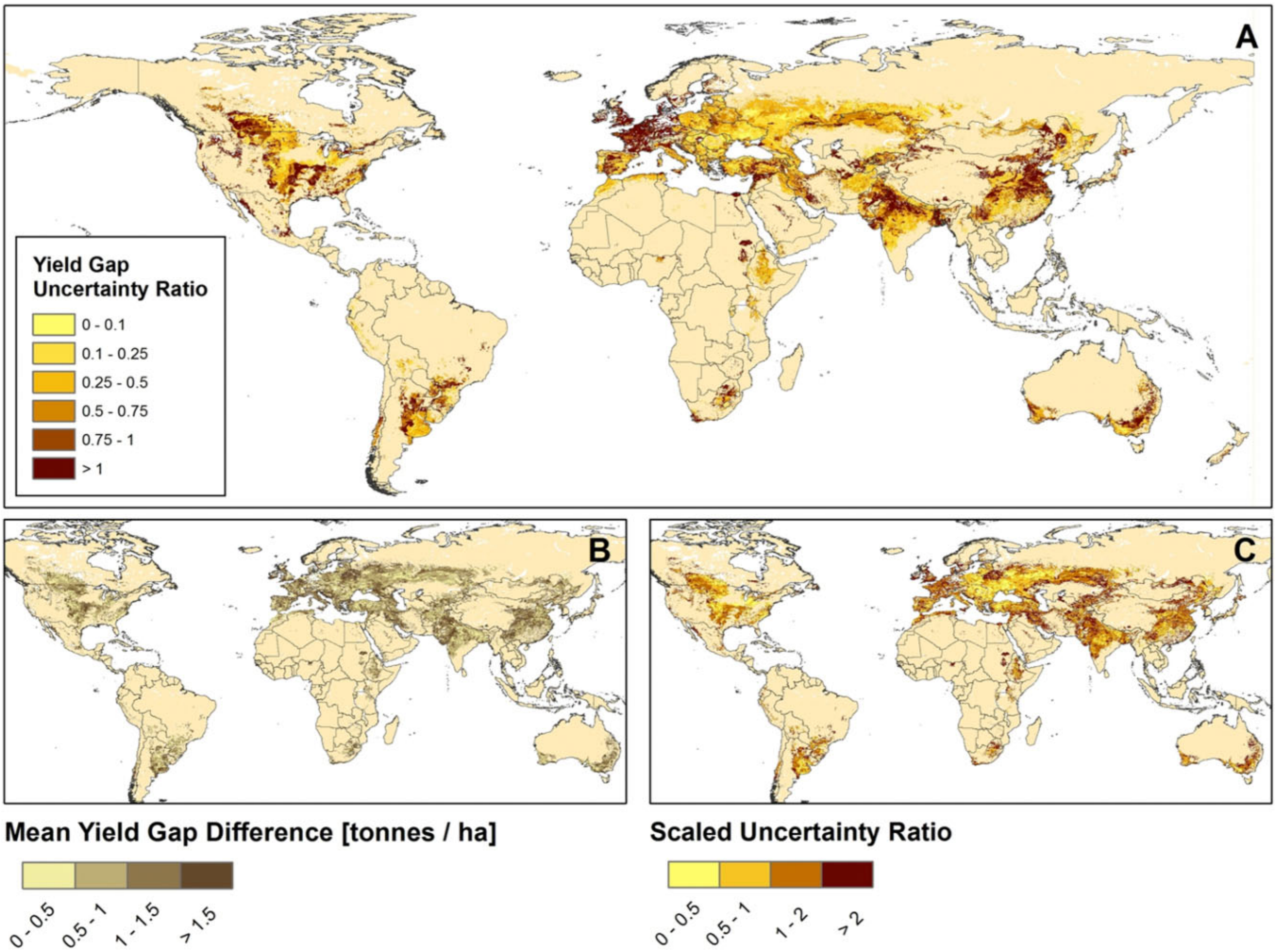

Figure 5 Implications of model differences for estimated wheat yield gaps. (a) Yield gap uncertainty ratio: average model difference divided by average estimated yield gap. (b) Average difference in estimated yield gap divided by existing yield. (c) Average difference in estimated yield gap.

and insightful comments. The work presented here is partly sponsored by National Basic Research Program of China (973 Program: No. 2010CB951500), the Bill and Melinda Gates Foundation (Harvest Choice Program), CGIAR Research Program on Policies, Institutions and Markets, US National Science Foundation (AGS-1048967). These supports are greatly appreciated.

REFERENCES

Bondeau, A., Smith, P.C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., Gerten, D., Lotze-Campen, H., Müller, C., Reichstein, M. & Smith, B. (2007) Modelling the role of agriculture for the 20th century global terrestrial carbon balance. *Global Change Biology*, **13**, 679–706.

Center for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Institute (IFPRI); and World Resources Institute (WRI) (2000) *Gridded population of the world (GPW)*, version 2. CIESIN, Columbia University, Palisades, NY. Available at: http://sedac.ciesin.columbia.edu/plue/gpw.

Center for International Earth Science Information Network (CIESIN) Columbia University; International Food Policy Research Institute (IFPRI), The World Bank; and Centro Internacional de Agricultura Tropical (CIAT) (2004) *Global rural–urban mapping project (GRUMP), alpha version*. Socioeconomic Data and Applications Center (SEDAC), Columbia University, Palisades, NY.

Deryng, D., Sacks, W.J., Barford, C.C. & Ramankutty, N. (2011) Simulating the effects of climate and agricultural management practices on global crop yield. *Global Biogeochemical Cycles*, **25**, GB2006, doi: 10.1029/2009GB003765.

FAO (2001) *Global forest resources assessment 2000*. Forestry paper no. 140. Food and Agriculture Organization of the United Nations, Rome, Italy.

Fischer, G., Nachtergaele, F.O., Prieler, S., Teixeira, E., Tóth, G., van Velthuizen, H., Verelst, L. & Wiberg, D. (2013) *Global agro-ecological zones (GAEZ)* version 3.0. Available at: http://gaez.fao.org/Main.html# (accessed October 2013).

Foley, J., Defries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E., Kucharik, C.J., Monfreda, C., Patz, J., Prentice, I.C., Ramankutty, N. & Snyder, P. (2005) Global consequences of land use. *Science*, **309**, 570–574.

Foley, J.A., Ramankutty, N., Brauman, K. *et al*. (2011) Solutions for a cultivated planet. *Nature*, **478**, 337–342.

Friedl, M., McIver, D., Hodges, J., Zhang, X., Muchoney, D., Strahler, A., Woodcock, C., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F. & Schaaf, C. (2002) Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, **83**, 287–302.

Fritz, S. & See, S. (2008) Identifying and quantifying uncertainty and spatial disagreement in the comparison of Global Land Cover for different applications. *Global Change Biology*, **14**, 1057–1075.

Fritz, S., See, L., McCallum, I., Schill, C., Obersteiner, M., van der Velde, M., Boettcher, H., Havlík, P. & Achard, F. (2011) Highlighting continued uncertainty in global land cover maps for the user community. *Environmental Research Letters*, **6**, 044005.

Giri, C., Zhu, Z. & Reed, B. (2005) A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. *Remote Sensing of Environment*, **94**, 123–132.

Golan, A., Judge, G. & Miller, D. (1996) *Maximum entropy econometrics: robust estimation with limited data*. John Wiley and Sons, New York.

Herold, M., Mayaux, P., Woodcock, C.E., Baccini, A. & Schmullius, C. (2008) Some challenges in global land cover mapping: an assessment of agreement and accuracy in existing 1 km datasets. *Remote Sensing of Environment*, **112**, 2538–2556.

Iizumi, T., Yokozawa, M., Sakurai, G., Travasso, M.I., Romanenkov, V., Oettli, P., Newby, T., Ishigooka, Y. & Furuya, J. (2014) Historical changes in global yields: major cereal and legume crops from 1982 to 2006. *Global Ecology and Biogeography*, **23**, 346–357.

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

Leff, B., Ramankutty, N. & Foley, J.A. (2004) Geographic distribution of major crops across the world. *Global Biogeochemical Cycles*, **18**, GB1009, doi: 10.1029/2003GB002108.

Lence, S.H. & Miller, D.J. (1998) Estimation of multi-output production functions with incomplete data : a generalised maximum entropy approach. *European Review of Agricultural Economics*, **25**, 188–209.

Licker, R., Johnston, M., Foley, J. a., Barford, C., Kucharik, C.J., Monfreda, C. & Ramankutty, N. (2010) Mind the gap: how do climate and agricultural management explain the "yield gap" of croplands around the world? *Global Ecology and Biogeography*, **19**, 769–782.

Liu, J., You, L., Amini, M., Obersteiner, M., Herrero, M., Zehnder, A.J.B. & Yang, H. (2010) A high-resolution assessment on global nitrogen flows in cropland. *Proceedings of the National Academy of Sciences USA*, **107**, 8035–8040.

Monfreda, C., Ramankutty, N. & Foley, J. (2008) Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles*, **22**, GB1022, doi: 10.1029/2007GB002947.

Mueller, N.D., Gerber, J.S., Johnston, M., Ray, D.K., Ramankutty, N. & Foley, J.A. (2012) Closing yield gaps through nutrient and water management. *Nature*, **490**, 254–257.

Müller, C. & Robertson, R.D. (2014) Projecting future crop productivity for global economic modeling. *Agricultural Economics*, **45**, 37–50.

Nelson, G.C., Rosegrant, M.W., Koo, J., Robertson, R., Sulser, T., Zhu, T., Ringler, C., Msangi, S., Palazzo, A., Batka, M., Magalhaes, M., Valmonte-santos, R., Ewing, M. & Lee, D. (2008) *Climate change: impact on agriculture and costs of adaptation*. IFPRI Policy Report. International Food Policy Research Institute, Washington, DC.

Neumann, K., Verburg, P.H., Stehfest, E. & Müller, C. (2010) The yield gap of global grain production: a spatial analysis. *Agricultural Systems*, **103**, 316–326.

Pongratz, J., Reick, C., Raddatz, T. & Claussen, M. (2008) A reconstruction of global agricultural areas and land cover for the last millennium. *Global Biogeochemical Cycles*, **22**, GB3018, doi: 10.1029/2007GB003153.

Ponti, L., Gutierrez, A.P., Ruti, P.M. & Dell'Aquila, A. (2014) Fine-scale ecological and economic assessment of climate change on olive in the Mediterranean Basin reveals winners and losers. *Proceedings of the National Academy of Sciences USA*, **111**, 5598–5603.

Portmann, F.T., Siebert, S. & Döll, P. (2010) MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: a new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles*, **24**, GB1011, doi: 10.1029/2008GB003435.

Ramankutty, N., Delire, C. & Snyder, P. (2006) Feedbacks between agriculture and climate: an illustration of the potential unintended consequences of human land use activities. *Global and Planetary Change*, **54**, 79–93.

Ramankutty, N., Evan, A.T., Monfreda, C. & Foley, J. (2008) Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles*, **22**, GB1003, doi: 10.1029/2007GB002952.

Rosegrant, M.W., Cai, X. & Cline, S. (2002) *World water and food to 2025: dealing with scarcity*. International Food Policy Research Institute, Washington, DC.

Siebert, S., Portmann, F.T. & Döll, P. (2010) Global patterns of cropland use intensity. *Remote Sensing*, **2**, 1625–1643.

Teixeira, E., Fischer, G., van Velthuizen, H., van Dingenen, R., Dentener, F., Mills, G., Walter, C. & Ewert, F. (2011) Limited potential of crop management for mitigating surface ozone impacts on global food supply. *Atmospheric Environment*, **45**, 2569–2576.

Teixeira, E.I., Fischer, G., van Velthuizen, H., Walter, C. & Ewert, F. (2013) Global hot-spots of heat stress on agricultural crops due to climate change. *Agricultural and Forest Meteorology*, **170**, 206–215.

Voldoire, A., Eickhout, B. & Schaeffer, M. (2007) Climate simulation of the 21st century with interactive land-use changes. *Climate Dynamics*, **29**, 177–193.

WDPA (2009) *World database of protected areas, annual release 2009*. Available at: http://www.wdpa.org/AnnualRelease.aspx.

Yang, D., Kanae, S., Oki, T., Koike, T. & Musiake, K. (2003) Global potential soil erosion with reference to land use and climate changes. *Hydrological Processes*, **17**, 2913–2928.

You, L. & Wood, S. (2006) An entropy approach to spatial disaggregation of agricultural production. *Agricultural Systems*, **90**, 329–347.

You, L., Wood, S. & Wood-Sichra, U. (2009) Generating plausible crop distribution maps for sub-Saharan Africa using a spatially disaggregated data fusion and optimization approach. *Agricultural Systems*, **99**, 126–140.

You, L., Ringler, C., Wood-Sichra, U., Robertson, R., Wood, S., Zhu, T., Nelson, G., Guo, Z. & Sun, Y. (2011) What is the irrigation potential for Africa? A combined biophysical and socioeconomic approach. *Food Policy*, **36**, 770–782.

You, L., Wood, S., Wood-Sichra, U. & Wu, W. (2014) Generating global crop distribution maps: from census to grid. *Agricultural Systems*, **127**, 53–60.

Zhang, X. & Fan, S. (2001) Estimating crop-specific production technologies in Chinese agriculture: a generalized maximum entropy approach. *American Journal of Agricultural Economics*, **83**, 378–388.

Additional references to the data used in this study can be found at the end of the Supporting Information at [URL].

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Appendix S1 Model methodology.
Appendix S2 Comparison of the downscaling methodologies.
Appendix S3 Gaussian filter sensitivity analysis.
Appendix S4 Cropland extent supplementary material.
Appendix S5 Supplementary figures for wheat.
Appendix S6 Supplementary analysis for rice and maize.
Appendix S7 Supplementary pixel-wise figures for wheat, rice and maize.

BIOSKETCH

The authors' research focuses on modelling agricultural land-use systems, in particular on the evolution and drivers of change in agricultural systems, as well as on estimating the impact of climate change and climate variability on food security.

Editor: Martin Sykes